

从0到1造一个人形机器人 需要哪些关键技术?

“你好,我是青龙人形智能机器人,能帮你做家务。”

“你看桌面上有什么?”

“我看看。桌面上有三个面包、两个水果。”

“你帮我清理一下吧。”

“好的,正在为您整理中,我先把面包和水果分类摆放。”

说话间,身高1米85、体重80公斤的“青龙”看向桌面,依次拿起面包和水果,分类放入筐中。现场展示了行走、对话、做家务等能力。

在2024世界人工智能大会上,全球首款通用人形机器人开源版机“青龙”惊艳亮相。“我们看到,人形机器人可以通过语言信息理解的意图,根据现场环境做出综合判断并执行任务。”现场工作人员介绍,这个任务看似简单,实际上体现了“大小脑”的感知、规划、决策以及控制能力,代表了国内人形机器人技术的领先水平。

“青龙”由国家地方共建人形机器人创新中心(以下简称中心)打造,该中心于今年5月,由工业和信息化部 and 上海市人民政府共同授牌,落户浦东。中心科研团队硕博占比约80%,研发团队长期从事人形机器人领域技术与软硬件研发。

拥有43个自由度的“青龙”,代表了目前人形机器人行业比较顶尖的硬件设计水平。“青龙”不仅在硬件参数上达到了国际领先水平,更在具身智能技术的集成与应用上展现出了特有的优势。“青龙”包含人形机器人平台技术、具身智能、数据集和智能训练场四大技术板块,我们已经进行了开源,希望更多人参与到人形机器人的技术创新中来。”中心首席科学家江磊说。

近日,《瞭望》新闻周刊记者走近“青龙”,与科研团队对话,看如何打造和训练一个“优秀”的人形机器人。

机械躯体:人形机器人的“骨肉”

“平台技术可以简单理解为机械躯体,是人形机器人的基础,包含行走与驱动系统、操纵与作业系统、感知与控制系统三大模块。”中心机器人平台技术负责人梁振杰介绍说。

本刊记者在现场看到,“青龙”一步步走上讲台,然后停下来向观众挥手打招呼,它走路的速度不算快,但步伐足够稳健,躯体结构与人体类似。

梁振杰介绍,“青龙”全身集成了43个主动自由度,实现从头部到手部、臂部、腿部、腰部和踝部的全尺寸设计。关节模组是躯体结构的核心组成单元,它一共搭载了10种、31个关节,最大关节扭矩达到了396Nm,峰值扭矩密度实现了200Nm/kg,用以实现高强度动力输出。

双足负责行走,双臂执行作业。梁振杰说,“青龙”的腿部系统追求轻量化、高刚度和低惯量,搭载了高扭矩密度的轴向电机,以此保障它在复杂地形中的稳态行走能力。其上肢配备了7自由度的机械臂与集成触觉感知的五指灵巧手,为完成精细操作和复杂任务提供了硬件基础。

“青龙”搭载了算力可达400TOPS的控制器和丰富的外部接口,用以满足人形机器人产品以及常规外部设备的使用要求。400TOPS意味着控制器每秒可以进行400万亿次操作,是当前非常强大的算力,能支持复杂的AI应用和高级别的自动驾驶功能。

平台整体集成了“视、听、触、嗅、动”五感融合设计,使人形机器人能感知周围环境。

“青龙”平台的核心技术经过了十多年的技术沉淀,标志着我们的人形机器人平

台技术实现从无到有的突破。”梁振杰说。

青龙核心研发团队是国内最早开展仿生腿足式机器人研究的团队之一,拥有近十年机器人行业技术积累,构建了仿生机器人核心技术体系,建立了机器人控制、感知、交互等核心技术群,支撑着“青龙”处在国内人形机器人技术领先水平。

具身大脑+小脑模型:人形机器人的智力

“青龙”拥有“朱雀”具身大脑和“玄武”小脑模型。

“朱雀”具身大脑是一个以多模态大模型为核心的机器人指挥调度中心,利用多模态大模型的感知能力、任务理解能力、记忆能力以及规划能力,帮助机器人完成任务。输入方式是文字和图像信息,支持语音交互,最终将任务决策信息输出给“玄武”小脑模型。

记者采访了解到,在当前发布版本中,“朱雀”具身大脑共集成3个大模型,分别是:科大讯飞星火大模型、上海人工智能实验室书生·浦语大模型、上海人工智能实验室书生·万象多模态大模型。

“朱雀”具身大脑拥有跨设备的调度框架。具体来看,书生浦语大模型和万象多模态大模型运行在本地服务器,其中,语言大模型实现用户意图识别与对话功能;讯飞星火大模型与浦语功能相似,但是运行在云端服务器;多模态大模型具备处理图像的能力,当识别到用户的任务与当前环境相关时,会启用多模态大模型进行环境感知。

