arXiv:2611.32185v2 [cs.RO] 25 Nov 2026

OVOK：一种用于基于视频的关节式机器人状态估计的递归式自我改进智能体框架

Cheng Dai[1*‡] Renkai Luo[1, 2*] Guanfeng Tang[1] Siming Li[1] Hao Zhao[1†‡]
AIR，清华大学 [2]清华大学求真书院

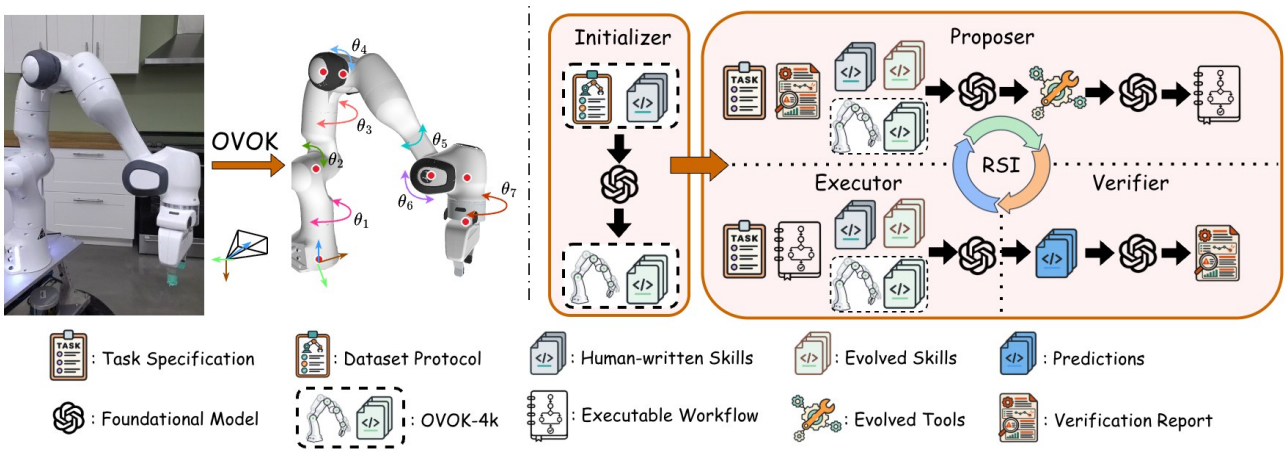

图1. OVOK（一段视频到运动学）概览。左：一个可执行工作流将RGB视频映射为相机内参、相机相对机器人位姿以及关节角度。右：初始化器为特定机器人构建数据集OVOK-4k；在递归自我改进过程中，提议者–执行者–验证者循环利用反馈积累技能、改进工具和工作流，直至形成精简的求解器。

摘要

从视频中恢复关节式机器人状态，有助于为操作、场景重建和动作规划提供公制空间感知。然而，这个问题存在病态性：不同的内参、相机到机器人位姿和关节角度，都可能解释相似的观测。以往的特定任务回归器通常假设内参已知，并通过耗时的数据收集和机器人专用训练来拟合单一机器人，但在不同场景间的迁移效果仍然不佳。基础模型无需训练即可零样本应用于任何机器人，但其估计误差较大。我们提出 *OVOK*（一段视频到运动学），这是一种用于该任务的递归式自我改进智能体框架。在人工编写的协议下，智能体式初始化器构建针对特定机器人的 *OVOK-4k*；随后，提议者–执行者–验证者循环根据反馈共同演进工具和可执行工作流。只有在验证器判定取得正向改进时，更新后的工具和工作流才会被接受。保留的工作流将被冻结，以实现准确推理。实验表明，演进后的工作流在 *DREAM-V* 和 *RoboKeyGen-V* 上取得了最先进的结果，提供了稳健且准确的估计。

1. 引言

基于视频的关节式机器人状态估计可在已知度量机器人模型的条件下，从 RGB序列中恢复相机内参、相机相对机器人位姿和关节角度。作为原始视频与物理实体之间的度量接口，它支撑着基于视觉的操作和闭环控制 [1–6], 4D 场景重建 [7–9], 以及细粒度运动规划 [10–18]。

从 RGB 中恢复这些量仍然具有挑战性：即使已知度量机器人模型，未知的内参和相机到机器人位姿也会使投影与机器人位姿耦合。
关节运动 [19, 20], 而自遮挡 [21], 部分可见性 [3, 22], 较小的图像占幅和成像退化——

‡项目负责人。*贡献相同。†通讯作者。

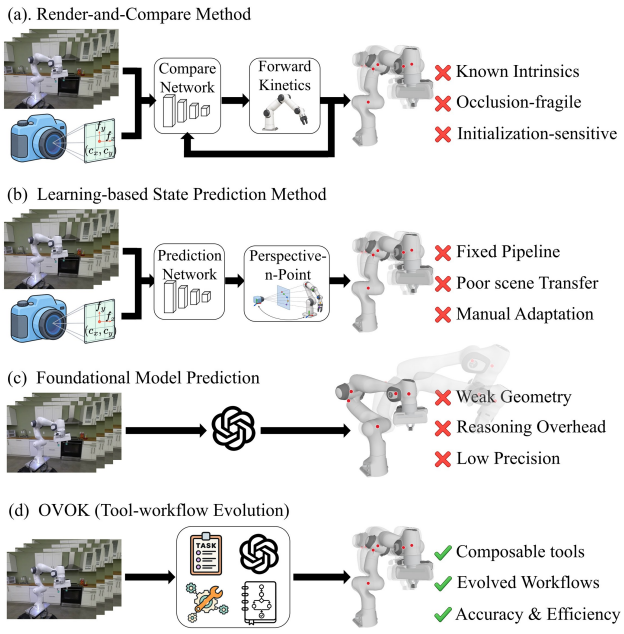

图2。对比范式（a–c）与 OVOK（d），展示了用于机器人状态估计的工具和工作流的演进。

（例如光照变化和模糊）仅留下微弱的几何证据 [23]。

如图 2所示，先前工作主要遵循两种特定任务的范式。渲染并比较的优化 [24–26] 在初始化良好时精度较高，但通常需要已知内参，且容易受到遮挡影响。基于学习的状态预测器 [19, 22, 23, 27, 28] 估计状态，但依赖机器人专用训练、手工设计的流水线，并且跨场景迁移能力有限 [23]。

除了这些范式之外，近期的多模态基础模型可通过提示直接从 RGB [29–33]输出关节式机器人状态。然而，这些估计在度量上仍不够精确 [34], ，反复进行视觉推理也会带来很高的开销。如图 2所示，零样本定位误差和几何不一致性揭示了语义理解与度量估计之间的差距 [35]。

弥合这一差距，需要的不只是更强的零样本预测器。递归自我改进（RSI）旨在让自主智能体从经验中获取知识，并逐步改进其求解器 [36–38]：给定任务反馈，多模态大语言模型 [39]提出修改方案并进行评估，只保留能够提高准确率或降低成本的修改。然而，在几何感知领域，大多数 RSI 系统改写的是编排逻辑或智能体代码 [36–38, 40–54]；它们很少重新设计用于生成度量测量的感知工具。

因此，我们引入 **OVOK** （One Videoto Kinetics），这是一种任务专用 RSI 框架，可演化出一个可执行的感知与几何工作流 [35, 55–57], ，用于仅根据视频并在已知统一机器人描述格式的条件下，估计完整的关节式机器人状态：相机内参、相机到机器人位姿和关节角度。基于智能体的 初始化器 构建了 OVOK- 4k，这是一个用于工具开发和工作流评估的合成数据集。在线演化期间，提议器–执行器–验证器循环根据反馈 [36, 58–62]积累技能，并修订工具和工作流。随后，将保留的工作流冻结，以用于推理。

具体而言，OVOK 针对图2所示的失效模式。首先，视觉模型作为工具被调用，以提供粗略估计，从而减轻遮挡脆弱性和初始化敏感性。其次，OVOK 对感知与几何工具及工作流进行演进，以消除先前方法中对已知内参的假设、固定流水线以及场景迁移能力弱等问题。第三，基于少量人工种子——一个初始化协议、一个初始技能–工具–工作流三元组，以及一套演进正则化与调度方案——OVOK 开展长时程自我发现与修订，最终保留一个度量精度和鲁棒性更高的紧凑工作流。

总之，t hi s wor k makes f our contr ibutions:

- **RSI 框架：**我们提出了 OVOK，这是一种递归式自我改进框架，可联合演化感知与几何工具及其可执行工作流，从未校准的 RGB视频中恢复关节式机器人状态。
- **通过 OVOK-4k 实现闭环：**由人工编写的初始化协议使初始化器能够构建 OVOK-4k，这是一个包含真实状态的视频集，可为工具适配、工作流选择和验证提供反馈。
- **面向长周期演化的轻量种子：**借助技能–工具–工作流种子以及演化正则化与调度方案，OVOK 能够持续实现长周期改进。
- **冻结的可迁移求解器：**演化后的工作流在推理时保持冻结，并在 DREAM-V 和 RoboKeyGen-V 上实现了最先进的准确率和鲁棒性。

2. 相关工作

2.1. 关节式机器人状态估计

当已知关节角度时，DREAM [1], CtRNet [2],和 CtRNet-X [63]利用视觉对应关系和运动学几何恢复相机到机器人位姿。SGTA- Pose [21]融合时间线索，而 MonoSE(3)- Diffusion [20]则通过条件去噪优化位姿。

同时估计关节运动和相机相对位姿，无需满足关节状态这一要求。RoboPose [24]采用迭代式渲染并比较优化，而 HoRo- Pose [19]将状态回归与学习式深度估计相结合。RoboKeyGen [23]和 RoPE-Flow [27] 分别通过扩散和图条件流匹配生成三维关键点。RoboPEPP [22] 采用联合掩码预训练来预测关节角度和关键点，随后进行几何位姿恢复。RoboTAG [28] 通过拓扑图对齐 2D 和 3D 分支，从而减少

对标注数据的依赖，并采用一致性监督。

基于深度的方法利用额外的几何观测。FEEPE [3]结合测量深度与运动学运动先验，以优化末端执行器位姿；而 SPDH [64] 则根据深度估计三维关节位置。这些方法依赖深度测量，因此在无法进行深度感知的场景中用途受限。

上述估计器通常假设相机内参已知 [22, 27], ，这限制了其直接用于未标定视频的能力。更广泛而言，手工设计的感知模块和求解器流水线需要进一步的工程工作，才能应对新的观测条件和失效模式。

2.2. 通用递归自我改进

除了通过训练 [65–71], 更新模型参数的方法外，近期研究还通过优化基础模型周围的运行框架来改进智能体 [72]。执行反馈指导运行框架的修订，而修订后的智能体则为下一轮提供经验 [38, 40, 43,44]。在迭代过程中保留的知识和可执行组件，进而支持进一步改进 [73, 74]。

这些方法可根据其修订对象进行分类。一类方法修改模型读取的内容，根据执行结果更新提示和记忆，而不改变模型参数 [58–60, 75, 76]。第二类方法修改模型能够执行的操作：技能逐渐积累为可复用的流程，工具则提供额外的计算量 [74, 77–81]。第三类方法修改组件的组合方式，通过搜索工作流和智能体实现 [37,41， 42, 45–47], 这依赖于信息丰富的反馈和可靠的候选项评估，以保留有用的修订 [48]。

2.3. 几何任务的自我改进

近期的基础模型（例如 GPT-6-Astra）展现出强大的三维空间理解和测量能力 [31, 82–85]。然而，几何任务仍然十分困难：必须在严格的延迟约束下，从视频中恢复相机相对位姿、关节角度和度量距离，而当前模型只能给出粗略估计，无法提供精确结果。造成这一差距的根源有两个局限性。首先，模型缺乏基于物体几何先验（例如统一机器人描述格式或网格）的显式度量空间推理能力 [35, 55, 86–88]。其次，它们既不提供用于测量的计算工具，也没有可调用这些工具的接口 [81]。

有两类方法分别应对这些局限性。第一类方法强化空间推理，通过反复检查、提出假设并根据观测 [33, 35, 55] 进行修订来细化估计；准确率有所提高，但推理时间也随之增加。第二类方法提供深度、标定和优化器等测量工具，并在推理过程中对其进行编排 [24, 25, 54, 79–81, 89]。它们的预测结果通常更加稳定且准确，但编排往往由人工设计，很少有方法能够提供一种高效的反馈驱动机制，用于修订工具和组合方式。这正是 OVOK 所要解决的空白。

3. 预备知识

状态与观测 设 $\mathcal{M}$为一个已知的机器人运动学模型：它由一棵连杆树、 n 个带有限位的旋转关节 $\mathcal{Q} \subset \mathbb{R}^n$以及每个连杆上的一个本体坐标系组成。一个构型 $q \in \mathcal{Q}$ 通过正向运动学确定机器人基座坐标系中每个连杆的刚体位姿 $\mathrm{FK}_j(q) \in \mathrm{SE}(3)$ j 。

一台内参为 K 的 RGB相机记录了一段视频 $I_{1:N} = (I_t)_{t=1}^{N}$，其中 $I_t \in \mathbb{R}^{H\times W\times 3}$。在第 t 帧，机器人状态为 $(\mathbf{T}_{c\leftarrow b},\ q_t)$ ，其中 $\mathbf{T}_{c\leftarrow b} \in \mathrm{SE}(3)$ 将机器人基座坐标系映射到相机坐标系，而 q_t 指定关节构型。对于定义在选定连杆原点处的运动学关键点 $\mathcal{J}$，其图像投影为

$$\mathbf{p}_{t,j} = \pi\big(K,\ \mathbf{T}_{c\leftarrow b},\ \mathrm{FK}_j(q_t)\big) \in \mathbb{R}^2, \qquad j \in \mathcal{J}. \quad (1)$$

在以 K 和 q_t 为条件时，正向运动学可提供三维关键点，其图像对应关系可在几何非退化的情况下通过 PnP [22, 24, 90] 实现恢复。

任务定义 给定 $I_{1:N}$ 和 $\mathcal{M}$，我们逐帧估计相机内参 K、相机相对位姿 $\mathbf{T}_{c\leftarrow b}$和关节角度 $\hat{}q_t \in \mathcal{Q}$ 。我们在 Franka E mika Panda 上实例化该方法，并将 $\mathcal{J}$ 作为七个连杆原点标记。

公式设计依据 先前工作通常使用单张 RGB 图像预测二维/三维关键点以及关节角度 [19, 22–24, 27, 28]。在已知内参的情况下，图像-模型对应关系可通过 PnP求解器 [90] 实现相机相对位姿的度量恢复。相较之下，我们基于已知度量模型 $\mathcal{M}$，直接从视频联合估计 K、$\mathbf{T}_{c\leftarrow b}$和 q_t ，原因有二。首先，将 K 视为已知通常意味着需要离线相机标定；这在实践中十分繁琐，也阻碍了未标定视频的录制。其次，视频是实际部署中的原生观测：时间一致性能够对受自遮挡或语义歧义影响的帧进行正则化，而度量机器人模型则提供绝对尺度，因此与单张未标定图像相比，估计问题的欠定程度较轻。

