arXiv:2611.30173v1 [cs.RO] 23 Nov 2026

自信满满地迷失：光子匮乏下稠密视觉 SLAM 的静默失败与边缘部署

Shrikrishna Bhagirath Rajule、Srinath Neelamagam 和 Aniket Bera，美国印第安纳州西拉斐特，普渡大学，邮编 47907

摘要——**基于学习得到的双视图几何构建的稠密视觉 SLAM 通常在光照良好的条件下进行评估；在低光照下，它会静默地失效。两帧完全由 9 级量化噪声构成、毫无场景信息的图像，能让 MASt3R-SLAM 的网络得到 0.936 的跨视角几何自一致性；而一对真实暗帧的得分为 0.935。结构信息匮乏时，网络会为两个视角臆造出相同的几何关系，系统则信任这种一致性：在一段暗光序列中，系统处理了 1605 帧，输出 7 个关键帧，没有拒绝任何因子图边，并报告了全帧率；然而，它接受的对应关系中，超过 95% 违反了真实极线几何。一项虚高的统计量被两个子系统读取，而它们使用的阈值方向恰好相反，因此关键帧选取陷入饥饿，而重定位几乎接受了每个候选项。在 ETH3D 暗光序列的关键帧对基线上，同一统计量无需真实值或额外计算，就能区分自信崩溃的运行与正常运行（样本内）。低光照适配权重使能够跟踪的轨迹比例从 1/6 提升至 5/6，重新调节两个阈值也能达到同样效果；但两者都无法得到准确轨迹，因为这种适配只是降低了对错误几何的置信度，而没有找到正确几何。将该机制回放到一台由足式机器人搭载的 Jetson AGX Orin 上后，它的特征在四条路线中最暗的两条上重现；具体表现则随路线几何和传感器编码而异。网络在暗处并未失去置信度，而是在凭空制造置信度。**

I. 引言

夜间检查隧道的机器人需要一种 SLAM 系统，既能在自身携带的计算机上运行，也能适应现有光照条件。基于学习式双视图几何构建的稠密视觉 SLAM 是最有力的候选方案：MASt3R-SLAM [1] 通过基础模型回归稠密点图和描述符，并在无需标定的情况下实时重建。然而，这类系统的所有定量结果都在光照良好的条件下测得：要么仅在边缘设备上评估吞吐量 [2],，要么仅在合成暗化条件下评估成对网络 [3]；它能否在暗处部署于边缘计算机上，尚未经过验证。本文探讨实际部署时会发生什么：系统在光子匮乏下会以何种具体形式失效、失效机制是什么、何种健康信号能以零成本揭示失效、现有补救方案的收益与代价是什么，以及其中哪些结论适用于足式机器人搭载的 Jetson AGX Orin。

显而易见的补救方案已经存在：Dark3R [4] 通过离线 LoRA 蒸馏 [5], , 将同一个 MASt3R 骨干网络适配到光子匮乏的原始图像；离线阶段可以获取每一幅图像，并对几何进行全局优化。SLAM 具有因果性，必须即时确定关键帧、匹配结果和位姿；蒸馏得到的低光先验能否经受这一转变，尚未经过验证。

我们发现，它作为一种行为得以保留，却没有保留为几何；而更值得关注的结果，是原本要修复的问题会对系统造成什么影响。加载 Dark3R 的将权重载入 MASt3R-SLAM，其他设置一概不变，ETH3D-SLAM 暗光序列 [6]中可用轨迹的数量便从 1/6 提升到 5/6。可用指的是满足工作量标准（≥ 50 关键帧数量足够，且对齐能够收敛），而非精度标准：所有暗光轨迹的误差都达到米级，且任一模型接受的对应关系中，超过 95% 都违反真实极线几何（第 IV-B 节）。调整原版系统的两个门控阈值，同样可将结果提升到 5/6（第 IV-A 节），但错误对应关系依旧存在；在随附的接受半径下，无论如何调整门控都无法摆脱退化状态，而适配则可以。在 sofa_暗场景_1 中，原版 MASt3R-SLAM 处理了 1 605 帧，输出 7 个关键帧，并报告 14.6 帧率；所有部署时可能监测的信号均显示正常：跳帧数为 0，Cholesky 分解失败次数为 0，重定位尝试次数为 0，没有因子图边被拒绝，且轨迹文件通过了标准评估。它对系统本身、评估协议（仅凭关键帧文件计算的绝对轨迹误差无法区分只有七个关键帧的地图与完整地图），以及任何下游使用者都毫无提示地，将一段 1605 帧的行走序列压缩成了一张只有七个关键帧的地图（图 1）。

这一机制是匹配器接受半径下网络输出的一种属性，并非任何门控阈值所致。MASt3R-SLAM 通过几何自洽性验证对应关系：经过网络跨视图点图传递的像素，必须落在另一视图自身点图的 10 厘米范围内。在近乎黑暗环境中，两项预测会变得虚假地一致；网络输出的几何信息彼此吻合，却对应着一个它无法看见的场景，因此有效性达到 0.93。当置信匹配操作数低于 0.333 时，关键帧规则便会触发；在静默运行中，该操作数的中位数为 0.85–0.89，因此在 1605 帧中仅触发六次。重定位在同一数值

升高于0.3 时接受候选项：在 plant_暗场景中，543 次尝试有 497 次被接受，并拒绝了其提出的边中的 21.5%。同一原因导致两种相反症状，我们称之为**静默退出**（标称帧率下关键帧生成不足）和**喧闹 退出**（反复抖动：无限制地重定位到错误锚点）。

诊断得出了一种无需任何计算开销的检测器。在每条因子图边上计算关键帧对基线处的几何匹配有效性，g；在 8 位 ETH3D 上的 24 次运行网格中，它能够以毫无重叠的方式区分置信度崩溃与跟踪运行（第 IV-B 节）。虚假的高 匹配率就是警报。这是系统状态的属性，而非输入的属性：对于相同的帧，原版读数为 0.929，适配模型读数为 0.371。它并不能证明剩余地图的质量，而其健康水平取决于相机。

我们的贡献包括：（1）**一种机制**，已在完整序列上验证，

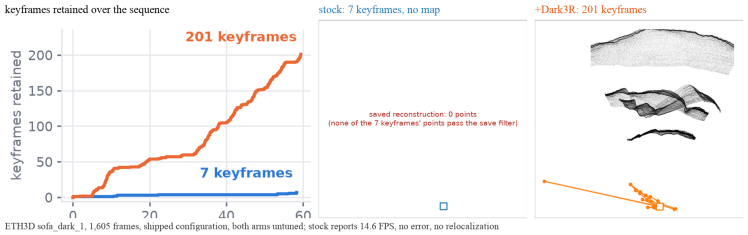

图 1. 静默失败。ETH3D 沙发_暗场景_1，1605 帧，出厂配置，双臂均未调优。左：整个序列中保留的关键帧，原版为 7 帧，适配后的机械臂为 201 帧。中、右：每次运行保存的重建结果，俯视图，并标出其关键帧位姿（方形：起点）。原版报告的帧率为 14.6，无误差，也未进行重定位，最终保存了一个空重建。

