

“炫技”后人形机器人何时去干活

▶ 本报记者 刘琴

6月6日,在北京智源大会“智源具身智能会客厅”环节,多位机器人企业负责人围绕人形机器人赛事、产业化落地、VLA(视觉—语言—动作)模型应用及其带来的挑战等话题进行了讨论。

机器人赛事是技术“验证场”

宇树科技 CEO 王兴兴表示,现在人形机器人还不能直接去家里干活,这是现在全球范围内机器人行业都面临的问题。宇树科技的终极目标,是人形机器人可以替代人类从事辛苦的工作。在实现这个终极目标前,对于企业已经取得的技术成果,希望能通过赛事和表演进行展示和普及,以实现一定商业化的宣传和推动作用。

在北京人形机器人创新中心总经理熊友军看来,机器人赛事一方面是知识普及的盛会,让大众对机器人现在发展状况及技术水平有比较全面的了解;另一方面,它也是很好的机器人训练场,对提升机器人的技术有很大帮助。此外,这也是机器人潜在客户了解和与机器人企业沟通的桥梁。通过比赛,会促成机器人加速产业化,在真实场景里做试点应用。

“我觉得具身智能当下的目标就是一定要推动产业化。虽然看到了具身智能炫酷的技能,但一定要反思这样的技能在新的环境里,在面对新的物品时,在用户非常‘挑剔’、需要成功率接近百分之百的情况下,它能不能成功。”北京大学助理教授、北京银河通用机器人有限公司创始人兼 CTO 王鹤表示,希望与后续赛事协同在操作层面对接起来,用赛事引领有价值、可落地的技能形成,不断地把具身智能的生态丰富起来。

上海交通大学教授、上海穹彻智能科技有限公司联合创始人卢策吾表示,人形机器人赛事很好地展示了机器人的性能,下一步期待举办机器人劳动技能比赛。

未来机器人形态呈多样化

在具身智能大潮下,机器人一定要做人形吗?这成为业内讨论的热门话题。

王兴兴表示,对于宇树科技来说,并不坚持要做人形机器人,企业也做机器狗。他认为,随着通用人工智能(AGI)的发展,机器人的形态将变得更加多样化。未来,工厂、医疗等

领域的机器人形态将远比现在丰富,甚至可能增加百倍。

为什么现在大量具身智能企业在做人形机器人?在王兴兴看来,是因为现在 AI 基于人的数据进行采集和训练,尤其是上半身与人的动作一致,数据采集方便很多,也方便 AI 做训练。对于机器人全身也是一样,比如跳舞、格斗比赛,如果将机器人做成别的样式,就没办法完成相应动作。

熊友军表示,具身智能不一定要局限于人形,但人形是一个比较好的载体。他认为,从市场应用情况看,具身智能市场最多的是走入商用和家庭,现在工厂场景只是一个开始,人形机器人市场的容量比其他机器人的容量要大。从人机交互角度看,未来人形机器人更像家庭伙伴、朋友,更容易被人接受。

“现在的环境大部分是为人设计的,部署人形机器人及其适应环境过程非常便利,人形机器人可以方便地使用人类的工具,不需要做太多的改造。虽然现在人形机器人的成本、技术不占优势,但是放在更长的时间范畴看,人形机器人会是一个很好的载体。”熊友军说。

VLA 模型仍存在局限性

2025 年以来,一种全新技术模式即 VLA(视觉—语言—动作)模型正在加速崛起。在无人驾驶领域,VLA 模型已经成为主流解决方案。据了解,VLA 模型整合视觉理解、语言推理与端到端控制于单一模型,使机器人能在复杂环境中推理和行动。

VLA 模型泛化性如何?王鹤表示,自动驾驶领域已经验证了端到端方案的优越性,VLA 模型的核心优势在于能够直接从视觉观测和自然语言指令中端到端地输出动作,无需中间产物。王鹤认为,人类除视觉的输入之外,还有味觉、触觉、嗅觉、听觉以及对温度的感知。所以 VLA 模型只能是一个起点,如果想达到人类级别的具身智能,还需要不断融入新模态。

VLA 模型现在最适合做什么呢?王鹤认为,主要是移动、抓取和放置。这几个技能主要依赖视觉,辅以触觉或力觉传感器,能在工业、商业非常广泛地应用。如果能够将 VLA 模型在工业、商业、服务等场景中部署成功,将标志着具身智能技