4. 方法

4.1. 概述

OVOK 是一个用于关节式机器人状态估计的工作流演化框架。其三个阶段（初始化、演化、推理）和四种角色（初始化器、提议器、执行器、验证器）均围绕基础模型组织。

初始化奠定数据基础。按照人工编写的协议（见附录A），初始化器搭建仿真与渲染流水线，并生成 OVOK-4k：3,600 个训练片段，用于开发工具

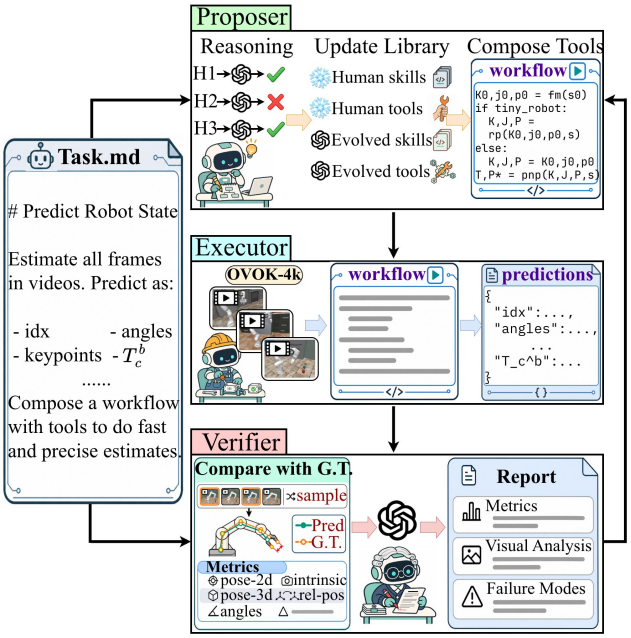

图 3. OVOK 循环。提议器提出假设，在训练集上修订工具，并组合出一个冻结的工作流。执行器在验证视频上运行该工作流，验证器则将所得指标和故障分析转化为下一次迭代的修改。

；另有 400 个验证片段，用于评估工作流。演化按照图3所示的循环进行。提议器提出假设、更新技能、在训练划分上修订工具，并组合工作流；执行器在验证集上运行冻结的工作流；验证器审计结果，将测得的误差和诊断出的失效模式转化为报告，指导下一轮修订。推理阶段则将演化后的工作流部署到真实视频中，高效地返回估计结果。

4.2. 演化空间

按照第 4.1 节所述，OVOK 围绕三类制品进行演化：技能 $\mathcal{S}$、工具 $\mathcal{U}$ 和可执行工作流 $\mathcal{W}$。技能 $s \in \mathcal{S}$ 记录任务知识 [59, 75, 91, 92]：相关工作、积累的假设和未解决问题。工具 $u \in \mathcal{U}$ 是可调用计算 [41, 78]。工作流 $\mathcal{W}$ 将工具绑定为程序 [46, 61, 93]。

演化策略 在演化过程中，我们施加一个结构先验：关节状态估计可以模块化分解为子问题。模块化分解符合任务的结构，也符合实践中的解决方式。当前方法通常采用模块化的串行流水线：关键点定位和相机相对位姿恢复在估计器 [1, 2, 19, 21–28, 63, 90] 中反复作为可分离的子问题出现。然而，部署需要自适应深度：观测良好的片段只需经过较短路径即可获得准确估计，而遮挡……被遗漏或含糊的内容需要更强的推理 [41, 42,47]。因此，将这些模块组合成可分支工作流能够覆盖比单一串行流水线更多的边界情况，将故障归因于负责的单元，从而使修复方案得以复用，并且让每次修订的成本都远低于重新训练端到端模型。此外，模块化设计高效且可扩展，并能通过对单个模块进行增量更新来适应新的观测。

技能 技能陈述了子任务是什么、何时适用，以及已从中学到什么；它是 OVOK 在迭代间传递知识的基本单元 [59, 77]。

技能分为两类。人工编写的种子使提议器能够在尚无任何证据时理解任务：问题定义以简洁的数学形式书写，先前工作则以清晰、结构化的语言加以总结。种子在整个运行过程中保持冻结，从而避免目标随着技能和工具的演化而偏移 [37]。

演化出的技能依据验证报告编写，只包含两项内容：已验证的假设，以及接下来需要解决的问题 [43, 44, 48, 49, 58, 72, 75, 91]。

工具。 工具分为两类。感知工具可在场景中实现稳健的实时模式识别，尤其是在观测受限的情况下。计算工具则通过确定性方法，互为补充地执行精确的几何计算 [35, 55, 79, 81]。

初始化时，会提供一组冻结的骨干网络（例如 Depth Anything 3 [94], DINOv3 [95], 和 V-JEPA [96]），以及一组可变工具，涵盖工作流可能需要的各种操作 [35, 81]。在演化过程中，会针对提出的假设修订工具，并将其累积到 $\mathcal{U}$ [46]中；其中，感知工具还需要进行任务特定适配。OVOK 支持在冻结的骨干网络上微调适配器，并在训练划分上训练小型任务专用模型，耗时仅需数小时 [97, 98]。

工作流。 在 OVOK 中，工作流是一个可执行 Python 程序 [41, 42, 45, 61], ，包含条件分支、数据依赖和工具绑定。它在工具修订后构建，并在执行期间保持不变。

4.3. 初始化

初始化构建数据基础。按照人工编写的协议，初始化器组装仿真与渲染流水线，并生成 OVOK-4k：包含 4,000 个配对视频片段的数据集，每个片段均包含 RoboCasa 厨房中 Franka Panda 的 RGB视频和仿真器状态 [1, 19, 23, 99, 100],

OVOK-4k 分为 3,600 个训练片段 $\mathcal{D}_{\text{tr}}$ 用于开发工具，以及 400 个验证片段 $\mathcal{D}_{\text{v}}$ 用于评估工作流，且房间布局彼此不重叠。每个片段仅在通过独立的定量检查后才会被纳入（见附录A）。

4.4. 演化

演化目标。 对于任务 τ（第 3 节），OVOK 的目标是：

$$(\mathcal{U}^\star, \mathcal{W}^\star) = \arg\max_{\mathcal{U},\mathcal{W}} J_\tau(\mathcal{U}, \mathcal{W}; \mathcal{D}_\mathrm{v}), \tag{2}$$

其中 J_τ 在固定的评估协议下衡量准确率和效率 [37, 46, 51, 58]（见附录B.3）。

演化进展。 为优化（式 2），OVOK 运行提议器–执行器–验证器循环。提议器读取技能、工具和先前的工作流以及验证报告，并提出 M 假设 $\{\delta_m\}_{m=1}^{M}$，针对求解器的不同方面。它从人工初始化的 $(\mathcal{U}_0、\ \mathcal{S}_0、\ \mathcal{W}_0)$ 开始，报告为空 $r_0 = \varnothing$ [53]（见算法 1）。给定这些假设，提议器在 $\sim\mathcal{D}_\mathrm{tr}$上修订工具，生成 $\mathcal{U}_i$，然后组合工作流。未通过廉价试运行探测的候选项会在完整验证之前被拒绝 [41, 42, 48, 58, 60, 101]。随后，执行器在验证集划分上运行组合后的工作流。

验证器分两遍检查每次运行结果。定量方面，它会测量二维/三维姿态、相机到基座位姿、内参、关节角度及相关量的误差，同时记录吞吐量。定性方面，它会抽取高误差片段，诊断其失败模式，并将其汇总为失效模式。两项分析都会写入验证报告，用于指导下一轮。只有在 $J_i > J^\star$时才会保留修订版本，即当其提高了精确度或效率时（见附录B.3）。

正则化与调度。 不受约束的大语言模型反馈可能错误归因、追逐存在偏差的局部收益，并采取不受控的编辑步骤 [36, 37, 75, 102]。因此，我们在 RSIloop 中注入两项人工控制。

首先，归因被限定在八种固定失效模式内：内参估计不准确、严重的$\mathbf{T}_{c\leftarrow b}$ 偏差、全局关节角度偏差、困难的基座方向估计（J1）、困难的末端执行器估计（J7）、遮挡/部分可辨识性、光照以及较低的计算效率。

其次，我们建议采用分阶段的修改规划：首先优先锁定相机内参 K，这些内参可以通过环境背景纹理轻松估计 [103, 104]，并且是度量恢复的前提；一旦 K 被判定为足够准确，便应专注于$\mathbf{T}_{c\leftarrow b}$ 和关节估计器，因为它们对最终位姿精度起主导作用；当这些几何基础满足验收标准后，后续修改可以在准确率不会大幅下降的情况下提升效率（见附录B.3）。

4.5. 推理

在实际使用中，OVOK 会冻结演化后的构型（$\mathcal{U}^\star$，$\mathcal{S}^\star$，$\mathcal{W}^\star$）。对于观测视频 X，系统会直接执行组合后的工作流：

$$\widehat{Y} = \textsc{Execute}(X; \mathcal{U}^\star, \mathcal{W}^\star). \tag{3}$$

算法1 初始化工作流 $\mathcal{W}_0$

需要： 视频 $I_{1:N} = (I_t)_{t=1}^{N}$、机器人模型 $\mathcal{M}$、工具 $\mathcal{U}_0$
确保： $\hat{K}$，$\hat{\mathbf{T}}_{c\leftarrow b}$，$\{\hat{q}_t\}_{t=1}^{N}$
$\hat{K} \leftarrow$ 一些化实验(I_1) [104]
对于 $t = 1, \ldots, N$ **执行**
  $(\hat{\mathbf{T}}_{c\leftarrow b,t}, \hat{q}_t) \leftarrow \textsc{HoRoPose}(I_t, \hat{K}; \mathcal{M})$ [19]
结束循环
$\hat{\mathbf{T}}_{c\leftarrow b} \leftarrow \frac{1}{N}\sum_{t=1}^{N} \hat{\mathbf{T}}_{c\leftarrow b,t}$
返回 $\hat{K}$, $\hat{\mathbf{T}}_{c\leftarrow b}$, $\{\hat{q}_t\}_{t=1}^{N}$

算法 2 OVOK 中的提议者–执行者–验证者循环

要求： 任务 τ、$(\mathcal{U}_0,\ \mathcal{S}_0,\ \mathcal{W}_0)$、$\mathcal{D}_\mathrm{tr}$、$\mathcal{D}_\mathrm{v} = (X_\mathrm{v},\ Y_\mathrm{v})$、迭代 N
确保： 已确认 $(\mathcal{U}^\star,\ \mathcal{S}^\star,\ \mathcal{W}^\star)$
$(\mathcal{U}^\star,\ \mathcal{S}^\star,\ \mathcal{W}^\star) \leftarrow (\mathcal{U}_0,\ \mathcal{S}_0,\ \mathcal{W}_0)$
$J^\star \leftarrow 0$, $r_0 \leftarrow \varnothing$, $r^\star \leftarrow \varnothing$
对于 $i = 1, \ldots, N$ **do**
  $\{\delta_m\}_{m=1}^{M} \leftarrow \textsc{Reason}(\tau,\ \mathcal{S}_{i-1},\ \mathcal{U}_{i-1},\ \mathcal{W}_{i-1},\ r_{i-1},\ M)$
  $\widetilde{\mathcal{U}}_i \leftarrow \textsc{Revise}(\tau, \{\delta_m\}_{m=1}^{M}, \mathcal{U}_{i-1}, \mathcal{D}_\mathrm{tr})$
  $\widetilde{\mathcal{W}}_i \leftarrow \text{组合}(\tau, \{\delta_m\}_{m=1}^{M}, \widetilde{\mathcal{U}}_i)$
  $(\widehat{Y}_i, \rho_i) \leftarrow \text{执行}\textsc{xecute}(X_\mathrm{v}; \widetilde{\mathcal{U}}_i, \widetilde{\mathcal{W}}_i)$
  $(J_i, r_i) \leftarrow \text{验证}(\tau, \widehat{Y}_i, Y_\mathrm{v}, \rho_i)$
  如果 $J_i > J^\star$ **则**
    $(\mathcal{U}_i, \mathcal{S}_i, \sim_{\mathcal{W}_i}) \leftarrow (\mathcal{U}_i, \text{合并}(\mathcal{S}_{i-1}, \{\delta_m\}_{m=1}^{M}), \widetilde{\mathcal{W}}_i)$
    $(\mathcal{U}^\star, \mathcal{S}^\star, \mathcal{W}^\star, r^\star, J^\star) \leftarrow (\mathcal{U}_i, \mathcal{S}_i, \mathcal{W}_i, r_i, J_i)$
  else
    $(\mathcal{U}_i, \mathcal{S}_i, \mathcal{W}_i) \leftarrow (\mathcal{U}_{i-1}, \mathcal{S}_{i-1}, \mathcal{W}_{i-1})$
  结束 if
结束 for
返回 $(\mathcal{U}^\star, \mathcal{S}^\star, \mathcal{W}^\star)$

单次运行会为每一帧返回相机相对位姿、关节角度、内参以及二维/三维关键点。

5. 实验

5.1. 实现细节

评估数据目前尚无用于机器人位姿和关节角度估计的成熟视频级协议。我们从 DREAM-real [1]和 RoboKeyGen-Franka [23],中提取运动片段，分别将其称为 DREAM-V 和 RoboKeyGen-V。

DREAM-V 包含来自不同相机设置（Azure Kinect、Kinect 360、RealSense 和 Orb）的 59 组配对视频片段，平均时长为 20.5 秒。RoboKeyGen-V 包含来自两台相机（RealSense 和 Azure Kinect）的 14 组配对视频片段，平均时长为 10.3 秒。两者均为移动中的 Panda 机器人在真实世界中的录制视频，编码帧率为 30 每秒帧数。在每段视频中，相机和机器人基座均保持相对静止。

实验方案 我们在统一的实验方案下评估三组方法。任务专用模型包括 RoboPose [24], HoRoPose [19], 和 RoboPEPP [22]。部分方法因缺少可用的公开实现而未纳入评估 [20, 23, 27, 28] 。

表1。DREAM-V 和 RoboKeyGen-V 上的机器人位姿估计。2 像素下的 PCK 和 20 毫米下的 ADD 均为成功率。 ✓：在测试时给定。 *：通过基础模型演化得到的 OVOK-solver。

方法	边界框	K	DREAM-V			RoboKeyGen-V		
			PCK ↑	ADD ↑	每秒帧数↑	PCK↑	ADD ↑	每秒帧数 ↑
特定任务方法								
RoboPose	×	✓	10.86	36.34	<u>29.79</u>	9.03	0.71	17.07
HoRoPose	✓	✓	17.09	46.19	**30.68**	12.82	1.66	<u>21.85</u>
RoboPEPP	✓	✓	20.60	62.78	23.43	<u>18.63</u>	2.04	13.06
基础模型								
Opus-5.5	×	×	3.49 (±1.54)	0.00 (±0.00)	19.42	5.09 (±1.17)	0.00 (±0.00)	**27.26**
GPT-6-Astra	×	×	1.02 (±0.46)	0.00 (±0.00)	18.99	2.74 (±1.08)	0.00 (±0.00)	21.45
GPT-6.1-Sol	×	×	0.62 (±0.52)	0.00 (±0.00)	8.53	1.77 (±1.01)	0.00 (±0.00)	3.91
OVOK 中的基础模型								
Opus-5.5*	×	×	<u>28.17</u>	<u>73.75</u>	2.09	17.73	<u>4.50</u>	1.52
GPT-6-Astra*	×	×	25.19	67.20	1.77	17.00	4.31	1.67
GPT-6.1-Sol*	×	×	**40.32**	**78.54**	2.47	**23.81**	**5.88**	1.93

