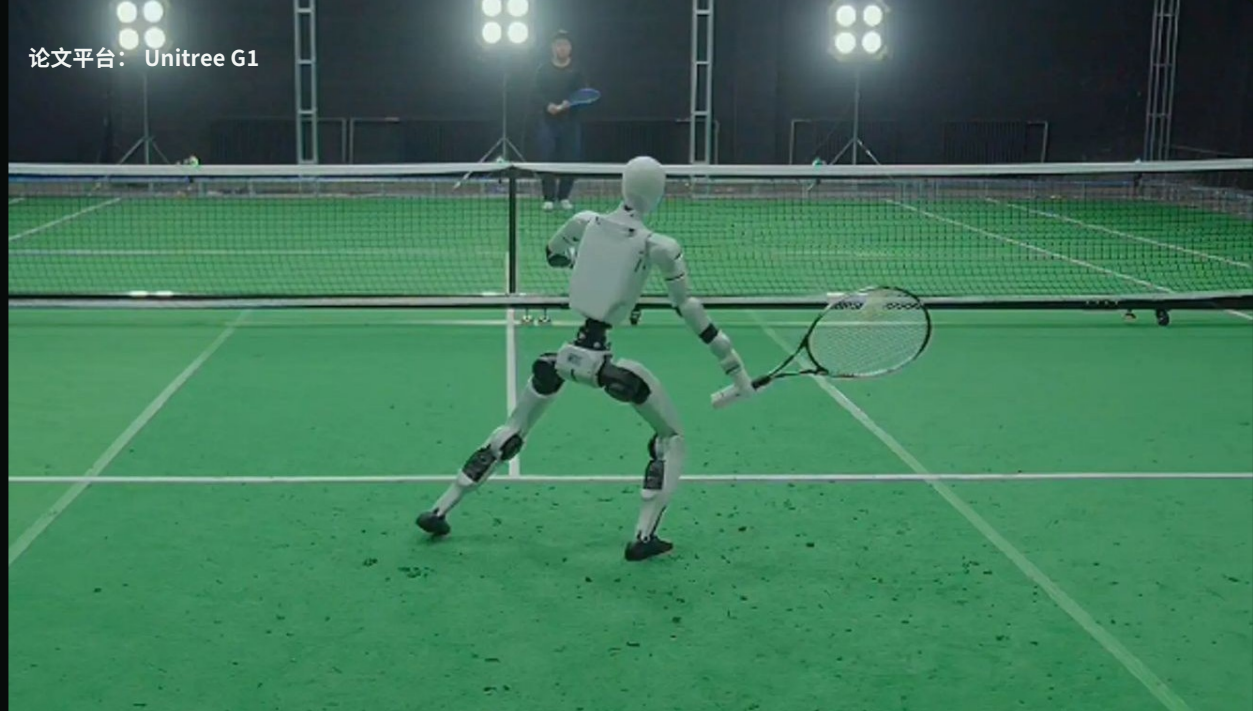


银河通用机器人 为什么会打网球

从 Unitree G1 的论文验证，到 Galbot ET1 的赛事演示：
LATENT 的 whole-body learning 路线，以及跨本体迁移真正说明了什么

