

7月9日,四川绵阳某工厂车间。一款名为“远征A2-W”的机器人正在流畅地搬运周转箱。它们灵活穿梭于货架间,稳当地将散落各处的周转箱搬运到指定位置。

更令人惊叹的是,当移动叉车突然驶过,或是工人临时走近作业区,他们还能灵巧避让。整个搬运过程中,机器人的动作自然顺畅,仿佛真的“看”懂了周围的工作环境,也

“想”明白了怎么执行每个搬运步骤。

这种能让机器感知环境、理解任务并自主动作的能力,就是具身智能(Embodied AI)。如今,方兴未艾的具身智能机器人,正逐渐让机器从冰冷的程序执行者,蜕变成能在真实世界自主行动的“智能伙伴”。

具身智能机器人——

当AI拥有了“身体”之后

■郭宁赵越董献洲

高技术前沿

能“看”能“想”能“动”

对于具身智能机器人的概念,大众或许仍感陌生。不妨设想一下这样一个场景——

你给家里的机器人下达指令:“把茶几上的空水杯拿到厨房去。”如果机器人只能沿着预设的路线移动,茶几被挪动了一点或者地上多了一个玩具,它就会迷路或者被玩具绊倒,就不能算是具身智能机器人。

面对同样的情况,具身智能机器人的反应是这样的——

它先用“眼睛”看,通过摄像头或其他“感官”扫描整个客厅,认出茶几,找到空水杯,还能定位厨房的位置。

再用“脑子”想,结合“拿”“空水杯”“厨房”的指令,它理解总结出“这是一个移动物体的任务”,随后进一步“思考”;如何抓握水杯才最稳当安全?如何规划避开地上物品的路线?甚至进行初步“推理”:主人让我把空水杯拿到厨房,很可能是要清洗,那么最好放到水槽旁。

最后“身体”动,它的机械臂会稳稳地抓起水杯,然后沿着规划的路线走向厨房,准确把杯子放下。

这个“感知—思考—行动”过程的闭环,正是具身智能机器人的核心功能:它将人工智能的“智慧”与机器实体的“动作”深度融合,实现信息世界和物理世界的互联。

支撑这一功能的,是众多智能技术的协同发展:让机器能看、听、触的多模态感知技术,理解重力、摩擦力等物理规律的世界模型技术,在动态环境中快速响应的实时决策与推理技术,以及让机器人既能认路又能灵巧操作的自主导航和精细控制技术等。

此外,还有将上述技术组合在一起的“慢慢搭配”智能系统架构。

以“擦黑板”为例。“慢系统”如同深谋远虑的军师,负责拆解复杂命令:找到抹布,走到黑板前,规划擦拭的动作顺序。

“快系统”则像一名身手矫健的运动员,一旦“慢系统”定下大方向,“快系统”就能根据实时看到的黑板情况(哪里脏、哪里湿),精细地调整机械臂的微小动作和力度,确保擦得干净又不会划伤黑板。

这种“深谋远虑”加“眼疾手快”的组合,本质上是对人类大脑的慢速推理(前额叶皮层)和快速直觉(基底神经节)两种处理机制的模仿,让机器人在面对动态环境时更加游刃有余。

具身智能的精髓,还在于“做中学”。这就像一个蹒跚学步的孩子,跌倒了无数次才学会平衡。人类智力的增长依赖于身体与环境的互动,具身智能机器人同样需要在真实或模拟的世界里反复“练习”。

谷歌公司人工智能研究实验室



7月26日,2025世界人工智能大会在上海举行。图为展览现场参观者与具身智能机器人互动。

新华社发

DeepMind的Genie模型、英伟达公司的Issac Sim平台,以及我国人形机器人创新中心今年3月发布的“格物”平台,就是供具身智能机器人训练学习的“虚拟世界”。以倒水训练为例,在这个高保真模拟环境中,机器人可以反复观看人类操作视频,并进行成千上万次的模拟练习,不断调整抓壶力度、倒水角度和移动速度。

最终,当它在现实世界拿起水壶时,就能像经验丰富的服务员一样,把水一滴不洒地倒入杯中。这种模拟训练大幅降低了试错成本,加快了机器人习得技能的速度,是具身智能机器人走向实用化的关键一环。

从“笨拙”走向“灵巧”