的公开实现。基础模型包括 Opus 5。5 [83], GPT-6 Astra [31], 和 GPT-6.1 Sol [105], , 均采用中等推理力度设置运行。我们还在 OVOK框架中运行这些模型。独立基础模型基线使用 10 次独立运行；OVOK运行 15 轮 RSI 迭代，并报告保留的最佳工作流。除非另有说明，定性比较默认使用 GPT-6.1 Sol 作为 OVOK骨干模型。所有实验均在沙箱中运行，主机配置为 AMD EPYC 7742（32 核）、一块 NVIDIA A800 GPU（80GB）、128GB 主机内存以及 1TB 存储空间。更多详情见附录C.1。

实验指标 对比实验主要考察位姿精度、关节角度精度和效率。我们报告的指标包括： $\mathcal{J}$中关键点的二维姿态 PCK（阈值为 2 像素）、三维姿态 ADD [106]（阈值为 20 毫米）、以度为单位的关节角度平均绝对误差，以及逐视频串行处理下的平均每秒帧数。对于 RSI 分析，我们则测量底层因素的准确性及效率：内参 K 采用 MAPE（平均绝对百分比误差）衡量，相机到基座位姿 $\mathbf{T}_{c\leftarrow b}$ 采用以米为单位的平均绝对误差衡量，关节角度采用以度为单位的平均绝对误差衡量，吞吐量则以每秒帧数衡量。

5.2. 机器人状态估计结果

表格 1报告了二维/三维姿态精度和吞吐量；表格2 报告了各关节角度的平均绝对误差。图4 和5 展示了用于比较的定性叠加图。

OVOK 在所评估的方法中实现了最先进的姿态精度和最低的平均关节角度误差，且无需额外输入。借助 GPT-6.1 Sol，其在 DREAM-V 上达到 40.32%/78.54% 的 PCK/ADD成功率，在 RoboKeyGen-V 上达到 23.81%/5.88%。这些提升是在

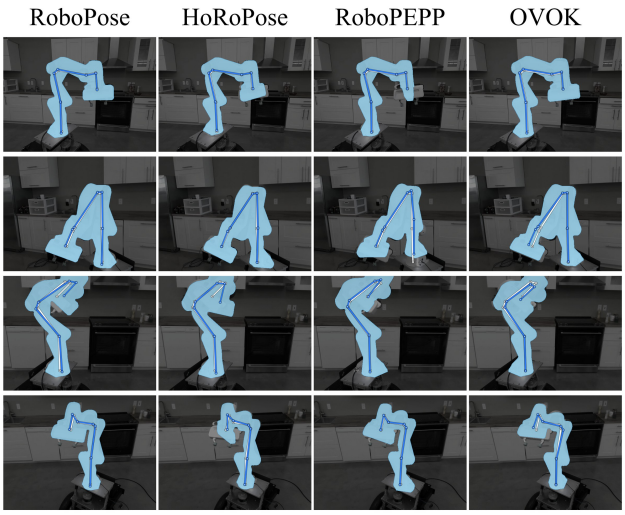

图4. DREAM-V 上的定性比较（从左至右：RoboPose、HoRoPose、RoboPEPP、使用 GPT-6.1 Sol 的 OVOK）。蓝色叠加表示预测结果；白色表示真实标注。

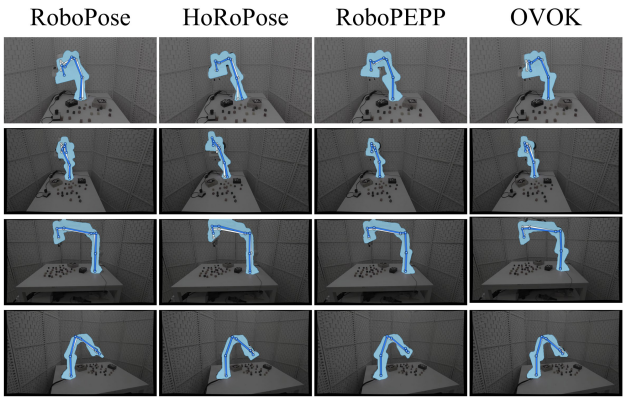

图5. RoboKeyGen-V 上的定性比较（从左到右：RoboPose、HoRoPose、RoboPEPP、使用 GPT-6.1 Sol 的 OVOK）。蓝色叠加层表示预测结果；白色表示真实标注。

吞吐量分别为 2.47 和 1.93 每秒帧数的情况下实现的。相对较高的延迟源于更复杂的任务设置和更长的多阶段推理流程。

单独使用基础模型查询在相同设置下几乎完全失效：ADD 成功率接近于零，平均关节角 MAE 仍达数十度。这表明其缺乏以关节机器人先验为基础的显式度量空间推理，因此直接推理无法确保运动学一致性，也无法恢复精确几何结构。

关节角度误差揭示了视觉估计变得困难的情况（表格2）。J1 依赖细微的基座外观线索，而随着机器人占地范围缩小，这些线索提供的信息会减少。在 RoboKeyGen-V 上，紧凑且接近极限的机械臂构型下的侧视图和自遮挡进一步增加了经DREAM训练的估计器的挑战，这与

表2。DREAM-V 和 RoboKeyGen-V 上的机器人关节角度估计。各项数据为每个关节的平均绝对误差，单位为度（越低越好）；Avg. 为所报告关节的平均值。 ✓：在测试时给定。 *：通过基础模型演化得到的 OVOK-solver。

方法	边界框	K	DREAM-V								RoboKeyGen-V							
			J1	J2	J3	J4	J5	J6	J7	平均值↓	J1	J2	J3	J4	J5	J6	J7	平均值↓
特定任务方法																		
RoboPose	×	✓	14.21	6.92	7.79	2.67	4.40	4.19	34.49	10.67	122.09	34.70	139.26	10.69	9.67	10.83	94.29	60.22
HoRoPose	✓	✓	17.31	3.75	7.24	3.42	4.99	6.09	31.03	10.55	108.34	32.62	126.76	**6.83**	5.70	7.12	19.68	43.86
RoboPEPP	✓	✓	66.13	**2.84**	4.24	**2.16**	4.00	**2.97**	–	–	119.47	34.12	133.47	13.69	10.54	9.56	–	–
基础模型																		
Opus-5.5	×	×	39.90 (±10.15)	23.71 (±13.87)	57.55 (±18.48)	25.43 (±18.71)	55.41 (±28.32)	43.05 (±9.15)	53.13 (±13.14)	42.60	127.64 (±16.85)	39.97 (±17.12)	124.41 (±13.78)	16.59 (±3.90)	43.92 (±12.26)	28.76 (±10.36)	33.82 (±14.76)	59.30
GPT-6-Astra	×	×	49.08 (±13.39)	21.27 (±6.22)	45.55 (±13.15)	18.21 (±4.97)	50.99 (±9.10)	47.38 (±14.38)	58.51 (±11.65)	41.57	125.26 (±21.96)	36.13 (±9.87)	135.46 (±20.35)	19.80 (±4.51)	47.03 (±13.53)	47.19 (±12.37)	42.36 (±15.58)	64.75
GPT-6.1-Sol	×	×	51.37 (±9.31)	27.87 (±5.52)	60.52 (±14.03)	19.66 (±5.17)	59.50 (±9.26)	46.99 (±10.36)	62.48 (±15.11)	46.91	110.27 (±8.25)	46.07 (±13.94)	117.60 (±6.68)	22.66 (±4.71)	49.46 (±10.33)	49.24 (±12.26)	48.36 (±13.28)	63.38
OVOK 中的基础模型																		
Opus-5.5*	×	×	17.54	5.48	**3.25**	3.12	3.93	4.33	14.95	7.51	27.71	**9.36**	13.27	12.35	15.20	11.60	17.57	15.29
GPT-6-Astra*	×	×	17.87	7.16	5.59	4.73	5.25	7.00	14.18	8.83	33.46	13.45	14.15	12.62	19.39	9.95	27.03	18.58
GPT-6.1-Sol*	×	×	**13.43**	6.26	3.82	3.54	**2.72**	3.85	**8.65**	**6.04**	**20.38**	9.83	**13.08**	7.38	**11.89**	**5.01**	**8.73**	**10.90**

运动学链观测不充分时较大的 J1–J3 误差相一致 [1, 19, 24]。OVOK 基于 OVOK-4k 构建，后者采样了多样的视角和场景 [23, 99], ，将平均关节角度误差降低至 6.04°（在 DREAM-V 上）以及 10.90°（在 RoboKeyGen-V 上）。在 RoboKeyGen-V 上，J1 和 J3 的改进尤为显著，其中 GPT-6.1 Sol 将平均绝对误差降低至 20.38° 和 13.08°。

在所有方法中，RoboKeyGen-V 上的位姿精度均显著下降（表格 1）。RoboKeyGen-V 上的相机–基座距离更大（1.29–1.82m），而 DREAM- synthetic 上的距离为（0.75–1.20m），后者是任务特定估计器所采用的通用训练集 [19, 22, 24]；由此产生的更小机器人占地范围和较弱的对应关系，使得在 20mm ADD 及 2像素 PCK 阈值下，不精确的定位代价更高。对于 OVOK，难度更为严格：二维/三维姿态精度由 K、$\mathbf{T}_{c\leftarrow b}$和 q共同决定，任何一个因素中的微小偏差都会通过投影被放大。即使采用结合几何工具与感知工具的进化工作流，高精度位姿仍然难以恢复。

定性比较展示了这些差异（图 4 和5）。在 DREAM-V 上，OVOK 通常能使投影后的机器人结构与观测到的机械臂更加贴合，包括连杆重叠导致关节难以区分的构型。在 RoboKeyGen-V 上，较小的机器人占地范围使基座对齐和机械臂形状方面的误差更加明显。OVOK 始终保持观测到的构型，与其在机器人位姿和关节角度准确率方面的提升相互补充。

5.3 RSI过程分析

图7 展示了三种 OVOK 骨干模型在 OVOK-4k 验证集上的 RSI 轨迹，图6 则展示了四个保留的 GPT-6.1 Sol* 在同一数据划分上的工作流。

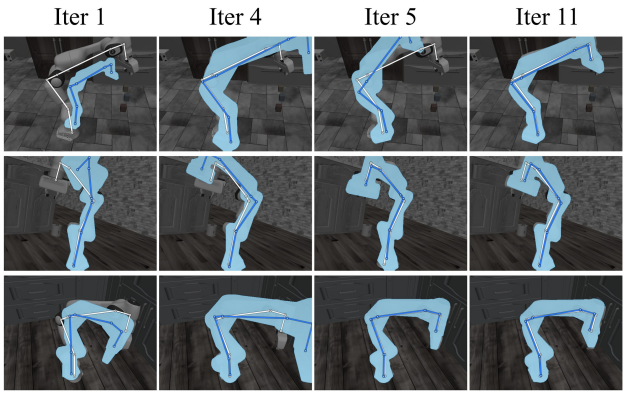

图 6. GPT-6.1 Sol* 在 OVOK-4k 验证集上的 RSI 工作流，列分别对应第 1、4、5 和 11 次迭代，三行对应三个片段。蓝色叠加层表示预测结果；白色表示真实标注。

不同骨干网络在相同的正则化与调度方案下采用不同的搜索策略。GPT-6.1 Sol* 最初将优化 K 与结合关节角度共同提升$\mathbf{T}_{c\leftarrow b}$ 视为基本相互独立的子任务，并行处理。在第 2–5 次迭代中，它修改内参工具以加快 K 收敛，同时冻结 DINOv3 [95]骨干网络，并训练一个 MLP 头来预测稀疏帧的关节角度、辅助二维关键点以及相机到基座的平移/旋转，从而获得精确的几何估计。在第 9 次和第 11 次迭代中，它转而冻结 HoRoPose [19]骨干网络，并微调 JointNet、RotationNet 和 KeypointNet，重点优化 J1 和 J7，从而同时提升准确性和效率。相比之下，GPT-6-Astra*在早期几乎将搜索重点全部放在内参上：它提出了许多以 K为中心的假设，并微调 Any-

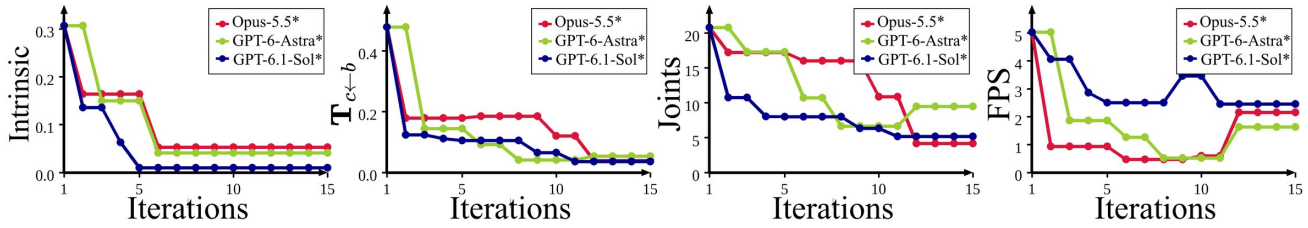

图7. 三种 OVOK 骨干模型（Opus-5.5*、GPT-6-Astra*、GPT- 6.1-Sol*）在 OVOK-4k验证集上经过15次迭代的 RSI轨迹。从左到右：相机内参MAPE、相机到基座位姿误差（m）、关节角MAE（°）和推理FPS。

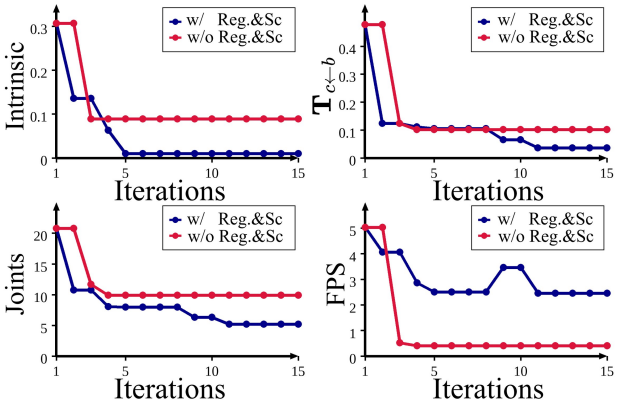

图8. 在 OVOK-4k验证集上，对 GPT-6.1 Sol* 的人工编写的正则化与调度方案进行消融。

标定 [104]解码器以实现精确标定，随后从第 5 次迭代开始转向 $\mathbf{T}_{c\leftarrow b}$ 和通过 HoRoPose/Robo-Pose 适配来处理关节，最后通过精简工作流复杂度来提高每秒帧数。Opus-5。5* 遵循了类似的分阶段路径。