并通过逐门控干预证实，它解释了稠密学习式SLAM为何会在低光照下悄然失效：网络在暗场景中并未失去置信度，而是凭空制造置信度；而同一个数值既会让关键帧选取因信息不足而停摆，又会让重定位接受所有结果；（2）**一种无需真实值、也无需额外计算的检测器**，并给出了其工作点以及在 66 次标注运行中的错误率；（3）**对补救方法进行对照比较**，包括重新调参、增强前端、权重自适应，以及一种无需训练、介于原版与适配权重之间的剂量方案。在出厂默认接受半径下，只有权重自适应能脱离退化状态；部分剂量则消除了其测得的唯一一种良光照退化，同时其绝对轨迹误差代价随剂量增加而上升；（4）**在机器人上进行范围测试**：在 Jetson AGX Orin 上（内存定律、Orin NX 上限、功耗平稳），并使用四足机器人的 8 位摄像头完成四条路线测试后发现，该机制的特征在两条最暗路线中重现；在三条无转弯路线中的两条上，权重自适应覆盖了原版所无法覆盖的一个级别；而故障呈现哪种形式，则因路线几何和传感器编码而异。

II. 相关工作

学习式稠密几何及其监测。DUSt3R[7] 将双目立体重构为点图回归，MASt3R [8] 加入了描述子头，MASt3R-SLAM [1] 则将模型嵌入因果流水线；后续的前馈方法 [9], [10]及基于它们的 SLAM 系统 [11] 都在光照充足的条件下进行评估。受损条件下的学习式跟踪器以持续漂移换取避免灾难性损失 [12]；一份针对六种跟踪器的低照度报告 [13] 未包含任何点图系统。我们以 MASt3R-SLAM 为基础，因为其适配采用 LoRA，作用于 VGGT 系列架构所缺少的交叉注意力投影层。运行时监测通过残差学习误差预测器 [14]或风险分数 [15]，，而*SLAM* 置信度陷阱 [16]指出，该领域发布的精确系统并未提供一致的不确定性度量；我们使用的是系统已有计算结果构成的、未经学习的统计量，用于识别一种无人描述的失败：系统报告一切正常，却什么也没有。LeanGate [17] 会跳过预测器判定为冗余的帧；第 IV-B 节描述了一种表面上的冗余恰恰意味着失败的情形。

低光照、增强、其他传感器、嵌入式计算。Dark3R [4] 在真实的光子匮乏原始图像上蒸馏 MASt3R，用于离线运动恢复结构；DarkVGGT [18]通过 LoRA 改造 VGGT，但推理时需要热成像相机，并将我们在 SLAM 内测量的日光代价命名出来。Twilight SLAM [19] 发现，增强对基于特征的流水线而言是效果参半的补救手段；DarkIR [20] 会降低稠密 SLAM（第 IV-D 节）。近红外照明 [21]、[22]和热融合 [18]都需要额外传感器；我们想知道，在需要这些传感器之前，被动相机的性能还能提升到什么程度。EC3R-SLAM [2] 测得 MASt3R-SLAM 在 Orin NX 上的运行速度为 0.1 帧率；我们并不与其比拼吞吐量。最接近的评估 [3] 报告称，MASt3R 在 Jetson 硬件上、合成弱光条件下的跟踪成功率为 93%，但其运行的成对网络没有包含在此处发生故障的关键帧和重定位子系统，而且采用的 2 米 30° 判据无法识别四关键帧地图。

III. 方法

A. 置信度如何融入 *MASt3R-SLAM*

对于一对关键帧（i, j），骨干网络通过一次前向计算生成四个点图： X^{ii}、 X^{jj} 以及跨视角的 X^{ji}、X^{ij}，并为每个像素生成置信度；系统构造对称匹配置信度 $Q_{i\to j} = \sqrt{Q^{ii}[\text{idx}_{i\to j}]\, Q^{ji}}$ ，并以 τ_Q (Q_conf，随 1.5 版本发布)作为门控。**几何匹配有效性** g 是对应关系自洽的像素所占比例：视图 j 中的每个像素都通过 X^{ji}传递，并由高斯–牛顿法投影到视图 i的射线上；若投影在图像范围内收敛，且落在该像素处 X^{ii} 的 dist_thresh $= 0.1$ 范围内，则接受该像素。描述符只会在此后细化像素；描述符和 Q 都不会改变有效性，因此 g 是网络的两项几何预测彼此一致程度的统计量，位于所有门控之前。**匹配比例** m是门控后的统计量，即 g的掩码及 $Q > \tau_Q$：有效且具有置信度。跟踪器的关键帧测试会读取 m的逐帧形式，并结合达到的不同关键帧像素比例 $\min(m_k,\ u_f)$；当该值低于 match_frac_thresh =0。333 时便创建关键帧；当 m 升至 reloc.min_match_frac $= 0.300$以上时，重定位会接受一个候选项。**这些规则针对近乎相同的阈值，指向相反的方向**；在正常光照下这并无大碍，却是我们观察到的两种失效模式的根源。

B. 调整骨干网络

Dark3R 提供了在光子匮乏的原始图像上训练的 LoRA 适配器。我们将其离线合并到基础权重中，生成一个可直接替换的检查点，其张量键集合与原版相同（1017 个张量），且无需更改 SLAM 代码：适配后的 机械臂， **+Dark3R**在表格中使用，并合并到运行名称中。恰有 288 个张量不同，其中编码器、解码器和第二个解码器块各有 96 个；预测头则逐位完全相同（Dark3R 的图示描述了对预测头进行微调；发布的检查点

未对其进行改动）。适配带来的变化并非仅限于匹配：在每个序列的 24 对相同输入帧对上，暗场景中双臂之间的中位相对点图变化为 168%，而明亮场景中为 11.5%；匹配比例变化了 32%（分母是原版自身的暗场景点图：结果在于变化方向，而非倍数）。这些帧对是某一帧与几帧之前的自身关键帧进行比较；其短基线有效性 g_{short} 在暗场景下为原版的 0.930、适配后的 0.631，在良好光照下双臂均为 0.89–0.90。对任何模型而言，短基线帧对都是自洽的，因此 g_{short} 并非检测器（第 IV-B 节）；发生变化的是暗场景下的数值。适配的作用是降低置信度，而非纠正错误：它消除了暗场景所制造的跨视图一致性，但其仍然接受的对应关系，与原版相比真实值仍同样错误。置信度更低，而非更正确（第 IV-C 节指出了这种影响源于权重）。

C. 调节适配剂量

由于合并时折叠低秩更新而未进行重新缩放，$\Delta = W_{\text{merged}} - W_{\text{stock}}$ 是发布的适配版本，而 $W(s) = W_{\text{stock}} + s\,\Delta$, $s \in [0,1]$ 则是在原版与完整合并之间的一条直线。我们在 SLAM 过程开始前构建一次 $W(s)$；所用操作只是简单的权重空间插值，而新颖之处在于剂量在因果 SLAM 循环中的作用。

D. 诊断失效及一个潜在缺陷

我们在因子图边的构建位置添加了只读诊断功能，由环境变量启用，逐次记录每条边方向上 Q, g, m 的分布、提供和拒绝的边数，以及明确的调用方标记。我们还发现，local_opt.window_size 会被读取却从未使用：边不断累积而不会被移除；在低光照条件下，每次重定位尝试都会增加边，导致 15.5 GiB GPU 在关键帧 273 处耗尽。我们实现了该参数所指示的窗口机制（只要边的任一端点位于窗口内，该边就会保留）。另一个独立分支也发现了相同缺陷；我们声称测量了该缺陷造成的代价。