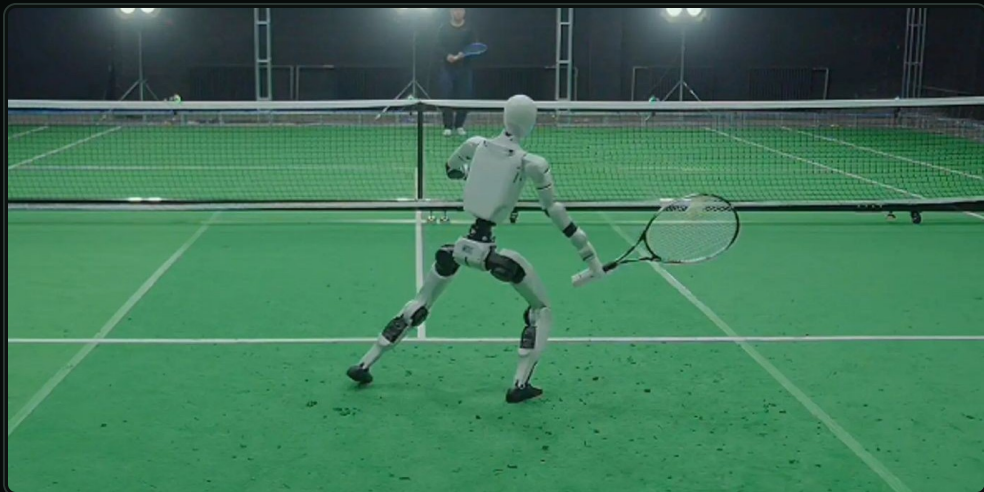
TECHNICAL BRIEF

ONE SYSTEM / TWO BODIES



先纠正一个关键事实： LATENT 先后出现在两台不同的机器人上

MAR 2026 • PAPER



UNITREE G1

- LATENT 论文的全部实验平台
- 29-DoF 商用小型人形
- 外部光学动捕 + 改装球拍末端

LATENT

AUG 2026 • PRODUCT



GALBOT ET1 • “银河星仔”

- 银河通用首款自研双足本体
- 8 月 22 日赛事演示使用 LATENT
- 迁移方式与控制栈尚未公开

同一个算法名字出现在两个本体上，是跨本体能力的线索；但它并不等于“同一组权重无改动直接迁移”。

论文真正证明的，是 LATENT 在 29-DoF Unitree G1 上成立

输入	球与机器人全局状态；真实部署由外部光学动捕提供
先验	正手、反手、shuffle、crossover 等 primitive motion fragments
表示	motion tracker → online distillation → correctable latent action space
策略	高层 PPO 组合技能，并对 latent 与右手腕动作进行修正
输出	50 Hz 全身关节目标；PD controller 转为执行器力矩



5 h
5 名业余球员动作片段

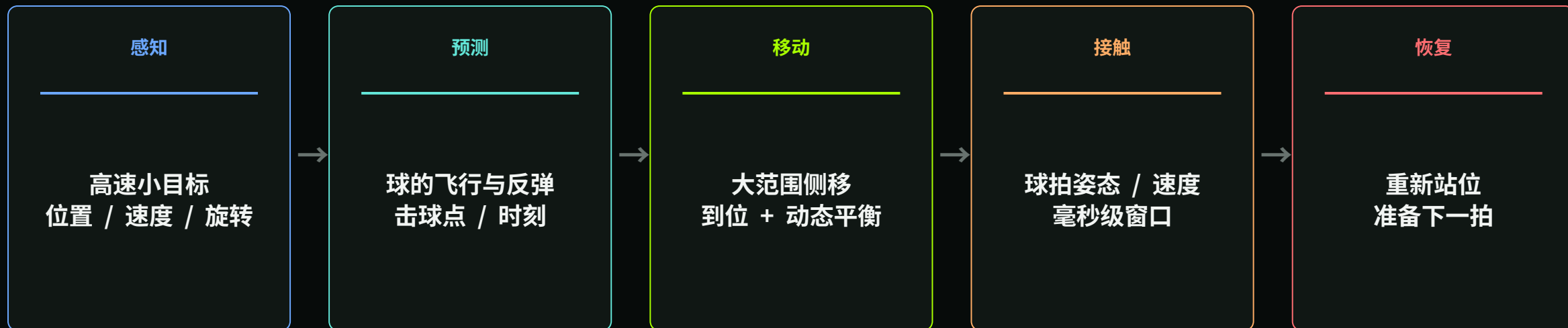
50 Hz
高层与低层控制频率

90.9%
真机正手回球成功率

77.8%
真机反手回球成功率

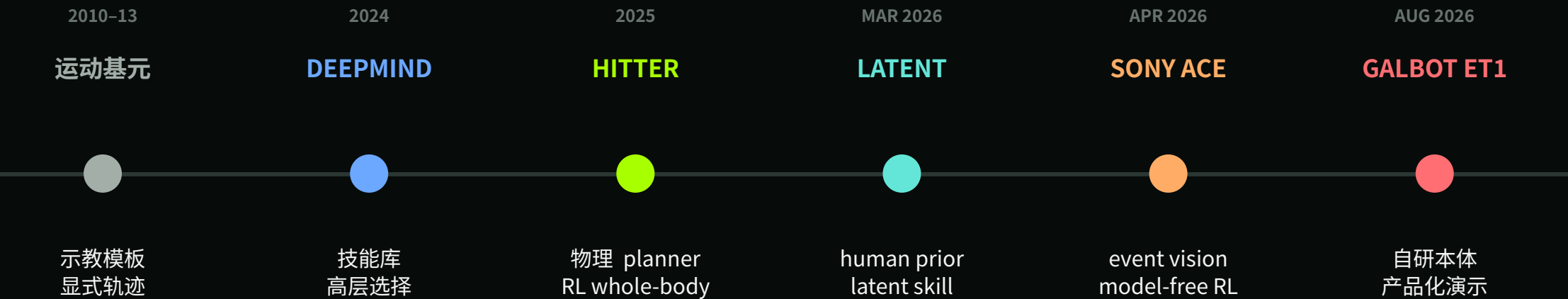
Source: LATENT paper, Sections 3–5 and Appendix B