尽管实现细节各不相同，成功的运行都遵循相同的宏观顺序：先锁定相机内 K参（可根据环境背景纹理轻松估计） [103, 104], ，然后优化$\mathbf{T}_{c\leftarrow b}$ 以及决定最终位姿精度的关节估计器，最后才为提高效率而精简工作流。如图6所示，GPT-6.1 Sol* 的早期工作流首先表现出较大的投影误差；后续修订逐步收紧骨架；而在更晚的迭代中，估计结果进一步优化，逐渐与观测到的手臂精确对齐。

5.4. 消融研究

如图 8所示，我们在 GPT-6.1 Sol* 的 OVOK-4k验证集上，对人工编写的正则化和调度方案进行了消融。移除这些控制后，早期的指标提升很快陷入平台期，运行过程也随之停滞；而采用调度方案的运行则继续降低 K、$\mathbf{T}_{c\leftarrow b}$以及关节误差。

具体而言，未受控运行从未重新发现第 5.3 节中的分阶段工作流。在最初三次迭代中，它将 RoboPEPP [22] 微调和 AnyCalib [104] 推断-方面的调整视为单一的精度捷径，而不是将内参与关节几何分开处理。在第 4–6 次迭代中，它又试图在几何准确率稳定之前削减计算量并提高每秒帧数，因此面向效率的修改未能通过验证。在第 6–8 次迭代中，它通过从弱初始化出发，交替细化 K、$\mathbf{T}_{c\leftarrow b}$ 和关节角来解决焦距与深度的歧义，这进一步加剧了这三个因素之间的欠定耦合。之后的迭代（第 9–14 次）加入了时间一致性模块。准确率提升仍然很小，而延迟却急剧上升，因此验证器拒绝了这些修改。纵观整个轨迹，智能体始终未能整合出一个经过规划的递归自我改进搜索。

这些行为凸显了无约束大语言模型搜索 [37, 47, 75, 102, 107–109] 的两种失效模式：(i) 归因错误：智能体无法找出验证报告背后的核心瓶颈；例如，当运行中期的几何误差仍然较大时，它会追求效率方面的修改，而不是提升估计质量；以及 (ii) 调度失败：从宏观上看，搜索缺乏基于物理原理的计划，例如 $K \rightarrow (\mathbf{T}_{c\leftarrow b}, q) \rightarrow$ 每秒帧数的分阶段聚焦方案；从更细的粒度来看，由失败驱动的局部修正会产生过于狭窄的假设，消耗迭代预算，却无法解决耦合估计问题。人工编写的失效模式词汇表限制了归因，而分阶段调度则确保了修改规划和粒度。

6. 结论

我们提出了 OVOK，这是一种用于基于视频的关节式机器人状态估计的递归式自我改进框架。在人工编写的协议下，初始化器构建 OVOK-4k，以闭合度量反馈循环；随后，提议者–执行者–验证者循环从一个轻量种子出发，在长时域内共同演进几何感知工具、感知工具及其工作流。保留的工作流在推理阶段被冻结，并在 DREAM-V 和 RoboKeyGen-V 上实现了最先进的度量精度和鲁棒性。OVOK 既为从视频中获取度量机器人状态提供了一条实用途径，也为在耦合物理约束下的其他几何感知问题中实现长时域自我改进提供了模板。

大语言模型使用声明

在论文准备过程中，我们全程使用 ChatGPT（OpenAI）和 DeepSeek 作为辅助工具，包括起草和重构文本、完善技术描述、改进语言表达、整理相关工作和 LaTeX 排版，以及针对论文提供审稿式反馈。在研究中，多模态基础模型（GPT-6.1 Sol、GPT-6 Astra 和 Claude Opus 5.5）被用作 OVOK 智能体骨干网络，并作为零样本基线（第 5 节）。基于大语言模型的助手还以辅助方式帮助完善作者编写的代码、实验分析、图表和合成数据协议。作者对方法、实验结果、解释及最终论文承担全部责任。

参考文献

[1] Timothy E. Lee、Jonathan Tremblay、Thang To、Jia Cheng、Terry Mosier、Oliver Kroemer、Dieter Fox 和 Stan Birchfield。基于单张图像的相机到机器人位姿估计。载于*IEEE* 国际机器人与自动化会议，第 9426–9432 页，2020 年。[2] Jingpei Lu、Florian Richter 和 Michael C. Yip。通过自监督仿真到现实迁移实现无标记的相机到机器人位姿估计。载于*IEEE* 计算机视觉与模式识别会议，第 21296–21306 页，2023 年。[3] Tianshu Wu、Jiyao Zhang、Shiqian Liang、Zhengxiao Han 和 Hao Dong。由基础特征驱动的在线末端执行器位姿估计：一种无标记且无需学习的方法。载于*IEEE* 国际机器人与自动化会议，2025 年。[4] Wenlong Huang、Chen Wang、Yunzhu Li、Ruohan Zhang 和 Li Fei-Fei。ReKep：用于机器人操作的关系关键点约束时空推理。载于第八届机器人学习会议论文集，第 4573–4602 页，2025 年。[5] Junhao Cai、Zetao Cai、Jiafei Cao、Yilun Chen、Zeyu He、Lei Jiang、Hang Li、Hengjie Li、Yang Li、Yufei Liu、Yanan Lu、Qi Lv、Haoxiang Ma、Jiangmiao Pang、Yu Qiao、Zherui Qiu、Yanqing Shen、Xu Shi、Yang Tian、Bolun Wang、Hanqing Wang、Jiaheng Wang、Tai Wang、Xueyuan Wei、Chao Wu、Yiman Xie、Boyang Xing、Yuqiang Yang、Yuyin Yang、Qiaojun Yu、Feng Yuan、Jia Zeng、Jingjing Zhang、Shenghan Zhang、Shi Zhang、Zhuoma Zhaxi、Bowen Zhou、Yuanzhen Zhou、Yunsong Zhou、Hongrui Zhu、Yangkun Zhu 和 Yuchen Zhu。InternVLA-A1：统一机器人操作中的理解、生成与行动。*arXiv* 预印本 *arXiv:2601.02456*，2026年。[6] Tianxing Chen、Yue Chen、Zixuan Li、Junyuan Tang、Kailun Su、Haoran Lu、Weijie Wan、Baijun Chen、Songling Liu、Haowen Yan、Honghao Su、Zhiyang Dou、Kaixuan Wang、Dandan Zhang、Yunze Liu、Yan Qin、Qiwei Liang、Qiwei Wu、Zijian Lin、Wenwei Lin、Yuran Wang、Minghua He、Tianshu Wu、Ruihai Wu、Jingquan Zhou、Kai-Chong Lei、Haibao Yu、Yuanfeng Ji、Weiyang Jin、Guanyu Lin、Xiaofan Li、Qi Xiong、Renjing Xu、Zhongyu Li、Wenhao Chai、Enze Xie、Ziwei Wang、Yao Mu、Hao Dong、Wojciech Matusik、Mingyu Ding、Wenbo Ding、Ping Luo 和 Masayoshi Tomizuka。RoboDojo：用于全面评估的统一仿真与现实基准

通用型机器人操作策略。*arXiv* 预印本 *arXiv:2607.04434*，2026年。[7] Yu Liu、Baoxiong Jia、Ruijie Lu、Chuyue Gan、Huayu Chen、Junfeng Ni、Song-Chun Zhu 和 Siyuan Huang。VideoArtGS：从单目视频构建关节式物体的数字孪生。*arXiv* 预印本 *arXiv:2509 17647*，2025年。。[8] Junhao Chen、Boran Zhang、Mingjin Chen、Henghaofan Zhang、Saining Zhang、Congcong Zhu、Hao Zhao、Ruqi Huang、Zhihao Li 和 Yufei Wang。一段视频，一个世界：将单目视频转化为物理四维场景。发表于欧洲计算机视觉会议，2026年。[9] Alexandros Delitzas、Chenyangguang Zhang、Alexey Gavryushin、Tommaso Di Mario、Boyang Sun、Rishabh Dabral、Leonidas Guibas、Christian Theobalt、Marc Pollefeys、Francis Engelmann 和 Daniel Barath。FunRec：从第一人称交互视频重建功能性三维场景。发表于 *IEEE* 计算机视觉与模式识别会议，第28446–28457页，2026年。[10]William Shen、Caelan Reed Garrett、Nishanth Kumar、Ankit Goyal、Tucker Hermans、Leslie Pack Kaelbling、Tom´as Lozano-P´érez 和 Fabio Ramos。可微分、GPU并行化的任务与运动规划。发表于机器人学：科学与系统会议（*RSS*）论文集，2025年。[11] Zhutian Yang、Caelan Garrett、Dieter Fox、Tom´as Lozano-P´érez 和 Leslie Pack Kaelbling。利用视觉语言模型指导长时域任务与运动规划。发表于*IEEE*国际机器人与自动化会议（*ICRA*）论文集，2025年。[12] Nishanth Kumar、William Shen、Fabio Ramos、Dieter Fox、Tom´as Lozano-P´érez、Leslie Pack Kaelbling 和 Caelan Reed Garrett。通过视觉语言模型生成的约束进行开放世界任务与运动规划。发表于<style id='27}机器人学习第九届会议（CoRL）论文集，2025年。[13]Denis Shcherba、Eckart Cobo-Briesewitz、Cornelius V. Braun 和 Marc Toussaint。利用语言模型进行任务与运动规划的元优化和程序搜索。发表于机器人学习第九届会议（*CoRL*）论文集，2025年。[14] Yoonwoo Kim、Raghav Arora、Roberto Mart´ín-Mart´ín、Peter Stone、Ben Abbatematteo 和 Yoonchang Sung。大型语言模型引导的部分可观测任务与运动规划状态估计。发表于*IEEE*国际机器人与自动化会议（*ICRA*）论文集，2026年。[15] Yixuan Huang、Bowen Li、Vaibhav Saxena、Yichao Liang、Utkarsh Aashu Mishra、Liang Ji、Lihan Zha、Jimmy Wu、Nishanth Kumar、Sebastian Scherer、Danfei Xu 和 Tom Silver。KinDER：面向机器人学习与规划的物理推理基准测试。发表于机器人学：科学与系统会议（*RSS*）论文集，2026年。[16] Matteo Merler、Bowen Li、Josh Roy、Yichao Liang、Qianwei Wang、Yixuan Huang 和 Tom Silver。用于广义任务与运动规划问题的编程智能体。*arXiv* 预印本 *arXiv:2609.30233*，2026年。[17] Jorge Mendez-Mendez。关于大型语言模型用于任务与运动规划的系统性研究——

与PDDL-Stream结合。arXiv预印本 arXiv:2510.00182，2025年。[18] Muyang Yan、Miras Mengdibayev、Ardon Floros、Weihang Guo、Lydia E. Kavraki 和 Zachary Kingston。利用视觉语言模型推理约束任务与运动规划。arXiv预印本 arXiv:2510.25548，2026年。[19] Shikun Ban、Juling Fan、Xiaoxuan Ma、Wentao Zhu、Yu Qiao 和 Yizhou Wang。具有未知状态的实时整体机器人位姿估计。发表于欧洲计算机视觉会议，2024年。[20] Kangjian Zhu、Haobo Jiang、Yigong Zhang、Jianjun Qian、Jian Yang 和 Jin Xie。MonoSE(3)-Diffusion：一种用于稳健相机到机器人位姿估计的单目SE(3)扩散框架。IEEE 机器人与自动化快报，10（11）：11832–11839，2025年。[21] Yang Tian、Jiyao Zhang、Zekai Yin 和 Hao Dong。机器人结构先验引导的时间注意力：基于图像序列的相机到机器人位姿估计。发表于IEEE计算机视觉与模式识别会议，第8917–8926页，2023年。[22] Raktim Gautam Goswami、Prashanth Krishnamurthy、Yann LeCun 和 Farshad Khorrami。RoboPEPP：通过嵌入预测式预训练进行基于视觉的机器人位姿与关节角度估计。发表于IEEE计算机视觉与模式识别会议，第6930–6939页，2025年。[23] Yang Tian、Jiyao Zhang、Guowei Huang、Bin Wang、Ping Wang、Jiangmiao Pang 和 Hao Dong。RoboKeyGen：通过基于扩散的三维关键点生成估计机器人位姿与关节角度。发表于IEEE国际机器人与自动化会议，2024年。[24] Yann Labb´e、Justin Carpentier、Mathieu Aubry 和 Josef Sivic。通过渲染与比较进行单视图机器人位姿与关节角度估计。发表于IEEE计算机视觉与模式识别会议，第1654–1663页，2021年。[25] Bowen Wen、Wei Yang、Jan Kautz 和 Stan Birchfield。FoundationPose：新物体的统一六维姿态估计与跟踪。发表于IEEE计算机视觉与模式识别会议，第17868–17879页，2024年。[26] Yi Li、Gu Wang、Xiangyang Ji、Yu Xiang 和 Dieter Fox。DeepIM：用于六维姿态估计的深度迭代匹配。发表于欧洲计算机视觉会议，第683–698页，2018年。[27] Liangjing Shao、Wanhao Liu、Zhiwei Fang、Rulin Zhou、Quanlu Zhang、Changjing Liu、Beilei Cui、Yiming Huang 和 Hongliang Ren。RoPE-Flow：通过二维图条件化三维流匹配进行基于单目视觉的机器人位姿估计。arXiv预印本 arXiv:2610.21086，2026年。[28] Yifan Liu、Fangneng Zhan、Wanhua Li、Haowen Sun、Katerina Fragkiadaki 和 Hanspeter Pfister。RoboTAG：通过拓扑对齐图进行端到端机器人构型估计。发表于IEEE计算机视觉与模式识别会议，2026年。[29] Gemini Robotics Team等。Gemini Robotics 1.5：通过先进的具身推理、思考和运动迁移，推动通用型机器人发展前沿。arXiv预印本 arXiv:2510.03342，2025年。[30] OpenAI。GPT-5.6 系统卡，2026年。https : //deploymentsafety.openai.com/gpt-5-6/gpt-5-6.pdf。