事实上,具身智能机器人的发展过程是从笨拙走向灵巧的进化史,其核心是解决“机器如何理解物理世界”的终极难题。

1950年,人工智能之父艾伦·图灵在其论文《计算机器与智能》中首次提出关于具身智能的设想——真正的智能需要拥有物理实体,通过与世界的互动来学习和进化。

70多年后的今天,这一理念正通过具身智能机器人技术变为现实,将机器人从“执行工具”升级为“认知主体”。

其实,真正意义上关于具身智能技术的探索时间并不算长。

2018年,谷歌开始训练机器人学习抓取积木,但成功率不足30%,机械臂

常把方块撞飞。

2022年,谷歌公司研发的RT-1模型与之前相比,有了明显的进步。机器人能理解诸如“把可乐放进冰箱”的多步骤指令,不过,当时的机器人如同需要手把手教的婴儿,想学会一个动作,需要为其反复演示上千次。

时间来到2024至2025年,具身智能机器人的发展迎来爆发期。例如,中国智平方公司推出RoboMamba模型,将机器人学习新动作的效率提升了数倍。以前研发人员需要大量演示才能让机械臂学会抓取新物体,现在只需微调参数就能使其快速掌握这项技能。

同期,开源社区也有了更大突破。美国斯坦福大学、加州大学伯克利分校、谷歌公司等联合推出OpenVLA模型,以不到主流模型七分之一的参数实现了更优性能。这意味着普通的研究者甚至学生,用消费级的显卡即可训练开发机器人算法,这极大促进了具身智能机器人的创新与普及。

今年,具身智能机器人迎来了“走出实验室”的关键节点:2月,武汉光谷“天问”机器人进入咖啡店“实习”,甚至可以制作咖啡拉花;而在美国,护士机器人Moxi已经在各地医院完成超过100万次医药配送;5月,美国特斯拉公司Optimus机器人开始在自有工厂“打工”,执行电池分装任务;7月,中国铁路设计集团研发的四足机器人,实现了对数据中心机房设备的高效自主巡检。

近日,两家具身智能机器人4S店分别落户北京和杭州。用户可以在店内体验、选购适合自身需求的“钢铁伙伴”,享受后续技能定制、软件更新和硬件维护服务。

从工厂车间里精准搬运的机械臂,到机房里自主巡检的机器狗,具身智能机器人正慢慢融入人类生产生活的多个方面。可以说,具身智能机器人不仅是机械躯体的进化,更是智能认知能力在物理世界的延伸。

化身人类“钢铁伙伴”

具身智能机器人正悄然卸下神秘面纱,化身我们身边的“钢铁伙伴”。工厂轰鸣的车间、温馨的居家角落,危机四伏的救援现场,甚至未来的无人战场,都将是它们可以大展拳脚的地方。

——工厂里的“超级工人”。训练后的机器人熟练度高,不会疲劳,未来可以替代人类从事繁重重复的搬运、组装等工作。

据统计,工厂车间里的“远征A2-W”机器人将搬运效率提升近30%,失误率却降至0.1%以下。这标志着具身智能机器人正开始成为切实提高生产力的“超级工人”。

——家庭中的“贴心帮手”。“奶奶,您该吃药了。”在深圳一家养老院,一台服务机器人来到老人身边,一边温和地发出语音提醒,一边稳稳地递上水杯和药盒。

虽然目前这些机器“保姆”还很稚嫩,但随着具身智能技术的升级,它们未来将学会整理房间、取送物品、陪伴聊天,甚至搀扶老人上下楼,并在紧急情况下自主呼叫医护人员。

——危险地带的“无畏先锋”。例

如,杭州某公司推出的山猫M20轮式机器人,能轻松应对坎坷山路、泥泞湿地和废墟障碍,适应-20℃至55℃的环境。这类机器人可以替代人类完成很多高危作业,成为沙漠油田、戈壁电站、火灾现场等极端环境中的冲锋勇士。

展望未来,未来机器人的形态可能像瑞士军刀一样,能进行多功能灵活组合。例如,根据任务需要,给强大的智能“大脑”配上不同的“身体”部件:需要精细操作,就配上灵巧的“手臂”;需要和人类自然互动,就配上类人的形态;需要征服野外崎岖地形,就配上多足或轮腿结合的“脚”……

具身智能机器人的“头脑”也会越来越聪明。它们会感知到更多的环境因素,理解更多的物理规律和人类常识,完成更加复杂的任务,甚至能预测接下来会发生什么,需要怎么行动来提供帮助,与人类的互动也会更加自然顺畅。

更重要的是,技术的开放共享和高效率的虚拟训练,将大幅降低其开发和训练成本。专家预测,在不远的将来,一些实用的家庭服务机器人价格或将大大降低,从而走进千家万户。

我们正见证一个人机关系重构的时代,具身智能机器人不会取代人类,而是拓展人类的能力边界——工人将从危险枯燥的流水线解放,士兵不再需要亲入雷区,老人可获得24小时照护……