中心具身智能负责人田翀说,在后续更新版本中,我们将微调参数量较小的语言模型,直接部署在机器人终端,作为大模型的补充。对于简单问题,机器人直接做出回复,从而减少与服务器通信的时间,实现更快速准确的交互与技能调度。

“玄武”小脑模型是任务执行模型,分为轨迹规划模块和运动控制器两个部分。首先,轨迹规划模块负责输出期望轨迹,它由端到端的机载视觉信息为驱动,以人类动作为模仿对象。然后,由运动控制器控制人形机器人做出相应的动作。

数据:人形机器人的“灵魂”

受访者表示,决定人形机器人智能程度的关键因素是数据。“数据是人形机器人的灵魂,数据越富集,‘大小脑’的智能越高、能力越强。”中心具身智能负责人邢伯阳介绍,“大小脑”需要非常多数据进行综合训练,包括但不限于多样化垂类场景训练数据、多模态语音数据、人体开源数据、运动捕捉数据、机器人本体数据、环境地形数据等。

通过“解剖”小脑模型,能看出数据对其的塑造作用。“我们首先收集了大量的人力数据,以模仿学习为基础打造行为策略和行为标准。然后利用控制理论和强化学习打造运动控制系统,在3个月内完成了小脑模型的算法开发和迭代。”田翀说,我们充分利用了人体运动数据,结合模仿学习策略,为机器人运动训练提供了精准的参考轨迹和运动标准。同时,我们还将传统的运动控制方法应用于数据收集工具的开发,并将这些思想融入到强化学习训练中。

通过技术融合,不仅提高了机器人运动控制策略的训练效率,缩短了训练周期,还确保了策略在实际应用中的安全性和可靠性,这是“玄武”小脑模型与其他小脑模型相比的一大特点。“当然,该模型还在发展初期,能力和功能仍需提升和完善。”田翀说。

借助AI,管控平台同时完成对数据

标注、管理与模型部署的需求,助力实现机器人数据采集、评估清洗、训推一体的全流程闭环。邢伯阳向记者解释说,通过仿真环境对采集数据实现快速的测试,对失败数据单元进行删除,完成对原始数据的清洗,再对神经网络模型进行训练,实现面向机器人端快速部署。

中心计划用3年时间完成“白虎”数据集的构建,一方面建设超过100个高精度运动采集设备,另一方面共享业界其他单位采集的数据,实现清洗后的高质量数据总量超过1PB,覆盖超过100个场景、2000个任务。

训练场:人形机器人的学校

“人为演示一个动作大约50次,机器人就能学会。”在2024世界人工智能大会的国家地方共建人形机器人创新中心展区,记者看到了一个小型的机器人“训练场”,现场工作人员如同一个熟练的“老师傅”做着生产线上的抓取动作,4个机器人则跟着他同步做出了一样的动作。

为什么要搭建训练场?中心技术负责人刘宇飞说,构建智能训练场,模拟产线及工业流水线,收集多模态数据,并搭建数据管理平台,将有效提升人形机器人在不同场景下的作业能力,加速具身智能技术的发展与应用落地。

“归根结底,训练场是数据生产和数据规模化的场地。”刘宇飞表示,训练场分为感、存、算、学、用五个部分。“感”和

“存”主要用来做云端数据的采集,包括要操作的数据、音频数据、自然语言、运动捕捉等。“算”和“学”是指机器人的技能模仿学习和强化学习,主要是做单臂、双臂以及全身行为的运动训练。

刘宇飞介绍,训练场数据的生产方法包括三个阶段:

第一步是打造单臂的单技能学习与作业对象环境的泛化能力。

第二步是基于协作臂把这套技术路线迁移到“青龙”的上肢。

“青龙”的上肢单臂拥有7自由度+6自由度的灵巧手,协作臂是单臂6自由度+夹爪;为了使灵巧手拥有更广的作业空间,没有加入手腕相机。迁移时,我们对采集数据实时优化,在时间上严格对齐动作轨迹和图像,以达到和三相机(1头部+2手腕)同样的学习效果。

第三步则是生产大规模、低成本的人类作业视频。

训练场不仅关注“小脑”模型的训练,还涉及“大脑”模型,如环境感知、行为控制、人机交互、云端网联等能力的训练。

“2024年,我们会在广东打造100+人形机器人的产品、100+人形机器人的训练场。到2027年,我们期待能够在多个城市、面向各类场景,搭建1000+人形机器人训练场来服务整个人形机器人生态。”刘宇飞说。

(《瞭望》新闻周刊记者 董雪 龚雯 实习生 宫雅婷)



实干成就未来 奋斗创造历史