[31] OpenAI。《GPT-6 Astra 系统卡片》，2026年。https: / / deploymentsafety . openai . com / gpt - 6 - astra.[32] Anthropic。推出 Claude Fable 5.1 和 Claude Mythos 5.1。Anthropic 模型发布公告，2026年。https://www.anthropic.com/claude-fable-and-mythos-5-1.[33] Kirill Mazur、Nikita Karaev、Matthew Chang、Jitendra Ma- lik 和 Nur Muhammad Mahi Shafiullah。AgentSTAR：基于单目视频的智能体形状跟踪与重建。arXiv 预印本 arXiv:2609.24487，2026年。[34] Laurence Liang。使用现成基础模型的免训练机器人姿态估计。arXiv 预印本 arXiv:2512.06017，2025年。[35] Yi Han、Cheng Chi、Enshen Zhou、Shanyu Rong、Jingkun An、Pengwei Wang、Zhongyuan Wang、Lu Sheng 和 Shanghang Zhang。TIGeR：面向机器人学的视觉语言模型工具集成几何推理。载于 IEEE 国际机器人与自动化会议，2026年。[36] Jenny Zhang、Shengran Hu、Cong Lu、Robert Lange 和 Jeff Clune。Darwin g¨odel machine：自我改进智能体的开放式进化。载于国际学习表征会议，2026年。[37] Peng Xia 等。RRSI：智能体运行框架的正则化递归自我改进。arXiv 预印本 arXiv:2609.24972，2026年。[38] Maxime Robeyns、Martin Szummer 和 Laurence Aitchi- son。自我改进编码智能体。载于 ICLR 2025 自我改进基础模型研讨会，2025年。[39] Tom Brown、Benjamin Mann、Nick Ryder、Melanie Sub- biah、Jared D Kaplan、Prafulla Dhariwal、Arvind Neelakan- tan、Pranav Shyam、Girish Sastry、Amanda Askell、Sand- hini Agarwal、Ariel Herbert-Voss、Gretchen Krueger、Tom Henighan、Rewon Child、Aditya Ramesh、Daniel Ziegler、Jeffrey Wu、Clemens Winter、Christopher Hesse、Mark Chen、Eric Sigler、Mateusz Litwin、Scott Gray、Benjamin Chess、Jack Clark、Christopher Berner、Sam McCandlish、Alec Radford、Ilya Sutskever 和 Dario Amodei。语言模型是少样本学习器。载于神经信息处理系统进展，页码 1877–1901，2020年。[40] Eric Zelikman、Elia na Lorch、Lester Mackey 和 Adam Tauman Kalai。自学优化器（STOP）：递归自我改进的代码生成。载于语言模型会议，2024年。[41] Shengran Hu、Cong Lu 和 Jeff Clune。智能体系统自动化设计。载于国际学习表征会议，2025年。[42] Jiayi Zhang、Jinyu Xiang、Zhaoyang Yu、Fengwei Teng、Xionghui Chen、Jiaqi Chen、Mingchen Zhuge、Xin Cheng、Sirui Hong、Jinlin Wang、Bingnan Zheng、Bang Liu、Yuyu Luo 和 Chenglin Wu。AFlow：智能体工作流生成自动化。载于国际学习表征会议，2025年。[43] Hangfan Zhang、Shao Zhang、Kangcong Li、Chen Zhang、Yang Chen、Yiqun Zhang、Lei Bai 和 Shuyue Hu。自运行框架：能够自我改进的运行框架。arXiv 预印本 arXiv:2606.09498，2026年。[44] Jiahang Lin、Shichun Liu、Chengjun Pan、Lizhi Lin、Shi- han Dou、Zhiheng Xi、Xuanjing Huang、Hang Yan、Zhen- hua Han、Tao Gui 和 Yu-Gang Jiang。智能体运行框架

工程：观测驱动的编码智能体运行框架自动演化。arXiv 预印本 arXiv:2604 25850，2026年。 。[45] Yoonho Lee、Roshen Nair、Qizheng Zhang、Kangwook Lee、Omar Khattab 和 Chelsea Finn。Meta-Harness：模型运行框架的端到端优化。载于语言模型会议（COLM），2026年。[46] Yaorui Shi、Yuchun Miao、Yuxin Chen、Jiayuan Zhang、Yueqing Sun、Xierui Song、Xiang Wang 和 An Zhang。PluginRSI：利用可复用插件递归改进智能体运行框架。arXiv 预印本 arXiv:2609.32423，2026年。[47] Siwei Wu、Jincheng Ren、Yizhi Li、Haau-Sing Li、Chen- gran Yang、Xudong Ma、Yuxuan Zhang、Weicheng Gu、Jian Yang、Riza Batista-Navarro、Chuanyi Zhang、Xianglong Liu、Ming Zhou、Bryan Dai 和 Chenghua Lin。Modu- larRSI：模块化且具通用性的运行框架递归自我改进。arXiv 预印本 arXiv:2609.14857，2026年。[48] Ruiyao Xu、Yanfei Che n、Zhongying CuiZhu、Bhavana Dalvi Mishra、Yifei Ming、Han Yu、Rujun Han、Chen-Yu Lee 和 Tomas Pfister。Harness^2：面向开放世界的递归智能体运行框架。arXiv 预印本 arXiv:2610 14224，2026年。 。[49] Seth Karten、Joel Zhang、Tersoo Upaa、Ruirong Feng、Wen- zhe Li、Chengshuai Shi、Chi Jin 和 Kiran Vodrahalli。持续运行框架：自我改进基础智能体的在线适应。arXiv 预印本 arXiv:2605.09998，2026年。[50] Zhengyao Jiang、Dominik Schmidt、Dhruv Srikanth、Dix- ing Xu、Ian Kaplan、Deniss Jacenko 和 Yuxiang Wu。AIDE：代码空间中的人工智能驱动探索。arXiv 预印本 arXiv:2502.13138，2025年。[51] Dhruv Srikanth、Bingchen Zhao、Dixing Xu、Yuxiang Wu 和 Zhengyao Jiang。人工智能研究智能体的递归自我改进。arXiv 预印本 arXiv:2609.26457，2026年。[52] Alexander Novikov 等。AlphaEvolve：用于科学与算法发现的编码智能体。arXiv 预印本 arXiv:2506.13131，2025年。[53] Sibo Zhu、Shicheng Fan、Xinyue Wang、Wenyi Wu、Kun Zhou 和 Biwei Huang。RSIAgent：新环境中的递归自我改进自主探索。arXiv 预印本 arXiv:2609.15364，2026年。[54] Kairui Hu、Siyuan Hu、Fangzhou Hong、Zhaoxi Chen 和 Ziwei Liu。具身图灵机：用于机器人递归自我改进的有状态代码。arXiv 预印本 arXiv:2610.12369，2026年。[55] Zeren Chen、Xiaoya Lu、Zhijie Zheng、Pengrui Li、Lehan He、Yijin Zhou、Jing Shao、Bohan Zhuang 和 Lu Sheng。具有几何约束的空间推理智能体。载于 IEEE 计算机视觉与模式识别会议，页码 38689–38699，2026年。[56]Ryan Lopopolo。脚手架工程：在智能体优先的世界中利用 Codex。OpenAI 工程，2026年。https://openai.com/index/harness- engineering/.[57] Prithvi Rajasekaran。长时间运行应用开发的运行框架设计。Anthropic 工程，2026年。https://www.anthropic.com/engineering/harness-design-long-running-apps.

[58] Haoran Ye、Xuning He、Vincent Arak、Haonan Dong 和 Guojie Song。通过智能体技能演进进行元上下文工程。载于《第43届国际机器学习会议论文集》，第148325–148368页，2026年。[59]Qizheng Zhang、Changran Hu、Shubhangi Upasani、Boyuan Ma、Fenglu Hong、Vamsidhar Kamanuru、Jay Rain- ton、Chen Wu、Mengmeng Ji、Hanchen Li、Urmish Thakker、James Zou 和 Kunle Olukotun。智能体上下文工程：为自我改进语言模型演进上下文。载于《国际学习表征会议》，2026年。[60] Lakshya A. Agrawal、Shangyin Tan、Dilara Soylu、Noah Ziems、Rishi Khare、Krista Opsahl-Ong、Arnav Singhvi、Herumb Shandilya、Michael J. Ryan、Meng Jiang、Christo- pher Potts、Koushik Sen、Alexandros G. Dimakis、Ion Sto- ica、Dan Klein、Matei Zaharia 和 Omar Khattab。GEPA：反思式提示词演进可以胜过强化学习。载于《国际学习表征会议》，2026年。[61] Alex Zhang、Zhening Li 和 Omar Khattab。递归语言模型。载于《第四十届神经信息处理系统年会》，2026年。已接受发表。[62] Alex Zhang、Zhening Li 和 Omar Khattab。管理不善的天才假设。研究博客，2026年。https: // alexzhang13.github.io/blog/2026/mgh/。[63] Jingpei Lu、Zekai Liang、Tristin Xie、Florian Richter、Shan Lin、Sainan Liu 和 Michael C. Yip。CtRNet-X：在真实世界条件下使用单目相机进行相机到机器人位姿估计。载于《IEEE 国际机器人与自动化会议》，第1914–1920页，2025年。[64] Alessandro Simoni、Stefano Pini、Guido Borghi 和 Roberto Vezzani。利用深度图进行三维机器人位姿估计的半透视解耦热图。《IEEE 机器人与自动化快报》，7(4)：11569–11576，2022年。[65] Ornith团队。Ornith-1.5：从自我支架构建到自我改进。Ornith博客，2026年。https://ornith. ai/ornith_1_5.html。[66] Eric Zelikman、Yuhuai Wu、Jesse Mu 和 Noah D. Good- man。STaR：通过推理进行推理自举。载于《神经信息处理系统进展》，2022年。[67] Weizhe Yuan、Richard Yuanzhe Pang、Kyunghyun Cho、Xian Li、Sainbayar Sukhbaatar、Jing Xu 和 Jason We- ston。自我奖励语言模型。载于《国际机器学习会议》，2024年。[68] DeepSeek-AI。DeepSeek-R1：通过强化学习激励大语言模型的推理能力。《自然》，645：633–638，2025年。[69] Andrew Zhao、Yiran Wu、Yang Yue、Tong Wu、Quentin Xu、Matthieu Lin、Shenzhi Wang、Qingyun Wu、Zilong Zheng 和 Gao Huang。绝对零：利用零数据进行强化自博弈推理。载于《神经信息处理系统进展》，2025年。[70] Chengsong Huang、Wenhao Yu、Xiaoyang Wang、Hong- ming Zhang、Zongxia Li、Ruosen Li、Jiaxin Huang、Haitao Mi 和 Dong Yu。R-Zero：从零数据开始的自我演进推理大语言模型。arXiv预印本 arXiv:2508.05004，2025年。[71] Zehan Qi、Xiao Liu、Iat Long Iong、Hanyu Lai、Xueqiao Sun、Wenyi Zhao、Yu Yang、Xinyue Yang、Jiadai Sun,

Shuntian Yao、Tianjie Zhang、Wei Xu、Jie Tang 和 Yux- iao Dong。WebRL：通过自我演进的在线课程强化学习训练大语言模型网页智能体。载于《国际学习表征会议》，2025年。[72] Lilian Weng。面向自我改进的脚手架工程。Lil'Log，2026年。https://lilianweng.github.io/posts/2026-07-04-harness/。[73] Meng Luo、Zechen Bai、Shanqing Xu、Yanlin Li、Kaixin Li、Mengchen Qian、Weifeng Lu、Ran Zhang、Yingtian Zou、Ziyang Luo、Wei Chen 和 Hongzhan Lin。迈向物理人工智能中的递归式自我改进。arXiv预印本 arXiv:2610.29677，2026年。[74] Haoyu Zhao、Zhengxu Yu、Zhiyuan He、Meng Fang、Rasul Tutunov、Haitham Bou-Ammar、Weilin Luo 和 Jun Wang。Memento 3：通过反思式规则手册实现基于模型的递归式自我改进。arXiv预印本 arXiv:2610.11794，2026年。[75] Mert Yuksekgonul、Federico Bianchi、Joseph Boen、Sheng Liu、Pan Lu、Zhi Huang、Carlos Guestrin 和 James Zou。通过反向传播语言模型反馈优化生成式人工智能。《自然》，639(8055)：609–616，2025年。[76] Mirac Suzgun、Mert Yuksekgonul、Federico Bianchi、Dan Jurafsky 和 James Zou。动态备忘单：利用自适应记忆进行测试时学习。载于《第19届计算语言学协会欧洲分会会议论文集》（第1卷：长篇论文），第7080–7106页，2026年。[77] Guanzhi Wang、Yuqi Xie、Yunfan Jiang、Ajay Mandlekar、Chaowei Xiao、Yuke Zhu、Linxi Fan 和 Anima Anandku- mar。Voyager：拥有大语言模型的开放式具身智能体。《机器学习研究汇刊》，2024年。[78] Chengyu Yang、Yunxiang Zheng、Zongzheng Li、Jia mei Li、Boxun Lan、Xiaoyu Liu、Xiaoyu Zhang、Xiaohong Chen 和 Ming Hu。通过程序化编译消除不确定性：自我演进的智能体技能。arXiv预印本 arXiv:2610.39508，2026年。[79] Zexi Li、Yehang Zhang、Wenqian Li、Haojian Huang、Chenxu Wang、Shiyuan Deng、Yangkai Wei、Tianyi Zhang、Binghui Xie、Bohan Zhou、Yifan Chang、Kaiwen Zhou、Ying-Cong Chen、James Cheng 和 Yinchuan Li。Robo-Harness K1：通过感知增强驾驭机器人操作智能体。arXiv预印本 arXiv:2609.29389，2026年。[80]Yanzhe Chen、Zechen Bai、Zhijun Cao、Wenzheng Zeng、Kevin Qinghong Lin、Yiqi Lin、Guoqiang Liang、Kevin Yuchen Ma、Qiming Huang 和 Mike Zheng Shou。Show-Harness：仅凭一个视觉语言模型智能体就能操作机器人。arXiv预印本 arXiv:2609.10522，2026年。[81] Kai Glantz 和 Clemens Grange。看见并不等于测量：面向视觉语言模型的工具增强度量空间推理。arXiv预印本 arXiv:2609.29073，2026年。[82] Achu Menon、Sravanthi Machcha、Sabrina Zou、Tzu Kit Chan 和 Jay Chooi。GPT-6 Astra 论机器人操作。Robocurve，2026年。https://openai.robocurve. org/gpt-6-astra/。[83] Anthropic。Claude Opus 5.5 发布公告。Anthropic模型发布公告，2026年。https://www.anthropic.com/claude-opus-5-5。[84] DeepSeek-AI。DeepSeek-V4：迈向高效的百万令牌上下文智能。arXiv 预印本 arXiv:2606.19348，2026年。[85] Bingran You、Wenjun Ke、Ria Rosenauer、Bo Liu、Wenbo Chen、Min Chen、Shuqi Xu、Yue Ma 等。FrontierPhysics：评估端到端物理学研究智能体。arXiv 预印本 arXiv:2611.06910，2026年。[86] Boyuan Chen、Zhuo Xu、Sean Kirmani、Brian Ichter、Danny Driess、Pete Florence、Dorsa Sadigh、Leonidas Guibas 和 Fei Xia。SpatialVLM：赋予视觉语言模型空间推理能力。载于 IEEE 计算机视觉与模式识别会议，2024年。[87] An-Chieh Cheng、Hongxu Yin、Yang Fu、Qiushan Guo、Ruihan Yang、Jan Kautz、Xiaolong Wang 和 Sifei Liu。SpatialRGPT：视觉语言模型中基于视觉依据的空间推理。载于 Adv. Neural Inform. Process. Syst.，2024年。[88] Xingyu Fu、Yushi Hu、Bangzheng Li、Yu Feng、Haoyu Wang、Xudong Lin、Dan Roth、Noah A. Smith、Wei-Chiu Ma 和 Ranjay Krishna。BLINK：多模态大语言模型能够看见，却无法感知。载于欧洲计算机视觉会议，2024年。[89] Kailin Wang、Haoxiang Jie、Yaoyuan Yan、Zhiyou Heng 和 Zhaosong Li。RegenHarness：一个具备证据门控递归自我改进能力的机器人智能体框架。arXiv 预印本 arXiv:2609.27612，2026年。[90] Vincent Lepetit、Francesc Moreno-Noguer 和 Pascal Fua。EPnP：一种精确的 $O(n)$ PnP问题解法。Int. J. Comput. Vis.，81(2)：155–166，2009。[91]Shuzheng Gao、Chaozheng Wang、Cuiyun Gao、Xiao-qian Jiao、Chun Yong Chong、Shan Gao 和 Michael Lyu。提示词炼金术士：用于测试用例生成的自动化面向大语言模型的提示词优化。arXiv 预印本 arXiv:2501.01329，2025年。[92] Runyu Lu、Yubo Wu、Ethan Kou、Haotian Lin、Letian Fu、Wenli Xiao、Ajay Mandlekar、Yinzhen Xu、Guanya Shi、Ken Goldberg、Ang Chen、Mosharaf Chowdhury、Jan Kautz、Yuke Zhu、Linxi Fan 和 Guanzhi Wang。ASPIRE：面向机器人学的智能体技能发现。arXiv 预印本 arXiv:2607.00272，2026年。[93] Alex Zhang 和 Omar Khattab。语言模型框架是组合泛化器。MIT OASYS 研究博客，2026年。https://alexzhang13.github.io/blog/ 2026/harness/。[94] Haotong Lin、Sili Chen、Junhao Liew、Donny Y. Chen、Zhenyu Li、Yang Zhao、Sida Peng、Hengkai Guo、Xiao wei Zhou、Guang Shi、Jiashi Feng 和 Bingyi Kang。Depth Anything 3：从任意视角恢复视觉空间。载于 Int. Conf. Learn. Represent.，2026年。[95] Oriane Sim´eoni、Huy V. Vo、Maximilian Seitzer、Federico Baldassarre、Maxime Oquab、Cijo Jose、Vasil Khalidov、Marc Szafraniec、Seungeun Yi、Micha¨el Ramamonjisoa、Francisco Massa、Daniel Haziza、Luca Wehrstedt、Jianyuan Wang、Timoth´ee Darcet、Th´eo Moutakanni、Leonel Sen-tana、Claire Roberts、Andrea Vedaldi、Jamie Tolan、John Brandt、Camille Couprie、Julien Mairal、Herv´e J´egou、Patrick Labatut 和 Piotr Bojanowski。DINOv3。《Transactions on Machine Learning Research》，2026年。