E. 评估协议

在评估配置下，流水线在固定硬件配置内具有位级确定性（重复运行得到的轨迹逐字节一致），因此每种配置仅报告一次运行结果；但这种确定性不适用于运行时间，也不适用于不同硬件之间（Jetson 的功耗模式变更导致 6 次运行中有 3 次的关键帧数量发生变化，其中一次变化了 43%）。文中采用三种轨迹协议，且绝不在同一张表中混用：经 Sim(3) 对齐后，将 关键帧位姿与 ETH3D 真实值进行比较，计算 ATE；将逐帧轨迹与原始场景的 COLMAP 参考结果进行比较；以及将关键帧位姿与机器人里程计进行比较，后者是参考值而非真实值（第五节）。所有帧率数据均附有关键帧数量。τ_Q 在两个序列（einstein_暗场景、plant_暗场景）上进行了调优，另外四个序列则留作测试；第四节 A 节中的门控阈值重新调优网格测试则覆盖了全部六个序列。

表 I 出厂配置，双臂未经调优，RTX 5080。Sim(3) 对齐后关键帧位姿的绝对轨迹误差（米）；“–” = 少于 50 个关键帧；“失败” = 对齐未收敛。

序列	帧数	原版 关键帧数	原版绝对轨迹误差	+Dark3R 关键帧数	+Dark3R 绝对轨迹误差
爱因斯坦 _dark	2035	320	1.522	358	1.639
植物_暗场景	1140	506	失败	304	**0.871**
绑架_暗场景	526	15	–	125	**1.217**
sofa_dark_1	1605	**7**	–	**201**	**1.320**
沙发 暗场景 2_ _	281	4	–	1	–
沙发 暗场景 3_ _	1187	**4**	–	**160**	**1.298**
可用		**1/6**		**5/6**	

IV. 实验

设置。 ETH3D-SLAM 暗光序列（Vicon 真实值，281–2035 帧）； $2\times$ RTX 5080；窗口大小为 120，单线程，内参已标定；机械臂**原版**， **+DarkIR**（先增强，再使用原版），**+Dark3R**。若轨迹包含 ≥ 50 个关键帧且 Sim(3) 对齐收敛，则视为可用；这是工作负载标准，而非精度标准。

A. 出厂配置下的主要结果

表 I 展示了从业者加载检查点后得到的结果：1/6 到 5/6，无需调参。数量并非本文的论点。原版系统在出厂门控下可用的数量为 1/6，在其自身最佳设置下为 2/6 τ_Q，而将两个匹配比例阈值都提高到 0.80 后则为 5/6：这是同一系统的三种调参结果。在 0.80 时，原版权重有 4 个序列的绝对轨迹误差低于适配模型，且整体达到 5/6。在采用匹配协议的 42 次此类重新调参的原版运行中，g 范围为 0.369–0.933，其中 41 次高于适配模型的上限 0.451（唯一例外是 kidnap_暗场景，数值为 0.369；这是原版系统在该场景下已有的数值）；在 0.80 时，重新调参后的静默序列仍保留了膨胀的统计量（门控后 m 0在 sofa_ 暗场景 _1/3 上分别为 0.828 和 0.848 ）。重新调参只是移动阈值以匹配膨胀的统计量，并不会消除膨胀；它既不触及匹配器，也不触及门控的运算对象，因此接受的对应关系与原版系统完全相同，其中 97–100% 违反极几何（见 IV-B 节）；匹配器的接受半径是另一个调节手段，该节对此另有讨论。阈值也无法通用于所有序列：在 0.80 时，每 1000 帧的关键帧数介于 52–352 之间，且有一个序列导致存储溢出。sofa_暗场景_2 （281 帧）在出厂阈值下得到 4 个和 1 个位姿，而在 0.80 的重新调参原版系统下得到 ≥ 50：这是配置的局限，并非序列长度所致。einstein_暗场景的绝对轨迹误差呈双稳态（求解器容差变化时为 1.503 或 1.639），因此该行不作排序比较。

B. 机制与检测器

三条序列与两种相反症状收敛于 $m \approx 0.7$ ，相对于接近 0.3 的阈值（图 2）：门控后中位数 m 0.728 与 0.691，出现在静默的沙发_暗场景_1以及沙发_暗场景_3，位于 g 0.929 与 0.927；植物_暗场景的重定位边为 0.699，位于 g 0.831（其全部调用上为 0.826），而正常工作的爱因斯坦_暗场景上为 0.072，位于 g 0.562。适配模型以同一路线关闭两个出口：其门控后 m 在沙发_暗场景_1的关键帧对边上为 0.219，低于阈值，且其

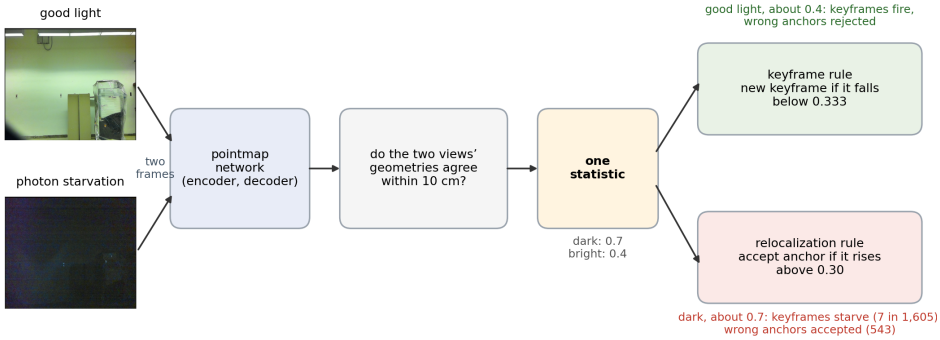

图 2。机制。两帧输入点图网络；当两个视图的几何预测之间的差异在 10 cm 以内时，即接受一组对应关系；置信比例是两条规则读取的唯一统计量。当置信比例低于 0.333 时，关键帧规则触发；高于 0.30 时，重定位规则接受一个锚点。在光线良好的情况下，该统计量接近 0.4，两条规则都能正常工作。结构信息匮乏时，网络会凭空构造出与自身预测一致的几何结构，使统计量升至约 0.7，导致关键帧选取不足（1605 帧中仅有 7 个关键帧），且重定位接受错误锚点（plant_dark 序列中 543 个锚点有 497 个错误）。帧来自机器人的夹爪相机，灯光开启，曝光时间为 1/100 s。

