

人形机器人的“拖延症”怎么治？

本报记者 姬晓婷 实习生 贺紫依

你是不是也有这样的疑问：为什么工业机械臂完成动作又快又准，但摊煎饼、做咖啡的机器人，不仅动作慢，还更容易失误？号称具身智能，能够通过对外在世界的感知做出动态反应的机器人，为什么看起来“呆呆”的？

机器人动作慢、反应迟钝的原因，若是从技术层面剖析，可以被简单归纳为一个词——时延长。时延越短，具身智能机器人商用化越能多一分胜算。



图为中兴通讯轻量化全尺寸人形机器人

反应慢，谁在“拖后腿”？

一台合格的具身智能机器人，具备一套包括感知、决策、执行的完整闭环系统。在感知层，机器人需要通过摄像头、激光雷达等视觉传感器，电子皮肤等触觉传感器，麦克风等听觉传感器完成对世界的感知；在决策层，机器人借助端侧处理器和预训练世界模型处理获取的信息；在执行层，机器人通过电机、关节、机械臂、灵巧手等作用于现实世界，改变环境状态。

之所以很多具身智能机器人看起来“呆呆的”，恰恰在于上述环节不同程度地存在处理时延。

在感知前处理阶段，涉及多传感器数据的接人与融合，视频流的时间戳对齐、多源数据同步、压缩编码及空间裁剪等预处理操作，该环节会存在基础性的固定延迟。

在认知决策阶段，机器人基于视觉-语言-动作（VLA）大模型进行场景理解、任务推理与行为规划。该环节的延迟主要来自端侧推理硬件的算力峰值与模型结构的计算复杂度，通常在数百毫秒量级，成为系统延迟的主要构成部分。

机器人对世界的理解有赖于计算模型的完善。现阶段，世界模型（VAM）的能力与人类认知水平仍存在不小差距。中兴通讯副总裁、人形机器人产品总经理崔卓告诉《中国电子报》记者，开源世界模型准确率只有 60%~70%，在机器人对世界的理解力不足的情况下，要保证单次任务的成功便需要调用更多的本地算力，进而耗费更多运算时间。此外，一些特定任务，如透明包装袋、异型工件的识别，算法计算量更大，机器人更易出现漏抓、错抓，甚至产生大批量残次品的情況。

在执行控制阶段，机器人将 VLA 输出的高层动作指令下发至下层运动控制系统（如小脑控制器），后者完成逆运动学求解与动力学约束解算，进而生成关节及灵巧手的驱控信号。该阶段延迟取决于伺服执行机构的带宽、机械传动特性及实时控制周期。机器人的数据传输架构也不够精简。

现阶段，上面提到的机器人三大技术环节分别由机器人上搭载的不同处理器控制——感知采集模块、全局决策大脑、运动控制模块在硬件上相互独立，就像三个人分处不同房间，需要反复传纸条沟通。单传感器数据预处理耗时就动辄几十毫秒到上百毫秒，交由 VLA 大模型完成任务识别、决策、规划，又会消耗数百毫秒。多模态数据在不同芯片间来回搬运，不仅叠及时延，还拉高整机功耗，缩短伺服、灵巧手使用寿命。

灵境智源联合创始人、高级副总裁须海江表示：“当前行业方案中，从感知到决策的路径割裂，且数据都依赖搭载本地的处理器进行推理计算，任务的完成需要从内存中反复搬运数据，便会造成时延和功耗的增加。”

机器人处理时间长，另一重要原因是不同组件之间时延测试基准不统一。传感器、伺服电机、计算芯片、整机通常分属不同企业独立开发，各段软硬件都会产生数毫秒微小延迟，行业此前又长期缺少统一时延测试基准，上下游零部件时序参数互不兼容，也可能大幅增加整机企业联调优化成本。

控时延，按什么标准？

如果不能缩短时延，就会被市场淘汰。记者通过采访发现，时延已成为人形机器人是否会被市场淘汰的硬性标尺。那些仅能在固定静态环境完成演示、无法适配动态场景的“演示样机”正在加速被市场淘汰。在接受《中国电子报》记者采访时，须海江表示，实时响应速度、复杂动态场景任务完成率这两个指标正在从机器人开发的“加分项”变成“及格线”。