[96] Adrien Bardes、Quentin Garrido、Jean Ponce、Xinlei Chen、Michael Rabbat、Yann LeCun、Mahmoud Assran 和 Nico-las Ballas。重新审视特征预测：从视频中学习视觉表征。载于 Int. Conf. Learn. Represent.，2025年。[97] Konstantinos Bousmalis、Giulia Vezzani、Dushyant Rao、Coline Devin、Alex X. Lee、Maria Bauza、Todor Davchev、Yuxiang Zhou、Agrim Gupta、Akhil Raju、Antoine Lau-rens、Claudio Fantacci、Valentin Dalibard、Martina Zam-belli、Murilo Martins、Rugile Pevceviciute、Michiel Blokz-ijl、Misha Denil、Nathan Batchelor、Thomas Lampe、Emilio Parisotto、Konrad ˙Zołna、Scott Reed、Sergio G´omez Col-menarejo、Jon Scholz、Abbas Abdolmaleki、Oliver Groth、Jean-Baptiste Regli、Oleg Sushkov、Tom Roth¨orl、Jos´e En-rique Chen、Yusuf Aytar、Dave Barker、Joy Ortiz、Mar-tin Riedmiller、Jost Tobias Springenberg、Raia Hadsell、Francesco Nori 和 Nicolas Heess。RoboCat：一种能够自我改进的机器人操作通用智能体。《Transactions on Machine Learning Research》，2023年。[98] Ilia Kulikov、Chenxi Whitehouse、Tianhao Wu、Yixin Nie、Swarnadeep Saha、Eryk Helenowski、Weizhe Yuan、Olga Golovneva、Jack Lanchantin、Yoram Bachrach、Jakob Foer-ster、Xian Li、Han Fang、Sainbayar Sukhbaatar 和 Jason Weston。Autodata：用于创建高质量合成数据的智能体数据科学家。arXiv 预印本 arXiv:2606 25996，2026年。。[99] Soroush Nasiriany、Sepehr Nasiriany、Abhiram Maddukuri 和 Yuke Zhu。RoboCasa365：用于训练和基准测试通用型机器人的大规模仿真框架。载于 Int. Conf. Learn. Represent.，2026年。[100] Dongzhou Cheng、Taoran Yi、Ye Fang、Xing wu Zhang、Fan Feng、Yixuan Li、Gengxiong Zhuang、Rongze Wang、Shuai Yang、Wei Song、Weizhi Xue、Minyan Wu、Jie Gui、Jiaqi Wang 和 Tong Wu。使用视觉语言模型智能体进行上下文内机器人学习。arXiv 预印本 arXiv:2609.19138，2026年。[101] Johannes Moll、Jean-Philippe Corbeil、Jiazhen Pan、Martin Hadamitzky、Daniel Rueckert、Lisa C. Adams 和 Keno Bressem。GRASP：面向自我改进大语言模型智能体的门控回归感知技能提议器。载于《2026年自然语言处理经验方法会议论文集》。计算语言学协会，2026年。[102] Suizhi Huang、Mei Li、Han Yu 和 Xiaoxiao Li。TextRes-Net：通过深度残差调优，在复合人工智能系统中解耦并路由优化信号。载于 Int. Conf. Mach. Learn.，2026年。[103] Alexander Veicht、Paul-Edouard Sarlin、Philipp Linden-berger 和 Marc Pollefeys。GeoCalib：通过几何优化学习单图像标定。载于欧洲计算机视觉会议，2024年。[104] Javier Tirado-Gar´ın 和 Javier Civera。AnyCalib：面向模型无关单视图相机标定的流形学习。载于 Int. Conf. Comput. Vis.，第 8044 8055 页，2025年 –。[105] OpenAI。GPT-6 Astra 系统卡片附录：GPT-6.1 Sol，2026年。https://deploymentsafety.openai.com/gpt-6-1-sol。

[106] Stefan Hinterstoisser、Vincent Lepetit、Slobodan Ilic、Stefan Holzer、Gary Bradski、Kurt Konolige 和 Nassir Navab。基于模型的训练、检测以及在高度杂乱场景中对无纹理三维物体的姿态估计。载于 计算机视觉——*ACCV 2012*，第 548–562 页，2013年。[107] Michael Nguyen、Wei Chen Tan、Nurul Aisyah Hassan、Arvind Raman、Li Hua Lim、Ahmad Faiz Razak 和 Jia Hui Wong。智能体框架优化的价值何在？自进化大语言模型智能体中的局部收益与预算拆分陷阱。*arXiv* 预印本 *arXiv:2609.02889*，2026年。[108] Yihua Zhu、Qianying Liu、Weixu Qiao、Xuan Ren、Weiwei Xu、Wenbo Li、Wei Wang、Ruijia Chen、Xinmiao Luan、Yin Luo、Hao Huang、Xiang Zheng 和 Hidetoshi Shimodaira。我的 FAULT：将自我诊断作为自进化智能体强化学习中的信用分配。*arXiv* 预印本 *arXiv:2610.01161*，2026年。[109] Qianshu Cai、Yonggang Zhang、Jun Nie、Maohao Ran、Hua-jiang Zheng、Jun Song、Xinmei Tian、Yike Guo 和 Wei Xue。智能体运行框架可靠自进化理论。*arXiv* 预印本 *arXiv:2609.08175*，2026年。[110] Bin Lin、Yang Ye、Bin Zhu、Jiaxi Cui、Munan Ning、Peng Jin 和 Li Yuan。Video-LLaVA：通过投影前对齐学习统一视觉表征。载于自然语言处理经验方法会议论文集，2024年。[111] Gedas Bertasius、Christoph Feichtenhofer、Du Tran、Jianbo Shi 和 Lorenzo Torresani。从稀疏标注视频中学习时序姿态估计。载于神经信息处理系统进展，2019年。[112] Cheng Dai、Jiale Lin、Bingxuan Song、Yifei Chen、Jiashuo Chen、Xin Yuan 和 Fanglin Bao。基于 HADAR 的热红外高光谱图像复原，2026年。[113] Cheng Dai、Jiale Lin、Hongyi Xu、Bingxuan Song、Ziyang Xie 和 Fanglin Bao。TeX-1500：用于温度-发射率-纹理分解的配对真实世界 LWIR 高光谱数据集与基准测试，2026年。[114] Jiashuo Chen、Cheng Dai、Yanan Hu 和 Fanglin Bao。T^2纹理：稀疏扰动的热纹理成像，2026年。[115] Guanfeng Tang、Zhiyuan Wu、Jiahang Li、Ping Zhong、We Ye、Xieyuanli Chen、Huiming Lu 和 Rui Fan。TiCoSS：在联合学习框架中强化语义分割与立体匹配之间的耦合，2025年。

OVOK：递归式自我改进智能体框架：基于视频的关节式机器人状态估计

补充材料

目录。

•A 初始化
•B 演化– B.1 人工初始化– B.2 目标与技能– B.3 提议器设置– B.4 验证器设置

•C 实验– C.1 实验设置– C.2 局限性与未来工作

表格 4. OVOK-4k 准入门槛。

组别	Gate
可见性	机器人掩码 $\geq 2\%$的 640×480 (每帧) 屏幕外顶点 $\leq 40\%$ 机器人遮挡 $\leq 15\%$；底座遮挡 $\leq 0.5\%$ 间隙 ≥ 0.15 m（场景）/ 0.20 m（机器人） 距离 0.85–2.55 m；基座部分处于视野内
物理正向运动学	$\leq 10^{-4}$ m；重投影 ≤ 0.1px 穿透 $= 0$；动力学回放 $\|\Delta q\| < 10^{-5}$ 片段内基座/相机静止 $f_x{=}f_y \in [298, 625]$； $c_x \in [300, 340]$, $c_y \in [160, 280]$
运动	≥ 3 个关节，范围为 > 0.08弧度，以及手 ≥ 0.08米 或腕部族方向变化为 ≥ 0.35弧度
多样性	≥ 40 训练集 / 每个族群12个验证片段 相机网格占用率 $\geq 50\%$ ≤ 1每次滚动执行的接受窗口

A. 初始化

初始化构建了数据基础，供后续工具修订和工作流评分使用。不同于以往的特定任务数据集 [1, 23, 28], , OVOK-4k 着重涵盖更具挑战性的观测：变化的内参 K、多样的视点和距离、自遮挡、较小的屏幕占用面积、截断，以及片段中复杂的光照。表格 3 比较了 OVOK-4k 与两种主导以往 Franka RGB 估计器的合成语料库：DREAM-synth [1]（也是 RoboPose、HoRoPose 和 RoboPEPP 通常使用的训练来源 [19, 22, 24]）以及 RoboKeyGen 的 SimRGBD-Franka [23]。

按照图中的人工编写协议。9，基础模型初始化器会实例化 RoboCasa [99]仿真与渲染流水线（50 种布局 $\times$ 50 种风格），并生成成对的 RGB / 模拟器状态片段，其中 Franka Panda 的基座在每个片段中固定于世界坐标系。每个通过筛选的片段均为 640×480 uint8 RGB，采样率为 25 赫兹，共 75 帧；在同一片段内， K 和$\mathbf{T}_{c\leftarrow b}$ 保持不变，但在不同片段间会变化。训练集（3,600）和验证集（400）使用互不重叠的房间布局 ID；只有满足数值准入门槛的片段才会纳入 OVOK-4k（见表4）。

所有代码和数据集将于 2026年 12 月前公开，网址：https://github.com/dccc2025/ovok_官方。

B. 演化

B.1. 人工初始化

演化始于一个轻量级人工编写种子（$\mathcal{U}_0$， $\mathcal{S}_0$， $\mathcal{W}_0$）。整个运行过程中，该种子始终保持冻结，从而避免目标随着技能和工具的演化而漂移。它为提议者–执行者–验证者循环指明方向，但既不干预探索，也不规定最终保留的求解器。

初始化技能。 种子技能让提议器能够在尚无任何验证证据时理解任务。 数学问题定义(图 10) 以简洁的几何与运动学形式阐述估计目标。 背景与动机(图 11) 以结构化语言概述先前工作，并记录了这样一项结构先验：关节状态估计可以模块化地分解为多个子问题。两者共同明确了要估计的内容，以及如何组合工作流的先验。

初始化工具。初始化还会在三个工具库中生成一组精简工具 $\mathcal{U}_0$： 可变工具、冻结的预训练模型和基础模型视觉 [35, 46, 79, 81]。可变工具是可编辑的几何与系统原语：标定、PnP、联合优化、渲染、分割、跟踪和测试。冻结模型提供感知和

表3. 用于Franka RGB机器人状态估计的合成训练集。

	DREAM-synth [1]	SimRGBD-Franka [23]	OVOK-4k
规模	∼100k 图像	4k × 3 帧	4k × 75 帧
场景	随机化三维场景	Blender场景	RoboCasa 厨房
时间单位	单个样本	3帧片段	75帧片段（25赫兹）
相机距离	0.75–1.20 米	未指定（测试时为 1.29–1.82 米）	0.85–2.55 米
内参 K	已知	已知	未知；$f_x{=}f_y \in [298, 625]$, $c_x \in [300, 340]$, $c_y \in [160, 280]$
运动	独立关节样本	模拟运动	门控展开（表 4）
监督	二维关键点	位姿、关节、掩码、深度	RGB + $(q, K, \mathbf{T}_{c\leftarrow b})$

任务专用预测器，工作流可调用这些预测器：视觉骨干网络 Depth Anything 3 [94], DINOv3 [95], 和 V-JEPA 2 [96], , 以及机器人估计器 RoboPose [24],HoRoPose [19], 和 RoboPEPP [22]。这些权重保持冻结；演化过程可以添加适配器，或在训练划分上训练小型任务专用头，可训练参数最多为 25M [97, 98]。基础模型视觉提供轻量级问答接口，用于视觉检查或粗略的零样本初始化。

初始化工作流。种子工作流 $\mathcal{W}_0$ 十分简洁（主论文中的算法 1 ）。AnyCalib [104] 根据第一帧估计内参；随后 HoRoPose [19]在固定的 K条件下恢复逐帧位姿和关节角度，之后 $\hat{\mathbf{T}}_{c\leftarrow b} = \frac{1}{N}\sum_{t=1}^{N} \hat{\mathbf{T}}_{c\leftarrow b,\ t}$。该工作流为演化提供了一个可运行的端到端基线，以及衡量 J_τ的起点。