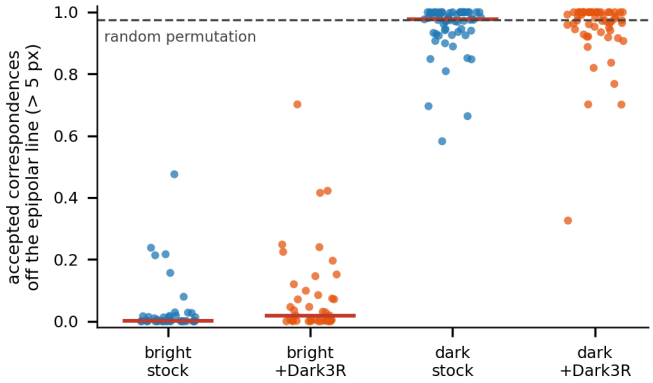

图3。得到的对应关系是错误的。对于基线为 40 帧的每一对帧，统计每个分支所接受的对应关系（有效且通过置信度门控）中，与真实极线的距离超过 5 px 的比例；红色条形表示中位数；虚线表示对相同像素进行随机排列所得的结果。明亮帧对来自两个 ETH3D 序列（每个分支 42 对），暗场景帧对来自五个暗光序列中的三个（每个分支 61 对）。

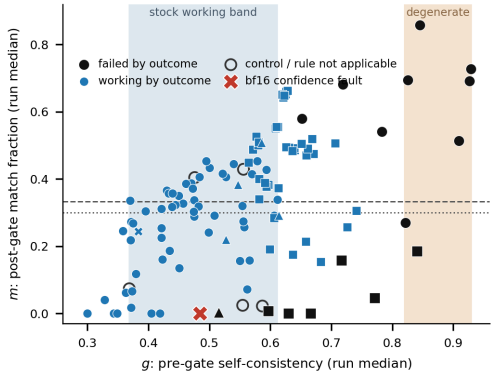

图4。(g, m) 平面，每个点代表 121 次有读数的运行中的一次。圆形表示 ETH3D 8 位，方形表示 Dark3R 原始数据，三角形表示重新编码为 8 位的厨房场景；色调分别表示原版工作区间（0.369–0.611）和退化区间（0.821–0.929）；虚线和点线分别表示 m 上的关键帧规则（0.333）和重定位规则（0.30）。标签不使用 g 或 m：ETH3D 暗光运行在 ≥ 50关键帧时处于工作状态且对齐收敛，在 ≥ 100 重定位尝试时反复抖动，其他情况则静默；明亮场景运行在发生 ≥ 10 Cholesky 分解失败时处于退化状态；原始数据运行只要曾经重定位，就标记为丢失。空心标记：被劫持的_暗场景对照，以及两次剂量不足的沙发_暗场景_2 运行。

重定位边界 m 在植物_暗场景上为 0.0065，因此 97.5% 的候选项被拒绝。每次仅干预一个阈值即可确认这种分离：只有关键帧阈值会改变饥饿症状（关键帧从 7 到 49），只有重定位阈值会改变反复抖动。

一致性是错误的吗？极线测试（图3）。ETH3D 提供标定和位姿，因此可以根据真实极几何，对每个分支接受的对应关系进行评分。在来自五个暗场景序列和两个明亮序列的相同图像对上（两个基线共 172 个暗场景图像对；我们报告其中帧数为 40 的 85 对，因为 10 帧基线足够短，任何模型都能保持自洽），原版接受的对应关系中，暗场景下有 97–100% 与极线相距超过 5 px，而明亮场景下为 0.0–1.5%；适配分支则分别为 96–99% 和 0.2–7.9%；对这些相同像素进行随机排列，得分为 97–98%。原版在暗场景下接受了全部像素的 58–91%，适配分支则为 3–65%。因此，原版人为制造的一致性是错误的一致性；适配分支接受的对应关系同样是错误的，所以解除关键帧选取饥饿状态的，是对错误对应关系降低置信度，而非恢复正确对应关系；并且 g并不能证明几何正确：合并/沙发_暗场景_1 的读数为 0.371，但其接受的对应关系中有 97% 是错误的，且其在一个 $\sim 3\times 4$米房间中的绝对轨迹误差为 1.32 m。

门自身的操作数，每一帧。记录跟踪器的关键帧测试在 36 次回放中的表现（每次均复现了原始关键帧数量），直接揭示了饥饿状态：在全部六次静默原始运行中（沙发_暗场景_1/2/3，增益为 τ_Q 1.5 和 1.0），$\min(m_k,\ u_f)$ 在 99.3–100% 的帧上不低于 0.333，中位数为 0.85–0.89；在 277–1 604 个跟踪帧中，该规则仅触发 0–6 次。同样的回放给出了否定结果：静默运行中的逐帧有效性 g_{short} 读数为 0.92–0.93，而运行正常的 kidnap_暗场景对照为 0.88–0.91；相隔几帧的两个视图对任何模型而言都是自洽的，因此检测器是关键帧对统计量，其证据只有在因子图调用时才会出现，而每次静默运行只有 2–6 次调用。故障抑制了自身的监控证据；首次读数要等到第二个关键帧出现，而关键帧由运动决定（一次明亮场景原版运行也等了 279 帧）：三次出厂配置静默运行分别在第 286、150 和 231 帧（分别是序列开始后 10.5、5.6 和 8.5 秒；序列时长为 59、44 和 10 秒）才出现读数；而两次 $\tau_Q = 1.0$单关键帧原始运行中始终没有读数。一旦出现读数，便已具有诊断意义：在沙发_暗场景_1中，原版的六次读数全都接近 0.93。

检测器。在 24 次运行网格中（原版在 $\tau_Q \in \{1.0, 1.5\}$，适配版在 $\{1.5, 3.0\}$，20 次读数；四次单关键帧运行

输出为零，而整个运动序列中因子图调用为零本身就是警报，并会在计算任何统计量之前进行检查），g 呈现出互不重叠的区间：原版退化为 0.821–0.929（$n = 5$），原版正常工作为 0.369–0.611（$n = 5$），适配版为 0.328–0.451（$n = 10$）。绑架_暗场景（0.369）未能触发该警报，这是正确的：脚本化绑架并非退化。**标签绝不取自** g，且关键帧数量与 g 在两个方向上都相互独立：植物_暗场景在 506 个关键帧时剧烈波动，且为 g 0.826；机器人在 L3 的原版运行保持 58 个关键帧，数值为 0.86–0.88；明亮序列每 1000 帧保留 14–26 个关键帧，数值为 g 0.36–0.53。**错误率与工作点**。在按结果标记的 66 次 ETH3D 运行中（11 次失败、55 次正常工作；排除三次绑架对照），ROC曲线下面积为 0.95；阈值设为 0.65 时，11 次失败中有 9 次超过该值，55 次正常运行则无一超过。这两个数值均为样本内结果（66 次运行包括 20 次网格运行，且阈值也是在这些运行上选定的）；两次漏检是沙发_暗场景中的部分剂量运行_2 （281 帧中分别有 4 个和 41 个位姿），依据 ≥ 50-关键帧规则，在其能力边界处被标记为失败。**部署者应该采用什么**。目前还不存在可迁移的阈值。绝对阈值 0.65 会在健康的机器人行走中产生误报（0.72，见第五节）；相对于明亮通行情况的比值也有重叠（正常运行达到 1.63×，已观测到的失败从 1.65× 起）。可迁移的是排序关系：在同一台相机和同一接受半径内，若 g 高于该相机健康运行的水平，就表明存在人为制造的一致性（这一现象出现在机器人最暗的两条路线中；而在较亮的两条路线中并未出现，因为无论失败是明显还是不明显，都无法形成排序）；明显的退出可通过重定位次数和门控后m读出；对于机器人，可将地图与里程计进行核对，后者无需标定。**它呈现出等级变化**：沿着三条暗光序列上的 $W(s)$， g 在两端都有读数的八个步骤中全部下降（沙发_暗场景_1 0.929、0.783、0.578、0.371，分别对应 $s = 0$、0.5、0.75、1；随着关键帧数量增至 7、18、102、201），并且在六条暗光序列中的五条上下降；唯一例外是绑架_暗场景对照。**它无需真实值**：绑架_2 未提供任何数据，而 g 识别出退化的机械臂（原版，0.719，78 次 Cholesky 分解失败；另一条机械臂为 0.370）。唯一的亮光退化发生在合并/绑架_1，数值为 0.845；但在绑架_2中情况反转：绑架会使任一机械臂都不稳定；三条摇晃序列则未显示运动影响（双臂为 0.554–0.606）。