网球把五个通常能分开解决的问题，压进了几百毫秒



乒乓球把感知与击球时序推到极限；网球进一步放大了移动范围、全身协调和动作先验的重要性。

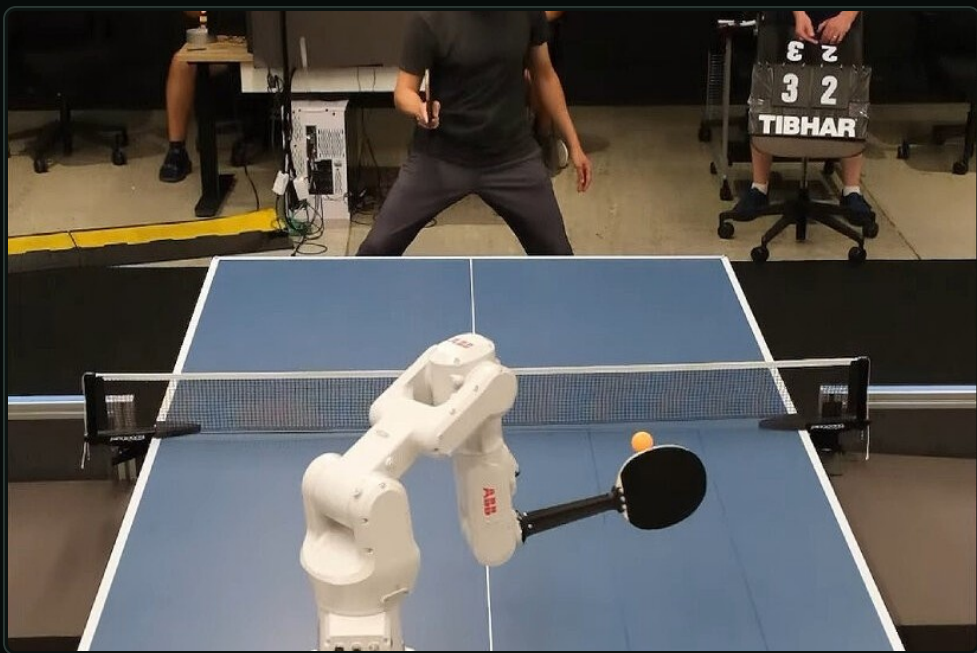
机器人打球的技术史，一直在改变“分层方式”