B.2. 目标与技能

公式 2 仅针对工具 $\mathcal{U}$ 和工作流 $\mathcal{W}$ 最大化 J_τ 。验证分数为 EXECUTE(X_v;$\mathcal{U}, \mathcal{W}$)，因此 J_τ 取决于可执行对；技能 $\mathcal{S}$ 并非该指标的参数。 $\mathcal{S}$ 作为搜索记忆保留：当某次修订提升了 J_τ，已接受的假设会合并到 $\mathcal{S}$ 中，并由下一 REASON阶段使用，而推理仍运行 EXECUTE(X; $\mathcal{U}^\star, \mathcal{W}^\star$)。主论文中的算法 2 明确说明了这种区分。

B.3. 提议器设置

如主论文所述，每个演化轮次运行三个提议器阶段：REASON、REVISE 和 COMPOSE（算法 3）。REASON 读取先前的验证报-告 r_{i-1} ，判断哪些编辑有所帮助，并输出 M 假 - 设 $\{\delta_m\}_{m=1}^{M}$。REVISE 根据这些假 - 设 $\mathcal{D}'_{\text{tr}}$ 在设计划分上更新工具：它可以重写可变工具、微调适配器，或训练小型任务专用模-$\widetilde{\ }$型。随后，COMPOSE 从 $\mathcal{W}$ ∼$\{\delta_m\}$ 和修订后的工具集 $\mathcal{U}$构建候选工作流。

为提高 RSI 的效率，我们不会将每个假设都直接送去进行完整验证。3,600 个片段组成的训练集 $\mathcal{D}_{\text{tr}}$ 在开始时只划分一次，分为设计划分 $\mathcal{D}'_{\text{tr}}$ (3,550 个片段) 和固定的试运行探测集 $\mathcal{D}_{\text{dry}}$ (50 个片段)；两个集合互不重叠，并在各次迭代中保持固定。REVISE 基于 $\mathcal{D}'_{\text{tr}}$进行拟合；若某个假设未能在$\mathcal{D}_{\text{dry}}$ (算法3)[42, 60, 101]上取得改进，则会被提前淘汰。只有得到试运行探测支持的候选项，才会进入组合阶段和验证阶段执行器–验证器循环。

算法 3 带训练集试运行拒绝的提议器

要求： 任务 τ、($\mathcal{U}$、 $\mathcal{S}$、 $\mathcal{W}$)、报告 r、训练集 $\mathcal{D}'_{\text{tr}}$和 $\mathcal{D}_{\text{dr}}$ $_{\text{y}}$
确保： 修订后的工具 $\widetilde{\mathcal{U}}$，工作流 $\widetilde{\mathcal{W}}$，已接受的假设 $\mathcal{H}$

$J_{\text{base}} \leftarrow$ 得分($\mathcal{W}, \mathcal{U}; \mathcal{D}_{\text{dry}}$)
$\{\delta_m\}_{m=1}^{M} \leftarrow$ REASON(τ, $\mathcal{S}$, $\mathcal{U}$, $\mathcal{W}$, r, M)
$\mathcal{H} \leftarrow \varnothing$, $\widetilde{\mathcal{U}} \leftarrow \mathcal{U}$
对于 $m = 1, \ldots, M$ **执行**
  $\mathcal{U}^{(m)} \leftarrow$ 修订 ($\tau, \delta_m, \mathcal{U}, \mathcal{D}'_{\text{tr}}$)
  $\mathcal{W}^{(m)} \leftarrow$ COMPOSE(τ, δ_m, $\mathcal{U}^{(m)}$, $\mathcal{W}$)
  $J^{(m)} \leftarrow$ 分数($\mathcal{W}^{(m)}$, $\mathcal{U}^{(m)}$; $\mathcal{D}_{\text{dry}}$)
  如果 $J^{(m)} > J_{\text{base}}$ **则**
    $\mathcal{H} \leftarrow \mathcal{H} \cup \{\delta_m\}$
    $\widetilde{\mathcal{U}} \leftarrow \mathcal{U}^{(m)}$, $J_{\text{base}} \leftarrow J^{(m)}$
  结束如果
结束循环
$\widetilde{\mathcal{W}} \leftarrow$ 编排 (τ, $\mathcal{H}$、 $\mathcal{U}, \mathcal{W}$)
返回 $\widetilde{\mathcal{U}}, \widetilde{\mathcal{W}}, \mathcal{H}$

多项接受。 算法 $J_i > J^\star$ 和$J^{(m)} > J_{\text{base}}$ 中的比较2和3决定是否保留一个假设。单个标量可能导致搜索过早停滞，或掩盖 K、$\mathbf{T}_{c\leftarrow b}$和 q之间的补偿误差。在运动学图像模型 $\mathbf{p} = \pi(K, \mathbf{T}_{c\leftarrow b}, \text{FK}(q))$的指导下，我们保留一个未归一化的四项向量

$$\mathbf{J} = (J_1, J_2, J_3, J_4), \tag{4}$$

其中， J_1 是 K的相对误差， J_2 是$\mathbf{T}_{c\leftarrow b}$的平均绝对误差， J_3 是关节角度的平均绝对误差（以上指标均越低越好），而J_4 是串行端到端帧率（越高越好）。相对改进和退化均遵循这些方向。只有同时满足以下条件时，假设才标记为受支持：(i) 至少一项指标改善超过 20%，或至少两项指标

提升幅度超过 10%，而其余每个术语的表现下降幅度均小于 2%；（ii）串行帧率至少保持在 0.3。

编辑计划。 上述门控在整个运行过程中固定不变；但接下来 REASON REASON 应尝试什么则并非如此。人工制定的计划（图 13）会建议编辑优先级，但不设定僵化的迭代次数配额。首先锁定 K，它可以根据背景纹理进行估计，也是恢复度量尺度的前提。一旦判定 K 足够可靠，就以$\mathbf{T}_{c\leftarrow b}$ 和关节估计器为目标，因为它们对最终位姿精度影响最大。之后，只有在准确率不会下降时才提高每秒帧数。提出假设时，REASON 会结合这份计划与先前的报告 r_{i-1}。

B.4. 验证器设置

执行器在 $\mathcal{D}_{\text{v}}$ 上运行组合工作流后，验证器会返回分数 $\mathbf{J}_i$ 和报告 r_i（算法 2）。如图 3所示，报告包含三个部分——定量指标、定性诊断和失败归因，并据此引导下一轮提议器工作。这遵循了近期 RSI 实践：由验证而非标量训练损失来指导后续修订 [36, 48, 51, 58–60, 75]。

在定量方面，验证器会在验证集划分上评估 $\mathbf{J} = (J_1, J_2, J_3, J_4)$，并使用仅用于诊断的辅助二维/三维姿态统计数据。只有满足第 B.3 节中的多项规则，才会保留修订版本。在定性方面，验证器会抽取高误差片段并诊断其失败模式。归因随后将这些模式映射到八种典型模式，且仅限于以下模式：内参不准确、严重 $\mathbf{T}_{c\leftarrow b}$ 偏差、全局关节角度偏差、困难的基座朝向（J1）、困难的末端执行器（J7）、遮挡/部分可辨识性、光照以及计算效率低。

下一阶段的 REASON 会读取 r_i 以及第B.3节中的编辑计划。因此，指标决定是否保留修订版本，而报告决定下一步要修改什么。

C. 实验

C.1. 实验设置

所有实验均在沙箱中运行，使用 AMD EPYC 7742 的 32 个核心、一块 NVIDIA A800 GPU（80GB）、128 GB主机内存，以及 1TB 存储空间上限（见第 5.1 节）。此处我们补充基础模型实验。独立基线通过官方 API 调用 Opus 5.5、GPT-6 Astra 和 GPT-6.1 Sol，并采用中等推理力度设置；每种构型重复运行 10 次。同一组 API 也用作 OVOK 骨干模型：我们运行 15 次 RSI 迭代，并在 DREAM-V 和 RoboKeyGen-V 上评估保留的最佳工作流。除非另有说明，定性比较和消融实验均使用 GPT-6.1 Sol。

稀疏采样基础模型推理。视频视觉语言模型在稀疏采样帧上进行查询，而非以原生帧率 [110], ，视频姿态估计器则通过时间传播，将这些关键帧稠密化 [111]。为公平比较，我们采用相同的协议：每隔 10 帧查询一次基础模型，并保持每次估计结果，直到下一个关键帧（算法4）。视频处理和关键帧查询均按串行方式进行；报告的每秒帧数使用纯处理毫秒数（API计算时间）。图15 和16 展示了在 DREAM-V 和 RoboKeyGen-V 上使用的提示；每条提示都要求提供查询帧中的七个 URDF连杆二维关键点和七个关节角度。

算法4 稀疏采样基础模型推理

要求： 视频 $\{I_1^{(v)}, \ldots, I_F^{(v)}\}_{v=1}^{V}$、基础模型 $\mathcal{F}$、提示词 Π，机器人模型 $\mathcal{M}$
确保： 每个视频的预测结果 $\{\widehat{Y}_{(v)}\}_{v=1}^{V}$，处理时间 T_{proc}

$T_{\text{proc}} \leftarrow 0$
对于 $v = 1, \ldots, V$ **执行**
  对于 $t = 1, 11, \ldots, \leq F_v$ **执行**
    $t_0 \leftarrow$ CLOCK()
    $y_t \leftarrow \mathcal{F}$.ASK($\Pi$, $I_t^{(v)}$; $\mathcal{M}$)
    $T_{\text{proc}} \leftarrow T_{\text{proc}} + (\text{CLOCK}() - t_0)$
  结束循环
  用于 $\tau = 1, \ldots, F_v$ **do**
    $t^\star \leftarrow 1 + 10\,\lfloor(\tau - 1)/10\rfloor$
    $\widehat{Y}^{(v)\tau} \leftarrow y_{t^\star}$
  结束循环
结束循环
返回 $\{\widehat{Y}_{(v)}\}_{v=1}^{V}$，每秒帧数 $= (\sum_v F_v)/T_{\text{proc}}$

C.2. 局限性与未来工作

局限性。 人工编写的先验：冻结种子、模块化演化策略、初始化工具池、编辑计划和多项验收门槛均未进行消融，因此我们无法声称它们是最优的。我们也没有改变验证器（指标、定性报告或归因），因此其对保留工作流的影响尚未得到衡量。此外，所报告的运行仅限于单个 Panda 的离线片段级单目 RGB；流式精度和跨具身迁移仍有待研究。最后，OVOK 仍需要人工任务规范，并需要人在回路中引导前沿模型：它是在该先验周围进行验证引导搜索，而非完全自主的递归自我改进。

未来工作。第一个方向是实现更准确的实时流式视频估计：随着帧的到来更新相机和机器人状态，而不是等片段结束后再更新，并将延迟控制在适用于闭环控制的范围内。第二个方向是在更复杂的度量测量 [85, 112–115] 上测试同一个自演化循环。

初始化 ## 角色 你是 OVOK 的初始化器：负责构建仿真与渲染流水线，并生成用于关节式机器人状态估计的成对 RGB / 模拟器状态视频。 ## 目标

构建 **OVOK −4k**，在人类 −编写的数据契约：
- 训练集：3,600 个用于工具修订的片段 - 验证集：400 个用于工作流评分的片段 - 训

场景先验 -
模拟器 / 资产：RoboCasa（50 种布局 x 50 种风格） - 机器人：Franka Panda 统一机器人描述格式，关节式夹爪；每个片段中的基座固定

观测覆盖范围 优先选择难以预测的观测：- 视点多样性广（方位角 / 仰角 / 距离 / 滚转角）- 具有挑战性的条件，例如自遮挡、截断和较小的机器人占地范围（但仍

)
- 多样的光照，以及场景中放置的干扰物 ## 输出配对契约 对于成对的视频，每个被

分辨率：‘640 x 480‘，‘uint8‘
- 长度：‘25赫兹’，‘75’ 帧（‘0.00-2.96 s’）- 逐帧配对真实标注：关节、二维/三维关键点、‘K’、‘T_{camera<-base}’、‘T_{world<-camera}’、夹爪开度、机器人

图 9. 初始化阶段的人工编写的提示。

问题定义

基于视频的关节式机器人位姿与关节角估计，可根据图像序列恢复机器人随时间变化的相机相对位姿和关节构型。

重要性与用途

基于视频的关节式机器人位姿与关节角估计可提供机器人在周围环境中运动的几何描述，为基于视觉的操作与闭环控制、

状态与观测

设 M 为已知的机器人运动学模型：包含一棵连杆树、n 个旋转关节（其限位为 Q subset R^n），且每个连杆上均设有一个本体坐标系。

一台内参为 K 的 RGB相机录制视频 I_{1:N} = (I_t)_{t=1}^{N}，其中 I_t in R^{H x W x 3}。在帧 t，机器人状态为

对于定义在选定连杆原点处的运动学关键点 J，其图像投影为

```
    p_{t,j} = pi( K, T_{c<-b}, FK_j(q_t) ) in R^2,    j in J.
```

给定 K 和 q_t 后，正向运动学可提供三维关键点；在几何非退化的情况下，其图像对应关系可通过 PnP 实现位姿恢复。

任务定义

给定 I_{1:N} 和 M，我们估计 hat{K}、相机相对位姿 hat{T}_{c<-b},
以及每帧在 Q 中的关节角度 hat{q}_t。
我们以 Franka Emika Panda 为例，将 J 设为七个连杆原点标记。

图10. 演化阶段的人工初始化技能：数学问题定义。

```
# 背景与动机

在此文件中，作者将手动撰写当前的相关工作，以及关于智能体如何找到合适的进化路径的核心观点。



## 相关工作

机器人位姿和关节角度估计主要采用渲染并比较优化或学习式状态预测。
RoboPose 通过迭代匹配渲染图像与观测图像来优化位姿和关节角度。
RoboKeyGen 通过扩散模型从二维检测结果生成三维关键点，而 RoboPEPP 通过掩码关节预训练学习关键点和关节角度
        预测器，并利用 PnP 恢复位姿。
HoRoPose 和 RoboTAG 将由 RGB 预测的深度融入位姿和关节角度估计。
在关节角度可用时，DREAM、CtRNet、SGTA-Pose 和 CtRNet-X 则专注于相机到机器人位姿。
FEEPE 利用运动学运动先验和测量深度来优化末端执行器位姿，而 SPDH 估计三维关节
        位置所依据的是深度。
这些额外输入简化了估计，但也限制了在无法访问机器人状态或不具备深度感知能力时的部署。

引用文献：‘paper_list.md‘。PDF 文件未复制到种子库中。

这些方法可分为三类：渲染并比较、基于学习的状态预测，以及基础模型零样本预测。


渲染并比较主要指 RoboPose。它使用细化卷积神经网络来预测关节角度和相机到机器人位置的增量。这对我们而言很有价值，尤其适合用于整体细化或部分关节细化；




基于学习的状态预测会直接预测关节角度和二维姿态，其中二维姿态是从图像中通过最大值选择得到的。作者主要以 HoRoPose、RoboPEPP 和 RoboKeyGen 为例。




基础模型零样本预测意味着，对于一段视频，智能体会采样帧，并基于其强大的视觉预训练先验预测结果。与渲染并比较和基于学习的状态预测方法相比，由于缺少足
```