具身智能技术的终极目标并非创造“替代人类的机器”,而是建立人机共生的新生态。当机器人能理解“请小心拿放,这是爷爷最爱的茶杯”背后的情感价值,能在火灾现场自主判断“优先抢救儿童卧室”的道德选择,技术才真正拥有了温度与智慧。

的作战人员获取实时战场影像的同时,可通过脑机接口实现对无人机集群的免手操控。据美国防务新闻报道,澳大利亚国防创新中心与悉尼科技大学将微软AR设备“HoloLens”和脑机接口技术结合,用以操控车辆和机器人,未来或将用于突破战场上对方的拦截。

目前,脑机接口技术已开始应用于增强认知感知、精准操控智能装备与战场医疗革新等领域,对战场进行有效赋能。尽管发展仍面临多重现实挑战,但随着一项项技术突破,我们有理由相信,脑机接口技术通过以“人机联网”为特征的战场智能网络,可实现指挥员、战斗员与辅助决策系统、无人化作战平台之间的意识指令控制和智能自主协同,或将在未来战场上发挥更大作用。

论见

科技云

科技连着你我他

■本期观察:强裕功 王红锦 何昌明

近年来,电动垂直起降飞行器(eVTOL)逐渐走进人们的视野。本期,我们共同关注3种不同整机构型的电动垂直起降飞行器。

多旋翼eVTOL



多旋翼eVTOL主要依靠多个独立旋翼协同运作,实现垂直起降与飞行,常见构型包括4旋翼、6旋翼和8旋翼。

工作时,电机驱动旋翼高速旋转,通过调节各旋翼转速差,控制多旋翼eVTOL的升力、姿态和方向。

据悉,这种构型的电动垂直起降飞行器安全性较好,即便某些旋翼出现故障,剩余旋翼仍能保证飞行器安全降落。以德国Volocopter公司研发的VoloCity电动垂直起降飞行器为例。该机型最大航程35公里,巡航速度110公里/小时,采用可快速更换的电池模组,只需5分钟就能完成能源补给。

凭借出色的垂直起降和悬停能力,未来多旋翼eVTOL将主要活跃于低空短途场景。例如,多旋翼eVTOL应用于城市里,既可以承担通勤人员跨区域短途出行,减少拥堵问题,也可融入物流网,完成小件物品的即时配送。此外,多旋翼eVTOL如果应用于低空旅游领域,还能为游客提供独特的观赏视角。

复合翼eVTOL



复合翼eVTOL融合了固定翼飞机与旋翼飞行器的优势,采用“垂直起降+水平巡航”双模式设计方式——垂直起降阶段,分布在机身四周的旋翼为飞行器提供升力;进入水平飞行状态后,旋翼停止工作,此时由固定翼产生升力,飞行器切换至类似传统飞机的飞行模式。

相比多旋翼eVTOL,复合翼eVTOL的巡航速度和航程均有所提升,更适配中长途场景,既能承担城际或偏远地区的物资运输任务,弥补地面物流的短板,也可以搭载设备开展大范围测绘和农林巡检工作。例如,国内某航空公司的V2000CG电动垂直起降飞行器,最大航程250公里,巡航速度达200公里/小时,是全球首款2吨级纯电动无人驾驶eVTOL,其运载能力等同小型直升机。如果遇到灾情等紧急情况,这种复合翼eVTOL还能帮助工作人员远距离投送救援物资、长时间监测灾情等。

倾转旋翼eVTOL



倾转旋翼eVTOL的核心技术,在于其旋翼系统能够在水平与垂直方向之间灵活倾转。

垂直起降时,其旋翼保持垂直状态,提供向上升力,使eVTOL具备直升机的起降性能;当eVTOL水平飞行时,旋翼倾转至水平方向,转变为类似螺旋桨飞机的拉力模式,从而实现eVTOL的高速巡航。

凭借出色的垂直起降与高速巡航能力,倾转旋翼eVTOL或将成为军民两用的热门产品。在民用领域,倾转旋翼eVTOL可应用于城市间快速运输、跨区域应急救援等;在军用领域,其可执行战术运输、特种作战、侦察警戒及医疗后送等多样化任务。

美国V-22“鱼鹰”倾斜旋翼机,是此类飞行器的典型代表之一。该飞行器巡航速度446公里/小时,航程超1600公里,最大起飞重量27.4吨,可搭载24名士兵或9吨物资。该飞行器具备远程高速灵活的垂直投送能力,适配多种任务场景。