术真正地走向了实用化。

在卢策吾看来,VLA 模型的扩展性虽然强大,但仍存在局限性。具身智能的发展需要逐步落地,同时也要有长远目标即推动通用智能的发展。相比无人车,机器人的任务和交

互复杂度高,通用空间很大。VLA 模型需要不断加入新的模态信息,不断压缩它的不确定性,如力反馈和物理世界理解,以增强模型的鲁棒性和适应性,从而推动 VLA 模型逐步走向通用智能。

Physical Intelligence 联合创始人兼 CEO Karol Hausman 表示,VLA 模型最大的瓶颈在于算法,他认为,现在数据是海量的,需要提升和改善 VLA 模型的算法以运用这些无尽的数据。



在“具身智能体验区”,人形机器人为观众倒橙汁饮料。

刘琴/摄

人形机器人各显神通展“绝活”

▶ 本报记者 刘琴

表演格斗、整理桌面、倒饮料、拿取食品……在6月6-7日举行的北京智源大会期间,多款人形机器人向观众展现各种“绝活”,让更多人见识到了人形机器人的“十八般武艺”。

在大会“智源具身智能会客厅”一开场,宇树科技 G1 人形机器人即在现场“大展拳脚”,刺拳、勾拳、踢腿、闪避……一套流畅的搏击动作吸引了全场目光。

据悉,该款人形机器人在不久前举行的全球首个以人形机器人为参赛主体的格斗竞技赛事中获得冠军。“过去无论是春晚上的秧歌表演还是格斗比赛,我们希望给大家真正展示一下当前全球机器人发展的情况,并且能带动整个机器人行业的发展。”宇树科技 CEO 王兴兴说。

接下来,在全球首个人形机器人半程马拉松比赛获得冠军的选手“天工”人形机器人,在现场展示了升级后的上肢能力——“天工”2.0 完成了“水果摆盘”的精细化服务演示。

“‘天工’2.0 升级在于上肢,让机器人有了更多的自由度,比如灵巧手、负重能力。”北京人形机器人创新中心总经理熊友军介

绍说。

在现场,北京银河通用机器人有限公司研发的人形机器人 Galbot 展示了其在商业货架零售场景中的最新应用。“我中午又饿又热,Galbot 给我搞点吃的和喝的。”当北京大学助理教授、银河通用创始人兼 CTO 王鹤向机器人发出指令后,机器人 Galbot 回答道:“根据您提到的,为您推荐安慕希酸奶和果粒爽果冻。”随即在会场的商品摆架上完成了拿取任务。

据了解,人形机器人 Galbot 已经能够识别并抓取袋装薯片、桶装果冻等多种形态的商品,展现出更高的泛化能力。王鹤介绍,该款人形机器人展是一款轮式底盘机器人,基于自研的端到端大模型,能让机器人在琳琅满目货架上精准抓取相应货品。

此外,王鹤介绍说,银河通用研发的人形机器人已经应用在北京 7 家无人药店,24 小时由机器人在里面捡药并对接骑手。今年年底前,该企业计划在北京、上海两市开设 100 家这样的无人药店。

在北京智源大会智源 AI 成果体验区的“具身智能体验区”,

多款人形机器人同样展现了智能收纳、做汉堡等超能力。

“请给我接一杯橙汁。”在一位观众发出指令后,只见人形机器人从容地拿起桌面上一个玻璃杯,放到橙汁盛放器下面,按动开关,接了一杯橙汁送到观众面前。

此展位展出的人形机器人来自北京智源研究院,该研究院具身多模态大模型研究员赵君凯接受本报采访时介绍道,研究院通过使用模仿学习算法,设计数据采集方案,采集人类操作数据并训练人形机器人上半身灵巧手操作技能,机器人通过头部相机可实时感知周围环境,当顾客语音输入想要的饮料种类后,机器人 大脑 robobrain 理解用户需求并调用操作技能,自主完成操作任务。

当前,人形机器人正一步步从科幻走向现实,从技术走向市场。“随着大模型能力提升,人形机器人会更加通用,结构上更加灵巧,具备自主完成复杂任务能力,进入工厂帮助人类完成危险岗位的工作,进入家庭帮助人们完成生活任务,比如房间整理,厨房做饭等,更好地服务人类生产生活。”赵君凯说。