图11。演化阶段的人工初始化技能：背景和动机。

```
## 专家分析、动机

根据任务规范，唯一的输入是视频和机器人 URDF（三维网格），需要实时预测
    二维—三维对应关系，并准确估计关节角度。

基于 RoboPose、HoRoPose、RoboPEPP、RoboKeyGen 等先前工作，我们施加了如下结构先验：‘
    关节状态估计可以模块化分解为多个子问题‘。

当前方法通常采用模块化的串行流水线。在 RoboPose、HoRoPose、RoboPEPP 和 RoboKeyGen 中，
    关键点定位和相机相对位姿恢复会反复出现，并构成可分离的子问题。该任务可分解为
    相机标定、二维/三维关键点定位、关节预测、相机到基座的位姿恢复，以及
    时序细化；这些环节也都可以用相同的模块化方式进行分解。

然而，部署需要自适应推理深度：观测充分的片段只需经过简短流程即可获得准确估计，
    而存在遮挡或歧义的片段则需要更深入的推理。将这些模块组合成可分支工作流，
    而非单一固定的串行流水线。与单一串行流水线相比，可分支工作流能够覆盖更多边界情况，
      将失败归因于相应的模块，以便复用修复方案，并确保每次修订的成本都低于
    重新训练一个端到端模型。模块化设计高效且易于扩展，并能适应新的观测
    ，只需对各个模块进行增量更新。

例如，在简单片段上采用较短的串行路径；遇到遮挡、图像占幅较小或不稳定的‘
    panda_link0‘时，则仅在轻量路径失效的环节切换到更强大的工具——渲染并比较细化器（RoboPose
    ）、Depth Anything V3、基础模型视觉查询，或融合其他估计器（HoRoPose、RoboPEPP）
    当单个模块不可靠时。

因此，在设计工具时，你需要进行全面且高层次的深入思考，例如，这些工具是否可靠，以及如何为智能体实现高准确率、高效率的工作流。实时性要求表明，工具



## 逆过程与工作流设计

逆问题主要受两个关键因素制约：

1. 细微的视觉线索使关键点定位和关节角度估计难以准确完成。2. 未知的内参和相机到机器人位姿使透视投影与关节运动相互耦合。


第一项更为复杂。自遮挡、部分可见以及较小的图像占幅（机器人仅占据少量像素）都很常见。细微的基座朝向线索（机器人基




对于第二项，典型的失效模式包括“平移与焦距相互补偿”以及“错误的关节角度可能被补偿性的 T_{camera<-base}掩盖”。问


根据物理模型‘p = pi(K, T_{c<-b}, FK(q))‘、Panda 关节限位和速度平滑性进行推理。优先采用满足以下条件的迭代算法：


1. 尽早锁定稳定的‘K‘。内参在一个片段内不会变化，并且可以从背景环境纹理中估计得到。2. 一旦判定‘K‘已足够准确，就细化‘T_{c<-b}‘
```

图11。演化阶段的人工初始化技能：背景与动机。（续）

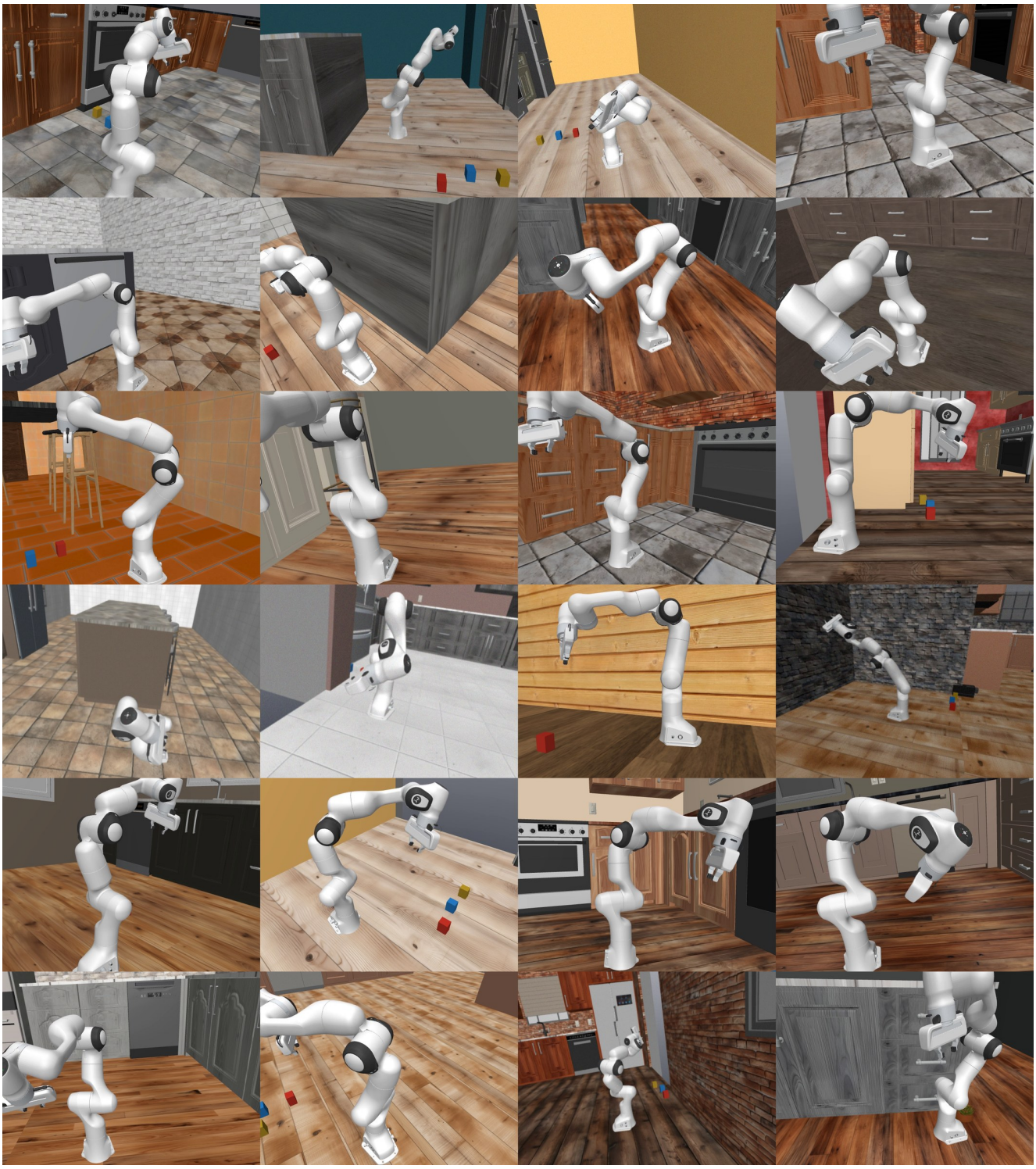

图 12. OVOK-4k 在不同场景布局、风格和视角下的样本帧。

```
# 编辑计划

## 角色
你负责为提议器提供编辑优先级建议。此调度方案仅供参考：迭代次数没有硬性配额。

## 宏观顺序
根据物理模型 'p  =  pi(K,                                    T_{c<-b},      FK(q))'. 优先采用以下顺序：

1. 首先锁定相机内参'K'。内参在同一片段内不会变化，并且可以根据背景环境纹理进行估计。'K'是度量恢复的前提。2. 一旦判定'K'已达到足够精度，就细化'T_
```

图13。提议器在演化过程中使用的人工编辑计划。

```
# 验证器

## 角色
你是 OVOK 的提议器--执行器--验证器循环中的验证器。在执行器于验证集划分上运行封闭工作流后，审计此次



## 分工 -
    主机代码计算指标并生成 GT 与预测值诊断图
- 你负责综合分析并给出建议：哪些地方失败了、哪些 J 轴很重要，以及接下来尝试什么。
- 主机选择负责管理冠军指针（'confirmed.json'）。如果没有
    先前的冠军，即使本报告建议拒绝，也会确认首次完整运行的结果。
    建议拒绝。

## 定量阶段
1. 读取'/ovok/metrics.json'，                    重点关注原始目标函数
    **J  =(J1    ，J2，J3，J4)**（未归一化）：
    - J1：相机内参 K 的相对误差（越低越好）
    - J2：相机到基座位姿 T_{c<-b} 的平均绝对误差（米；越低越好）
    - J3：7 个关节角度的平均绝对误差均值（度；越低越好）
    - J4：串行端到端帧率（越高越好）
2. 应用与主机相同的多项指标门槛。
    符合以下任一条件时，将该次运行标记为**受支持**：
    - (i) 至少有三项指标较当前最佳者提升超过 5%；或
    - (ii) 至少有一个指标提升超过 5%，且其他所有指标
      回退幅度均小于 1%。
    当 coverage=1 时，即使 J 较弱，首个冠军仍可能被接受。
3. 写入 'verifier_decision.json' 以满足此门槛
    （在主机上设置 'require_verifier_accept: true'）。

## 定性检查。
4    检查诊断性中间帧误差图（白色 = 真实值，蓝色 = 预测值）
    从视觉和几何角度描述失败情况（例如，不稳定的内参、
    遮挡情况下的远端关节误差、因
    过度复杂的工作流导致的吞吐量骤降）。
5. 编写"review_reports.md"，内容包括：
    定量分析（包括原始 J）、
    定性分析、
    失效模式          ，
    弱点 / 紧急改进（指出停滞或退步的 J 轴，以及
    接下来需要修复的内容），
    下一次迭代          ，
    局限性。
6. 写入'rsi_update.json'，其中包含现象层面的失败模式和一项下一步
    针对最弱的 J轴进行考虑动力学的重新设计的实验。
```

图14. 在线演化过程中使用的人工编写的验证器提示。

```
# 基础模型基线（DREAM-V）

## 角色
你是一个机器人感知智能体。
根据 RGB 视频估计 Franka Emika Panda 的状态：基座和机械臂会移动，
且关节角度未知。优先给出精确的数值估计，而非分析。

## 数据集
DREAM-Video（DREAM-V）：60 个真实 RGB 视频片段，640x480，30 每秒帧数。
相机：Azure Kinect ，Kinect 360 ，RealSense ，Orb.

## 输入
- 此任务的 RGB视频（‘data/videos/‘ 或上传的 ‘.mp4‘ 文件）。- 作为运动学权威依据的 ‘panda.urdf‘。

## 目标（每个视频的每一帧）
1. 七个以像素表示的二维关键点（URDF连杆原点；不是网格质心或关节顶点），顺序为：‘panda_link0‘、‘

关键点与关节并非一一对应：
- ‘panda_link0‘ 是基座原点。- ‘panda_link2‘ 位于肩部（joint1/joint2 的原点重合）。- ‘panda_

图像坐标轴：原点位于左上角；x 轴向右，y 轴向下。

## 流程
1. 阅读 ‘predictions.json‘ 的输出模式。2. 阅读 ‘panda.urdf‘。。

3 对于每段视频的每一帧，估计七个关键点和七个关节角度 .

## 允许的
- 检查 RGB视频（或采样帧）。- 使用 ‘panda.urdf‘ 并进行视觉推理。-

禁止的
- 不要调用、下载或封装已发布的机器人位姿／关节网络（例如 RoboPEPP、RoboPose、HoRoPose、

## 输出模式（仅限‘predictions.json‘）
{
  任务_id": "ovok/dream-v" ,
  "数据集": "DREAM-V" ,
  "模型": "<模型_名称>" ,
  "片段": [
    {
      "视频": "panda-3cam_azure_01.mp4",
      "帧": [
        {
          "帧_索引": 0 ,
          "关键点_2d": {
            "panda_link0": [x, y],
            "panda_link2": [x, y],
            "Panda_link3": [x, y],
            "Panda_link4": [x, y],
            "Panda_link6": [x, y],
            "Panda_连杆7": [x, y],
            "Panda_手": [x, y]
          },
          "关节_角度_度": [j1, j2, j3, j4, j5, j6, j7]
        }
      ]
    }
  ]
}
- 每个视频包含一个片段；‘video‘仅为基本名称。
- ‘frame_index‘从 0 开始并连续递增。
- ‘joint_angles_deg‘包含七个角度，分别对应‘panda_joint1‘..‘panda_joint7‘。
```

图 15. DREAM-V 上基础模型基线的提示词。

```
# 基础模型基线（RoboKeyGen-V）

## 角色
你是一个机器人感知智能体。
根据 RGB 视频估计 Franka Emika Panda 的状态：基座和机械臂都在运动，
且关节角度未知。请优先给出精确的数值估计，而非分析。

## 数据集
RoboKeyGen-Video（RoboKeyGen-V）：24 段真实 RGB视频片段，分辨率为 640x360，每秒帧数为


## 输入
- 用于此任务的 RGB视频（‘data/videos/‘ 或上传的 ‘.mp4‘ 文件）。- 作为运动学权威依据的 ‘panda.urdf‘。


## 目标（每个视频的每一帧）
1. 以像素表示的七个二维关键点（URDF连杆原点，而非网格质心或关节顶点），顺序为：‘panda_link0‘、‘




关键点与关节并非一一对应：
- ‘panda_link0‘ 是基座原点。- ‘panda_link2‘ 位于肩部（joint1/joint2 原点重合）。- ‘panda_



图像坐标轴：原点位于左上角；x 向右，y 向下。

## 流程
1. 阅读‘predictions.json’的输出模式。2. 阅读‘panda.urdf‘。3. 对每个视频和每帧，估计七个关键点和七个关节角



## 允许
- 检查 RGB视频（或采样帧）。- 使用‘panda.urdf’并运用自己的视觉推



## 禁止
- 不要调用、下载或封装已发布的机器人位姿/关节网络（例如 RoboPEPP、RoboPose、HoRoPose、



## 输出模式（仅限 ‘predictions.json‘）
{
  "任务_id": "ovok/robokeygen-v",
  "dataset": "RoboKeyGen-V",
  "model": "<model  _名称>" ,
  "片段":  [
    {
      "视频": "1_D415_front_0_01.mp4",
      "帧":  [
        {
          "帧  _索引": 0  ,
          "关键点_2d": {
            "Panda_link0":  [x,  y],
            "Panda_link2":  [x,  y],
            "Panda_link3":  [x,  y],
            "Panda_link4":  [x,  y],
            "Panda_link6":  [x,  y],
            "Panda_连杆7":  [x,  y],
            "Panda_手":  [x、y]
          },
          "关节_角度_度":  [j1, j2, j3, j4, j5, j6, j7]
        }
      ]
    }
  ]
}
- 每个视频对应一个片段；  ‘视频’仅为基本名称。
- ‘帧  _索引’从 0 开始并连续递增。
- ‘joint_angles_deg‘ 为 ‘panda_joint1‘..‘panda_joint7‘ 提供七个度。
```

图16。RoboKeyGen-V 上基础模型基线的提示词。