今年2月，由工业和信息化部人形机器人与具身智能标准化技术委员会组织领域内 120 余家科研院所、企业和行业用户单位共同编制的《人形机器人与具身智能标准体系（2026 版）》（以下简称《标准》）正式发布。《标准》明确将动态任务完成率、端到端闭环响应时间列为设备验收核心指标，明确大脑、小脑双协同架构整机端到端响应时间≤200ms。《标准》是我国首个覆盖人形机器人与具身智能全产业链、全生命周期的标准顶层设计，也给具身智能机器人的时延控制提供了明确的硬指标。

除了端到端响应时间，崔卓还提出了新的观察视角——动作规划频率与时延存在反比关系：规划频率越高，单个规划周期的可用时间越短，时延越低。动作规划频率数值越高，代表机器人响应速度越快，也就越能适配时延要求更严苛的场景。

崔卓表示，具身智能机器人适用的主流场景可以大致分为三类：商业服务类机器人动作规划频率普遍在 10~20Hz；3C 电子、汽车总装等精密工业场景的动作规划频率率达到 50Hz，感知、运算、传输、执行每层软硬件仅预留数毫秒处理窗口；家政机器人虽然整体时延门槛相对更低，但操作衣物、鸡蛋、瓷碗这类柔性、易碎物品时，依旧需要毫秒级力度实时调节。崔卓补充道，若采用无线通信，上行传输时限约 4 毫秒、下行传输时限为 1 毫秒，容错空间极小。

时延长，怎么降？

记者在采访中了解到，各企业正在采用不同的方式解决具身智能机器人处理任务过程中的时延难题。

崔卓强调了算法优化的路径：机器人通过世界模型提前预判遮挡、物体位移场景，再通过注意力模型自动过滤路过行人、闲置杂物等无关目标，只对风险物体启用高精度算力，便可以大幅减少重复推理耗时。

灵境智源瞄准的是机器人架构层。灵境智源研发的德沃夏克超异构计算架构，采用了“物理合一、逻辑分离”设计思路。“通俗来讲，就是把感知皮层、决策大脑、运动小脑三大模块集成在同一块芯片硬件平台，不用跨设备传输大量数据，从根源上削减传输延迟。”须海江向记者介绍道。

如果把中央脑类比为人类负责长远思考、逻辑规划的前额叶，那灵巧手、关节内置的末端协脑则对应与人体本能密切相关的小脑。通过搭载轻量化端侧 VLA 模型、

视源股份：从场景出发，为应用而生

本报记者 杨鹏岳 实习生 崔烁

柔性操作、无序分拣、无人零售、空地协同巡检、教育实训、商用清洁……在即将开幕的 2026 世界机器人大会上，视源股份将携 MAXHUB 产品矩阵及一系列场景化应用解决方案亮相。

在去年的世界机器人大会上，视源股份荣获了“2025 具身智能机器人创新产品”奖项。如今，凭借丰富的技术经验和优秀的应用能力，视源正将机械臂、机器狗等产品技术拓展至更多的行业和场景，秉持“为应用而生”的理念，让机器人技术在真实场景中产生切实价值。

技术加持彰显机械臂“硬实力”

放眼广阔产业场景，视源的产品和解决方案正拓展到更多的应用场景中。据了解，视源的应用解决方案已覆盖电力巡检、教学实训、无人零售等多个领域，正在为产品加装不同的“技能包”，响应行业真实需求。

在具身智能产业中，机械臂是支撑智能制造与发展新质生产力的关键执行载体之一。视源面向场景应用，推出了 MAXHUB D3 七轴仿人形双臂，延续“为应用而生”的理念，可广泛应用于物流分拣、商业服务、电力巡检等场景。以 MAXHUB D3 七轴仿人形双臂为载体，视源在这些典型工业场景中完成了模型与硬件的深度适配与联合验证。而具身智能要落地，不仅要有“聪明的大脑”，还要有一副经得起实战检验的“好身体”。