在 14 位原始数据输入下，故障性质有所不同，而 g 无法察觉这一点。在两个 Dark3R 原始场景中，以九档包围曝光和五种剂量进行测试（共 45 次运行），原始模型在暗场景档位下跟踪失败，且症状十分明显：每次运行会尝试重定位 306–358 次。对于这些运行，门控后 m 骤降至 0.000–0.159；而对于每次成功跟踪的运行，该数值为 0.154–0.663。同时，g的重叠范围为（丢失时 0.60–0.77，完整时 0.56–0.74）。**编码方式决定了故障类型**：厨房曝光包围序列经重新编码为相机输出的 8 位 sRGB 流后，原始模型在 1/1000 下的故障从明显失败变为静默失败（360 帧、单个位姿、零次重定位、零条因子图边、11.6 帧率）；而 sRGB 的阴影提亮使其能在 1/400 下跟踪，此时线性原始数据会导致跟踪失败。适配模型则呈现相反情况：它能在 1/1000 的原始数据下跟踪，却在相同帧的 8 位编码版本上陷入反复重定位（303 次）。六次在一个场景上的运行结果与第 IV-C 节的噪声配对结果一致。**两个盲点**：在合成曝光阶梯（包含三个明亮序列，并通过泊松噪声和读出噪声进行光子缩放）上，适配使跟踪范围扩展了约 7 分贝，但g的置信度区间变得模糊，达到 0.55–0.75；此外，它无法察觉一种会破坏置信度排序、却不影响几何结构的故障：在一段有 94% 的边被拒绝的运行中，bfloat16推理仍给出 0.484 的健康读数，而门控后 m （0.0001，对比 0.245）则能捕捉到这一故障。图4将 121 次运行置于同一平面上：在 8 位 ETH3D 系列中，除两次沙发暗场景剂量运行外，所有静默失败都位于所有正常运行的右侧，而 m 不存在全局阈值。

C. 消融实验：接受半径、无场景输入、部分合并

收紧接受半径只会掩盖故障，并不能修复它 （图 5）。将 dist_thresh 减半至 5 cm 后，静默退出依然存在（ sofa_ 暗场景_1 7关键帧为 g 0.86，sofa_暗场景_3 6 为 0.85）。在 2 cm 下，计数看似恢复正常：71 和 53 个关键帧， g 0.46 和 0.50，以及 4 次和 0 次重定位。但地图仍然错误，在约 3×4 m 的房间中，绝对轨迹误差为 1.24 和 1.28 m。半径决定哪些对应关系能够通过自洽性检验，而不是决定它们是否正确，因此，一次本可通过所有基于计数的检查的运行，仍然具有同样错误的几何结构。在 2 cm 下，原版模型在完成的五个序列中的四个序列上达到 可用状态 ，但仍是使用同一组错误对应关系的第三次重新调参；而适配模型在 5 cm 和 2 cm 下都在 6 个序列中的 4 个序列上丢失跟踪或溢出关键帧存储：随附的 10 cm 是原版模型的人为制造的一致性能够通过、而适配模型更稀疏的一致性仍能通过的半径，收紧该半径反而会移除补救措施。由于 g 是通过该半径定义的，其区间只能在单一半径下成立。

这种一致性是从噪声本身中制造出来的。两帧彼此独立、完全不含场景的 9 级量化噪声，使原版模型的跨视图自一致性达到 0.936；真实暗光图像对的数值则为 0.935。适配模型在相同噪声上的读数为 0.104。对于噪声对，置信度门控会拒绝所有内容；真实帧则提供了置信度。适配模型也有自己的盲输入（均匀白噪声，0.976）：每个网络都会在某种无结构输入上与自身达成一致，而形似 ETH3D 暗光的 9级输入，则是原版模型的盲输入。

合并的任何一半都不等同于适配。仅编码器块就能让沙发_暗场景下的关键帧选取摆脱饥饿状态（关键帧从 7 增至 68，重定位 44 次），同时 g 仍为 0.79，高于原版工作区间（0.369–0.611）；仅解码器块则完全无法跟踪该序列。只有完整合并才能同时改变关键帧选取和 g。

D. 负面结果：增强

离线应用 DarkIR-m（ 31×亮度提升），然后运行原版 MASt3R-SLAM，并对 τ_Q 的四个取值进行扫描，得到**0/6** 条可用轨迹；相比原版在其最佳设置下的 2/6 更差，甚至不如不做预处理，且任何情况下最多只有 13 个关键帧。增强帧的读数为 $g = 0.887$– 0.948，这是我们记录到的最高值，每次读数都是如此，但

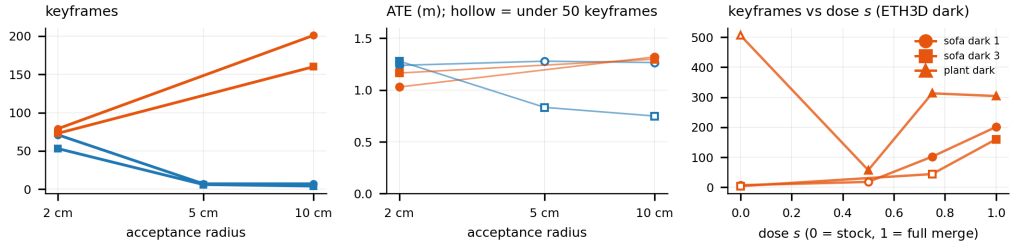

图 5. 消融实验。左图和中图：sofa_暗光_1 （圆形）与 sofa_暗光_3 （方形）上的接受半径，原版（蓝色）与适配模型（橙色）：在 2 cm 处，原版的关键帧数量恢复正常，而其绝对轨迹误差没有变化；空心标记表示关键帧数量少于 50 的运行，此时绝对轨迹误差不构成测量值。右图：三个 ETH3D 暗光序列中，关键帧数量与剂量 s 的关系；空心标记表示发生剂量不足或反复抖动的运行。

kidnap_暗场景对照（0.600）：提亮仅包含约九个强度级别的帧，会放大量化造成的伪梯度，并制造出退化信号。

E. 嵌入式部署

在 Jetson AGX Orin 上，对于关键帧数量为 4–394 的多次运行，实时 CUDA 分配仅跨越 6.91–9.12 GB，而分配器的预留空间跨越 7.3–54.3 GB：该预留空间是一个伪象；预分配的关键帧存储给出了一个规律，实时需求 ≈ 3.57 GB $+6.54$ MB $\times$ 关键帧_缓冲区 $+$最多 $\sim$2.2 GB 的地图增长，并且在工作地图上，模拟的 Orin NX 上限约为 128 个关键帧。两条机械臂上的功耗均稳定在 36.5–39.3 W；在六对实验中，适配模型的每关键帧能耗低 6.2–17.2$\times$ ，因为原版在相同功耗下会生成 4–10 个关键帧，而其每帧能耗为适配模型的 1.8–2.1$\times$ 倍；这是关键帧数量造成的效应，在机器人的梯子上会消失，此时双臂均生成关键帧（第五节）。