这条历史线没有走向“取消结构”。越高速、越接近真实交互，系统越需要明确：哪些变量交给物理，哪些交给策略，哪些压进运动先验。

Sources: DeepMind; HITTER; LATENT; Sony AI; Galbot ET1 releases

DeepMind 把击球执行与比赛级技能选择分开



专用机械臂 + 线性滑轨

把双足平衡问题拿掉，专心研究击球与策略。

低层

多个 specialized controllers

正手、反手、不同落点与速度被拆成多个技能。

高层

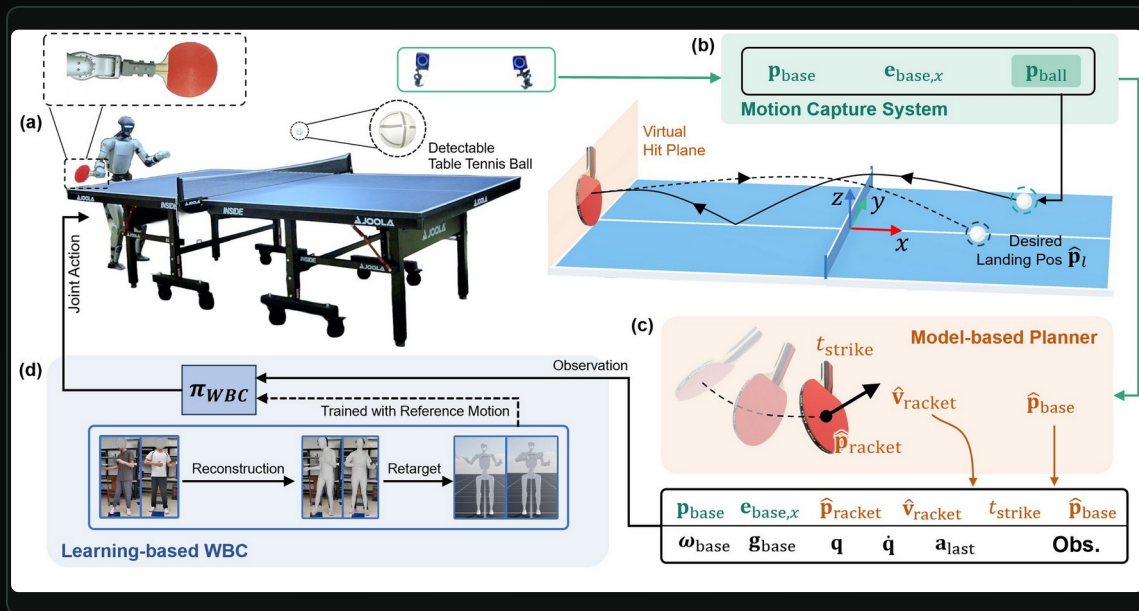
实时选择 skill，并适应陌生对手

13 / 29

真人比赛获胜；初学者 100%，中级约 55%。

智能主要放在 skill selection，而非让一个网络生成所有关节动作。

HITTER 让物理模型决定“打哪里”，RL 决定“身体怎么去”



01 状态估计

9 台 OptiTrack，360 Hz；平滑位置并估计速度。

02 轨迹预测

显式建模重力、空气阻力与台面反弹。

03 击球规划

给定目标落点，计算球拍位置、速度与时刻。

04 全身执行

PPO 输出 29 个关目标；PD controller 转为力矩。

106

最多与人连续击球

HITTER 最值得看的，是“哪里不需要学习”