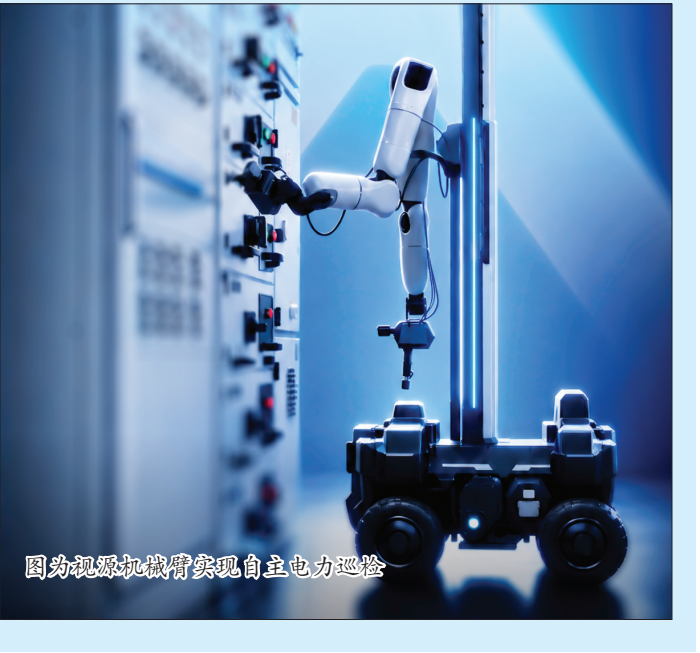
四大硬实力让 MAXHUB D3 七轴仿人形双臂成为具身智能落地的好搭档：长臂展与大负载设计让它能灵活覆盖大范围操作空间；高动态、低延迟的响应速度确保动作流畅敏感；兼容主流开发工具与仿真环境，开发者上手即用；支持多种操作方式和多维度数据精准对齐，为模型训练提供高质量数据底座。MAXHUB D3 七轴仿人形双臂的本质价值在于，它不只是一台孤立的机械臂，更是为具身智能模型量身打造的“数据采集+模型验证+场景落地”一体化硬件平台。

多元场景构建真实应用版图

如今，MAXHUB D3 七轴仿人形双臂正加速进入物流分拣、商业零售、家庭服务、



图为搭载视源机械臂的具身智能教学实验平台



图为视源机械臂实现自主电力巡检

电力巡检等多元场景。物流领域无序分拣快递包裹、家庭场景中完成衣物叠取、无人零售精准抓取商品、电力行业无人化巡检操作……一个个真实的案例构成了 D3 机械臂的场景应用版图。

电力行业是 MAXHUB D3 七轴仿人形双臂实际应用的前沿阵地。作为特种作业场景，电力巡检具有“高风险、高重复、高门槛”的特点。视源带来的电力巡检解决方案让机器人从“只能看”升级为“能操作”，在真实场景中高效完成巡检工作。搭载可升降机械臂的全向轮式底盘可以自主完成配电柜旋鈕操作、按钮点按、开关拨动等精细动作。通过视觉感知系统与 T-WAVES 模型，机械臂不仅能识别旋鈕位置，还能在旋转过程中实时判断是否打滑、受阻或接近机械限位，形成从视觉识别到物理操作的完整闭环。这套方案要解决的是一个实际困境：配电室等场景发现异常后，巡检和处置之间存在“人必须到场”的断点，而机械臂+轮式底盘的组合能够做到发现异常即时就地处置，避免人员入场的风险。

不仅是电力巡检，教学实训也是视源机械臂应用的重要延伸。视源推出的 AI+具身智能教学实验平台，面向高职院校的人工智

能、机器人、智能制造等专业，以双六轴柔性机械臂为核心，融合工业视觉、传送带、遥操作示教和边缘计算等设备，打通从数据采集、标注、模型训练到部署推理、实际验证的全流程。高校具身智能教学中的一个普遍痛点是“设备分散、软件割裂、数据不足、课程难建、真实项目难落地”，学生在课堂上学了算法，到实验室却找不到能验证的设备。视源的做法是把硬件、软件、课程和案例打包交付，配套 140 个行业场景实验案例，每个案例配有数据集、源码、预训练模型和实验指导书，同时开放全套接口支持院校二次开发。