F. 负面结果，以及剂量

实测负面结果：bfloat16推理在两个数据集上都会丢失暗光跟踪（chapel 仅跟踪了 500 帧中的 1 帧，进行了 246 次重定位尝试；ETH3D 进行了 1319 次尝试），因为 8 位尾数破坏了门控所依赖的置信度排序；fp16 则能保持该排序（500 帧全部跟踪成功）。

日光代价与剂量。在原始亮光场景的 500 帧逐帧轨迹上，适配检查点的绝对轨迹误差为 0.421，而原版为 0.178（+137 %；单个场景、单次曝光）；Dark3R 自己的干净图像表报告离线+18% [4]，而 DarkVGGT 将同一代价命名为 [18]；该数值随剂量 s单调上升。在四种曝光下遍历全部五个 Dark3R 原始测试场景（相对于 COLMAP 的逐帧绝对轨迹误差）后，出现了三种情况：光照良好时，完整合并从未是最佳方案；中度暗光下，部分剂量在 8 对 8 的比较中都以绝对轨迹误差胜过完整合并；极暗环境下，完整合并在 3 对 3 的比较中均为最佳，而部分剂量会导致跟踪丢失。在我们自己的协议中，两端情况都成立：ETH3D 暗光属于极暗环境（沙发_暗_1 7，18、102、201 个关键帧，曝光为 $s = 0$，0.5，0.75，1），而部分剂量完全消除了合并/劫持_1的亮光退化（g 0.439，$s = 0.5$ 相对于 0.845；Cholesky 分解失败次数从 46 降至 0）。每个 stock 与线性原始图像的比较都存在编码混杂因素（原版是在 sRGB 上训练的，仅阴影提亮就足以让它跟踪一个在原始线性数据上会丢失的厨房曝光包围序列）。可以设想一条规则，使 s 提高，直到 g 进入健康区间在 ETH3D 上；我们没有运行该规则，而在原始场景上它也无法奏效，因为 g 无法在那里区分丢失运行与完整运行。**健康水平是相机和流水线的属性**（ETH3D 工作状态下的原版为 0.369–0.611，原始 chapel 场景的完整运行范围为 0.60–0.74，原始 kitchen 场景为 0.56–0.68）：部署时会在同一相机的明亮通行过程中对其进行校准，按调用方读取g ，在计算任何统计量之前检查跟踪情况（因子图调用次数为零、重定位次数、门控后 m），并在机器人上根据里程计检查运动：覆盖率会标记每个丢失的无转弯行走过程，而起点到终点的范围则标记压缩地图（第五节）。

针对警报采取行动。一个包装器在每次因子图调用时读取 g，当该值首次超过 0.65 时，就地将在线模型切换为适配权重，且只切换一次。在 sofa_暗场景_1中，警报在首次读取时触发（帧 286， g 0.927），运行结果从 7 个关键帧增至 187 个（35 次重定位尝试，切换后为 g 0.399，187 个位姿上的绝对轨迹误差为 1.27 米；适配分支：201 个关键帧，1.32 米）；半剂量则产生 13 个关键帧。在机器人的 L3 行走中，警报在帧 179 触发（g 0.946），切换后的运行复现了适配分支自身的失败（47 个关键帧，而适配分支为 49 个，459 次尝试， m 0.000， g 0.453）：该操作在适配权重有效的地方也有效，并继承了其适用范围；切换本身没有造成损失（最多一次 Cholesky 分解失败）。单次运行、一个序列和一次行走。

V. 机器人范围测试

我们将机器人视为范围测试，而非部署：ETH3D 上的每项发现能否在硬件上的 8 位摄像头上复现，以及决定其能否复现的因素是什么？整个系统（包括两组权重集、进程内后端、滑动窗口因子图和检测器日志记录）均运行在由四足机器人 Boston Dynamics Spot 搭载的 Jetson AGX Orin 上，并回放由其夹爪相机记录的行走过程（图 6）。该相机输出 8 位 ISP 处理的视频流（平台上的任何相机都无法提供原始数据）；其控制参数均无单位。自动曝光、对焦、HDR 和白平衡均处于关闭状态，因此每次行走过程中的光照级别都是受控变量；与相机自身自动曝光的比较留待未来工作。帧以 10 Hz 的频率不压缩存储；参考轨迹是机器人在每幅图像采集时刻的里程计，即一个参考，而非真实值（漂移约为 1%）。

实验方案。在一个房间内，沿一条用胶带标出的往返路线行走，路线长约 24 米，每次行走包含 1 190–1 300 帧。先在灯光开启时行走两次，再在灯光关闭时以四个曝光级别进行测试，并提高增益，使帧中值保持在 40 左右：1/13 秒（L1）、1/40 秒（L2）、

表 II 机器人梯子、双臂、载荷 AGX Orin。每个级别行走一次（重复测试得出相同结论）。关键帧路径：关键帧覆盖的里程计距离 / 行走距离。 g 每次调用（后端/重定位）。J：每帧能耗。

级别	arm	关键帧数 重定位 成功/尝试	m g	关键帧路径（米）	J
灯光开启	原版	12 0/0	0.36 0.49	22.0/25.0	13.7
	+Dark3R	10 0/0	0.41 0.41	19.6/25.0	13.3
L1 1/13 秒	原版	17 0/0	0.37 0.49	21.6/23.3	13.6
	+Dark3R	10 0/0	0.36 0.44	19.8/23.3	13.4
L2 1/40 秒	原版	26 0/0	0.29 **0.72**	21.7/22.7	14.0
	+Dark3R	18 0/0	0.32 0.61	21.7/22.7	13.8
L3 1/100s	原版	58 2/539	**0.07 0.86/0.88**	19.1/25.0	33
	+Dark3R	49 1/459	**0.000** 0.45/0.49	**11.7**/25.0	31
L4 1/200s	原版	**2** 1/1297	0.000 –/0.73	0/23.9	41
	+Dark3R	**4** 0/1295	0.000 0.80/0.76	0/23.9	56