**0.5 s**

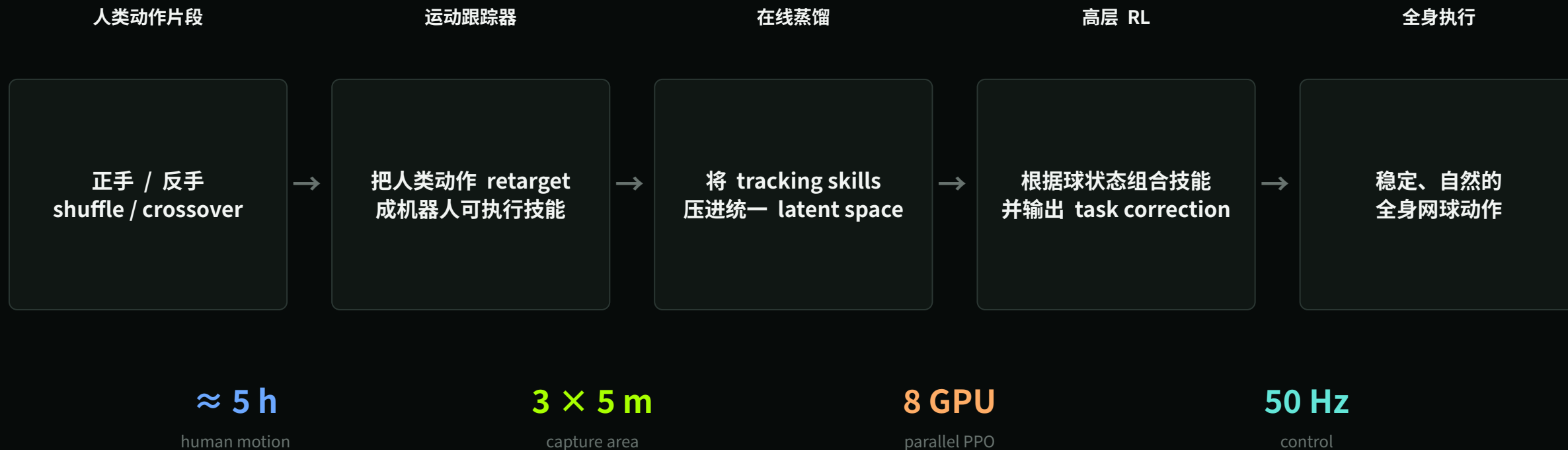
击球前：位置预测误差进入 7.5 cm 球拍半径阈值

0.3 s

击球前：时间误差低于 20 ms，一个 50 Hz 控制周期

哲学：planner 给出稳定、低维、可解释的 command interface；RL 处理 29-DoF 全身协调。

LATENT 先学会人的“运动语言”，再学习如何回球



高层策略不再直接搜索 29 个关节，而是在已经包含步法、转髋与挥拍结构的低维运动流形中搜索。

不完美的人类动作也够用——前提是它提供结构，而不是答案

传统目标	LATENT 的切法	换来的东西
完整比赛动作	只采 primitive fragments	采集成本显著下降
完整球与人轨迹	允许片段不完整	数据更容易扩展
精确手腕与球拍姿态	先学动作风格与协调	低层保留 human-like prior
大场地、长时间采集	任务精度交给 RL 修正	高层仍能针对来球优化

关键假设：人类数据只需提供 primitive skill 的结构先验；task-specific 精度可以由后续策略补回来。

Source: LATENT, data collection and method

运动先验必须引导策略，但不能把策略绑死



只做 imitation，机器人可能“像人”却打不中；只做 task RL，又容易利用 latent space 产生抖动和不自然动作。

本体换了，尚未被解释清楚的感知问题也换了

UNITREE G1 · 论文系统

50+

motion-capture
cameras

2048²

camera resolution

120 Hz

mocap update rate

19 × 15 m

capture area

≈ ¥350k

3 周场地与系统租赁

全球球位与机器人位姿由场地直接提供。

GALBOT ET1 · 公开演示



公开表述

赛事报道称 ET1 使用双目视觉感知来球，并以 LATENT 完成全自主对打。

尚未公开：是否仍有外部相机、球状态估计器结构、端到端延迟、控制频率，以及从 G1 到 ET1 改了哪些策略与接口。

Sources: LATENT GitHub deployment details; Galbot ET1 event coverage

四套打球系统，其实在回答四种不同的问题

系统	主要难题	高层结构	低层结构	被外包的部分
DEEPMIND 2024	竞技击球 / 对手适应	skill selection	learned low-level skills	双足 / 全身运动
HITTER 2025	高速人形击球	physics planner	RL whole-body controller	外部 mocap / 旋转简化
LATENT 2026	网球全身技能组合	latent skill policy	motion prior + RL	论文：外部 mocap
SONY ACE 2026	专业级速度与旋转	model-free rally policy	优化器 + 高速执行	专用硬件成本

没有哪套系统是“纯端到端”。差别在于结构化中间层是 skill、physics target、latent prior，还是轨迹优化器。

Sources: DeepMind; HITTER; LATENT; Sony AI / Nature

两个本体提示了可迁移性，但还没有证明“即插即用”

它提示了什么

- 同一套 tennis skill 表征可以被重新部署到不同尺度与动力学的身体。
- 银河通用开始把论文能力放进自研硬件与产品叙事。
- 本体不再只是实验耗材，而成为模型能力能否复用的测验。

它没有证明什么

- G1 与 ET1 共用同一组 policy weights。
- 29-DoF action space 可无缝映射到 ET1。
- 无需重新 retarget、微调或训练低层 controller。
- ET1 的视觉闭环已达到论文级可复现实证。

下一份真正有价值的材料，不是更漂亮的视频，而是 G1 → ET1 的 retarget、控制接口、训练量与感知栈说明。

下一步不是“再打得快一点”，而是补齐四个接口

01 ONBOARD PERCEPTION

在高速、遮挡与运动模糊下，机载相机能否稳定估计球的位置、速度和旋转？

02 TRANSFER DISCLOSURE

从 G1 到 ET1，retarget、action space、低层 controller 和训练量分别发生了什么？

03 GAME STRATEGY

从 return/rally 走向发球、旋转、落点控制、对手建模与真正的比赛决策。

04 SYSTEM COST

离开高规格场地后，感知、算力、关节性能与可靠性能否维持同样的闭环？

ET1 让 LATENT 的问题从“论文是否成立”变成了“能力能否穿过本体、感知和工程系统继续成立”。

更多机器人技术路线解读 持续更新在 ROBOATLAS.AI

ROBOATLAS.AI 是面向机器人产业的研究与数据机构，持续整理全球机器人公司、产品、模型、关键参数与技术路线，帮助从业者建立可比较的产业认知。

roboatlas.ai

行业交流、产品补充与合作联系



行业微信群

加入行业读者群



联系人二维码

报告咨询，产品补充

机器人 | 模型 | 公司 | 参数 | 技术路线