从电力巡检到高校教学实训，视源的应用场景覆盖了多个行业、多个领域。这种多场景的应用能力，来源于视源对机器人运动平台的理解、对软硬一体方案的整合、对行业真实需求的拆解，这种立足于现实场景的打法，让同样的产品技术在不同行业能解决不同的问题。

空地协同打造“无死角”巡检方案

除了机械臂，视源的每一个产品、每一套解决方案，都直接指向了场景应用。

在一些巡检场景中，传统的人工巡检搜

寻方式风险高、响应慢、覆盖难。如何用技术手段将搜索和巡检人员从高危环境中解放出来、最大程度覆盖巡检场景维度，成为机器人与具身智能行业备受关注的议题。

今年 6 月，工业和信息化部、国务院国资委联合印发通知，开展 2026 年度人形机器人与具身智能实景实训专项行动，其中把巡检搜索、检测分析等领域列为人形机器人、四足机器人的重点应用场景。

在这些领域中，视源已有一套独特的解决方案。从真实场景切入，视源构建起“水陆两栖”和“空地协同”两套高度场景化的应用解决方案，以 MAXHUB X7 智能机器狗为核心移动平台，根据场景需求灵活配置。

在两套方案中，空地协同的方案独具特色，以“机器狗+无人机”的组合实现“无死角”巡检。

空地协同方案将四足机器狗与无人机深度集成，实现“地面+空中”协同作业。机器狗搭载无人机起降平台，支持无人机自主起飞、降落与自动充电，无须人工干预即可完成全周期巡检任务。

该方案的核心在于消除单维度巡检盲区：机器狗覆盖地面通道、室内及狭窄区域，无人机补全高处、大跨度空间与屋顶视角，

两层叠加全面覆盖地面和空中两个维度，构建起完整的立体巡检网络。同时，自动起降与充电设计从根本上解决无人机组航受限、频繁更换电池中断作业的痛点，实现无人化持续作业。空地设备统一平台管控，任务编排、数据回传、异常告警一站式完成，运维管理零碎片化，为复杂场景下的全天候智能巡检提供一体化解决方案。

视源并不试图用单一产品或技术解决所有问题，而是依据需求进行组合，构建高度场景化的协同方案，使同一套设备能够在不同场景灵活配置，应对不同情境，最大化地发挥价值。

“为应用而生”：场景驱动的产业逻辑

2026 世界机器人大会以“人机共生，产需共融”为主题，当下机器人产业正在跨越从技术突破到商业应用落地的鸿沟，其关键在于响应行业需求，在真实场景中产生价值。

回看视源推出的应用解决方案，它们之间的关联不在于单一的产品或技术，而是思维方式的一致。这种思维方式正契合现实需要：机器人不是目的，解决场景问题才是目的。

视源的做法在于围绕真实的场景需求做可应用的解决方案。电力巡检方案对应的是让机器人“看得见、够得着、动得了”的操作诉求，教学实训方案瞄准的是“具身智能人才如何培养”的教育命题，空地协同方案回应的是“单一维度巡检盲区”的行业痛点，每一套解决方案的背后，都有明确的应用场景和具体的客户问题。

“为应用而生”不是一句简单的口号，它对应的是一套从场景出发、反向定义产品的研发逻辑。在这种逻辑下，技术指标是手段，场景闭环是目标。在视源的应用解决方案中，机械臂精度高低并不是唯一考量，重点是它不能代替人去拧那个配电柜旋鈕，机器狗能跑多快并不是唯一指标，重要的是它能不能替人进入危险场景。具身智能的价值，最后还要落在“实用”这个点上。

行业普遍预测 2026 年为具身智能的“应用元年”，当行业共识从参数比拼转向价值交付，从实验室走向实地应用，视源股份“为应用而生”的产业逻辑，正踩在一个恰到好处的节奏点上，而即将到来的 2026 世界机器人大会，将是检验这一逻辑的合适窗口。