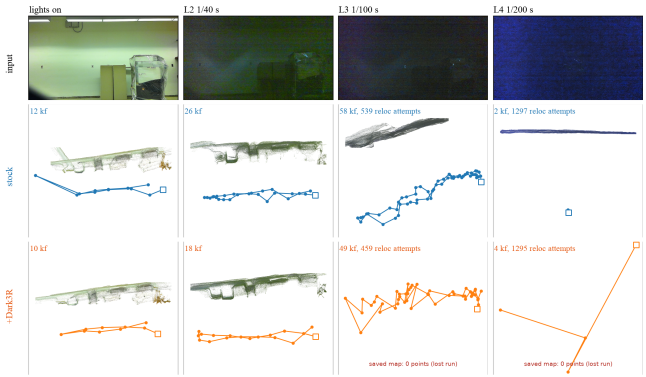

图 6. 机器人的梯子和第一个房间。上：夹爪相机捕获的五个级别中的四个（灯光开启时曝光 1/100 s、增益 0.05；L2 为 1/40 s，L3 为 1/100 s，L4 为 1/200 s，增益均为 0.40）。下：载荷 AGX Orin 上保存的重建和每次回放的关键帧位姿，分别展示原版和适配版；每次运行均采用各自的尺度（方形：起点）。双臂均能通过 L2 建立路线地图；在 L3，双臂均丢失路线并重定位数百次；在 L4，双臂均无法建立路线地图。

1/100 秒（L3，行走两次）、1/200 秒（L4）。每次行走都在载荷板上使用两组权重集回放（未经校准，窗口为 120，128槽存储区），并记录关键帧、重定位、 m、 g （每位调用者）、帧率和功耗；判断一次运行时，首先看其是否成功跟踪，然后再根据 g 进行评估。

机器人表现如何 （表 II，图 6 和图 7）。双臂都能在 *L2* 下跟踪，且 g 评估双臂所处环境的暗度：两条机械臂都没有重定位， g 在跟踪保持稳定时上升（原版为 0.49、0.49、0.72；适配后为 0.41、0.44、0.61）。原版在 L2 下正常行走的读数为 0.72，高于 0.65 的工作点，这是对正常运行的误报：级别是按相机校准的。在 *L3* 下，双臂都明显丢失跟踪：原版重定位 539 次，仅成功 2 次，其 m 降至 0.07；适配后的机械臂重定位 459 次， m 0.000，覆盖了半条路线。在 L4 下，双臂从最初几帧起就都丢失跟踪，且适配后机械臂的 g （0.80 / 0.76）也升至 ETH3D 工作范围以上：1/200 秒会使任一模型退化。面对同样明显的失败， g 反映出原版的崩溃和适配后机械臂未崩溃：在 L3 下，原版的 g 为 0.86 / 0.88，落在 ETH3D 的退化范围内；该相机并未用于建立任何范围。适配后机械臂的数值为 0.40–0.49；即使在这里，适配权重也消除了自我一致性，机械臂转而因置信度不足而失败。在这架梯子上，补救办法并未显现：不存在原版失败而适配后机械臂仍能跟踪的级别；半剂量下也没有，此时三个相同级别的结果仍是如此（短暂丢失

表 III

三条路线位于更明亮的房间内，且从不返回起点；双臂；512 槽存储。覆盖率：关键帧所跨越的里程计路径占已行走路径的比例。明亮对照场景下，双臂的覆盖率均为 93–97%。1 毫秒时的结构信噪比： $+4.7$、 -0.3和$+9.1$ 分贝（第一个房间的失效级别： -4.6）。

路线	快门	原版 关键帧数/1k 重定位	覆盖率%	+Dark3R 关键帧数/1k 重定位	覆盖率%
L形，12.5 米，沿 2.5 毫秒		13 0/359	**41**	123 0/0	**100**
	1 ms	110 0/0	100	152 0/0	100
横向移动，6.7 米，穿过	2.5 毫秒	538 0/129	**48**	298 0/0	**100**
	1 ms	28 2/298	0	90 0/246	25
横向移动，9.4 m，横穿	2.5 ms	238 0/0	100	73 0/0	95
	1 ms	120 0/281	0	30 0/387	0

在 L1 处），丢失相同的两个关键帧；在 L3 处，随着重定位追加的关键帧，溢出 128槽存储区。全部十四次回放在桌面端重复运行，关键帧数量完全相同，并且 g 在 0.04 以内。

三条从不回到起点的路线（表 III）。在更明亮的房间中，沿三条路线分别以明亮条件、2.5 ms 和 1 ms 快门、0.40 增益进行行走；1 ms 是相机的下限，此时静止起点的结构信噪比为 $+4.7$、-0.3 和$+9.1$ 分贝，而第一间房间发生失败的级别为 -4.6 分贝，因此这些路线中只有一条接近那种暗场景。一条路线呈 L形，相机朝向行走方向；另外两条是横移，相机朝向横向。所有明亮条件下的对照均能完成重建。由此可以得出三点结论。在 *2.5 ms* 下，适配覆盖了原版所丢失的一个级别，三条路线中有两条如此：原版分别保留路线的 41% 和 48%，并产生 359 次和 129 次失败的重定位尝试；适配后的机械臂则保留全部路线；第三条路线中两者都能跟踪。在 *1 ms* 下，两个机械臂在两条横移路线上都出现明显失败（覆盖率为 0–25%，尝试次数为 246–387， m 0.000，98–100% 的边被拒绝，每帧 45–57 J，而明亮条件下为 14–15 J）。在这四次失败运行中的三次里，所有关键帧都在静止起点期间创建，机器人一移动，跟踪便失败，因此其后端统计数据描述的是静止相机。在 *1 ms* 下的 *L*形路线上，两者都以另一种形式悄然失败：尝试次数为零， g 为其明亮级别下的数值；原版的关键帧速率是明亮条件下的五倍，并生成了一张看似完整的地图，但其终点到起点的范围相对于里程计的比例，随着曝光和增益降低而下降：明亮条件下为 0.86–0.87，2.5 ms 下为 0.55（适配后的 100% 地图），1 ms 下为 0.43–0.48，低增益下为 0.38。这是唯一一条相机朝向行走方向的路线上的隐性表现，而这个方向正是单目相机观测能力最弱的方向。 g 仅在三条路线中最暗的一条上呈现出相反的顺序：在 -0.3 分贝的横移路线上，原版在 1 ms 下的崩溃读数为 0.556，而其明亮条件下为 0.251（一次静止起点运行；该处移动中的适配运行则为 0.387，而明亮条件下为 0.257）；相比之下，在较明亮的两条路线上，每次失败运行的读数为 0.17–0.36，而明亮条件下的对照为 0.32–0.36。

代价，以及我们不作何种声称。跟踪两条机械臂时，代价相同：在 34–35 W 功耗下，每帧 13.2–13.9 J，相对于 10 Hz 录制速度为 2.6 FPS（表 II）；明显故障在 L3 级别下每帧耗能 22–34 J，在 L4 级别下为 41–56 J，是跟踪代价的 1.6–4× 倍，帧率为 0.75–1.7 FPS。相对于里程计的绝对轨迹误差（10–26 个位姿范围内为 0.24–0.73 米）并非一项结果。四个房间、四条路线、十八次行走，每一行都只对应一次运行；未测量精度。

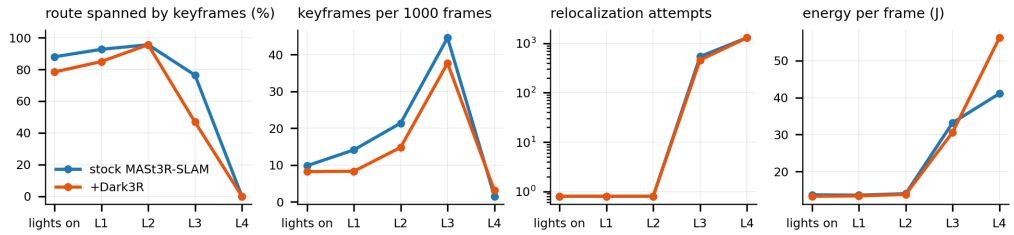

图 7。机器人在五个光照等级下的数字化梯子（表 II），涵盖双臂：由关键帧覆盖的路线（关键帧路径与行走路径之比）、每 1000 帧的关键帧数、重定位尝试次数（对数刻度；零值绘制在底部）以及根据完整功率轨迹计算的每帧能耗。在 L3 之前，双臂没有任何区别；到了 L3，双臂都出现明显失败。适配后的机械臂的关键帧覆盖了一半路线，而这种明显失败使每帧能耗增加了 2–4× 倍。

VI. 结论与局限

基于学习式双视图几何构建的稠密视觉 SLAM，在低光照下会因一种机制而失效：网络跨视图几何的伪自洽性使匹配统计量膨胀，超过两个子系统以相反方向读取的阈值，从而产生近乎空白却被报告为正常的地图，或在错误锚点上进行无界重定位。导致这一问题的量本身就是检测器：无需计算，也无需真实值。弱光适应通过降低置信度，而非找到正确几何，在 ETH3D 上修复了这两种症状；其良好光照下的代价随剂量增加，而检测器检测的是修复效果和良好光照下的退化，并非绝对轨迹误差的代价。在一台采用 8 位相机的四足机器人上，这一机制的特征在四条路线中最暗的两条上重现；在三条无转弯路线中的两条上，适配覆盖一个暗场景级别；而故障表现出的不同面向，则随路线几何和传感器编码而异。

局限性，每项仅陈述一次。 g 无法证明几何结构正确：工作带读数意味着网络并非退化性自洽，并不意味着地图正确（合并/沙发_暗场景_1：0.371，97% 错误）。**证据仅在关键帧处获得：** 一次无声运行中只有 2–6 次读数，首次读数出现在 5.6–10.5 秒之后；逐帧对应结果受短基线主导；原计划的宽基线探测未能实现。**范围。** g的证据来自 8 位 ETH3D 无声失败场景：其结果无法迁移到合成噪声场景，也无法观察到置信度崩塌（bfloat16、原始输入、机器人）或压缩地图；其健康水平按相机分别校准，且其工作点属于样本内。**单一系统；警报仅执行过一次。** 其他点图 SLAM 系统是否具有相同机制尚未测试；未运行基于特征或惯性的基线。两段明亮的无纹理序列在双臂上均读得 g0.31–0.34：低纹理不会抬高自洽性；纯旋转尚未测试。第 IV-F 节中由警报触发的切换，在每种设置下仅运行一次且只向前切换一次；每次运行期间光照保持恒定，因此光照恢复时能否恢复尚未测量。**其他局限。** 仅测试了一种增强器和一种适配方法；能耗比为范围值（确定性无法在功耗模式改变后保持）；里程计仅作为参考，未记录深度。代码和合并流程将发布；合并后的检查点无法使用（知识共享署名-非商业性使用-相同方式共享基础权重）。

致谢

使用了 AI 助手（ClaudeOpus 5）来起草和编辑文本，并编写生成图 1–7 的脚本；作者审阅并核实了所有此类输出。相关构想、实现、实验、测量、分析和结论均由作者独立完成。

参考文献

[1] R. Murai、E. Dexheimer 和 A. J. Davison，“MASt3R-SLAM：基于三维重建先验的实时稠密 SLAM”，载于 CVPR，2025 年。[2] L. Hu 等，“EC3R-SLAM：基于前馈三维重建的高效且一致的单目稠密 SLAM”，arXiv 预印本 arXiv:2510.02080，2025 年。[3] P. Kumar、A. Deepak 和 S. Kumar，“面向 GPS 拒止及退化环境中无人机导航的鲁棒视觉 SLAM：多范式评估与部署研究”，arXiv 预印本 arXiv:2605.03678，2026 年。[4] A. Y. Guo 等，“Dark3R：学习暗场景中的运动恢复结构”，载于 CVPR，2026 年。[5] E. J. Hu 等，“LoRA：大语言模型的低秩适配”，载于 ICLR，2022 年。[6] T. Schöps、T. Sattler 和 M. Pollefeys，“BAD SLAM：束调整直接 RGB-D SLAM”，载于 CVPR，2019 年。[7] S. Wang 等，“DUSt3R：让几何三维视觉变得简单”，载于 CVPR，2024 年。[8] V. Leroy、Y. Cabon 和 J. Revaud，“使用 MASt3R 将图像匹配锚定到三维空间”，载于 ECCV，2024 年。[9] J. Wang 等，“VGGT：以视觉几何为基础的 Transformer”，载于 CVPR，2025 年。[10] Y. Wang 等，“π^3：置换等变的视觉几何学习”，arXiv 预印本 arXiv:2507.13347，2025 年。[11] D. Maggio、H. Lim 和 L. Carlone，“VGGT-SLAM：在 SL(4)流形上优化的稠密 RGB SLAM”，载于 NeurIPS，2025 年。[12] A. S. Thomas、S. Agnihotri 和 M. Keuper，“失败还是漂移？评估合成扰动和真实世界扰动下的单目 SLAM”，载于 ECCV 的 NeuSLAM研讨会，2026 年。[13] O. Basystyi 等，“低照度环境中的 SLAM：项目报告”，arXiv 预印本 arXiv:2607.17699，2026 年。[14] S. Rabiee 和 J. Biswas，“IV-SLAM：用于同步定位与建图的内省式视觉”，载于 CoRL，2020 年。[15] J. A. Gaus、D. Häufle 和 W.-J. Baek，“SUPER——用于视觉惯性里程计的基于敏感度、感知不确定性的性能与风险评估框架”，arXiv 预印本 arXiv:2512.14189，2025 年。[16] S. Sansoni 和 S. R. Tosetti Sanz，“SLAM置信度陷阱”，载于 IEEE 国际高级机器人会议（ICAR），2025 年，第 733–738 页。[17] X. Xiong 等，“通过几何效用评分加速基于Transformer的单目SLAM”，arXiv 预印本 arXiv:2604.08718，2026 年。[18] M. Kweon 等，“DarkVGGT：利用热成像几何穿透黑暗，无需日光代价”，arXiv 预印本 arXiv:2606 11326，2026 年。。[19] S. P. Singh 等，“Twilight SLAM：在低照度环境中导航”，arXiv 预印本 arXiv:2304.11310，2023 年。[20] D. Feijoo 等，“DarkIR：鲁棒的低照度图像复原”，载于 CVPR，2025 年。[21] D. Kong、Y. Zhang 和 W. Dai，“采用主动照明的直接近红外-深度视觉 SLAM”，IEEE机器人与自动化快报，第 6 卷，第 4 期，第 7057–7064 页，2021 年 10 月。[22] N. Shankar、P. Ladosz 和 H. Yin，“CLEAR-IR：红外图像的清晰度增强主动重建”，arXiv 预印本 arXiv:2510.04883，2025 年。