arXiv:2611.30166v1 [cs.RO] 23 Nov 2026

借助深度—惯性里程计，迈向采用消费级激光雷达的开放式房间尺度扫描

Minsuk Jang

KAIST，大田，大韩民国。电子邮箱：minsukjang@kaist.ac.kr

摘要

如今，使用消费级激光雷达进行房间扫描依赖于 ARKit，而其跟踪和地图更新过程是一个黑箱：应用程序能够接收深度和位姿，却无法查看这些数据是如何生成的，因此扫描中的误差无法追溯其原因。本文研究了一种开放式替代方案在设备上是否可行，即其每个阶段都可以被检查。Mobile-LIO 是一种深度惯性里程计处理流程，使用手机的深度和惯性数据，不使用 ARKit位姿。在七段 ScanNet++ iPhone 录制内容中，它完成了每个场景，与 ARKit 派生参考的偏差为 0.2–2.0 m；相比之下，六种按照专用扫描仪配置运行的开放式激光雷达-惯性估计器，在 42 次运行中的 41 次至少出现了 4.46 m 的误差（六种估计器中有两种还未通过无噪声适配器检查，而 Mobile-LIO 设有它们没有的速度上限；该比较并非受控实验，也不对算法进行排名）。在 iPhone 15 Pro Max 上，经过手机代码修正的版本在三段时长为 7–13 秒的录制内容中的 15 次回放运行里，有 14 次能够跟上传感器；数据集结果使用的是修正之前的估计器。激光扫描网格展示了这一结果与封闭式处理流程之间的差距。在 3 m 以内，记录深度与网格的差异中位数为 1–5 cm；在三个留出场景中，与经数据集缩放并对齐到网格的 ARKit位姿融合得到的地图，与网格的距离中位数为 1.8–2.3 cm，而 Mobile-LIO 地图为 8–17 cm。由于该处理流程是开放的，因此可以进一步检查这一差距：在误差最大的场景中，无噪声的渲染深度将偏差从 1.97 m 降至 0.04 m；因此，记录深度也与估计器一起造成了误差，其在 5–7 m 处的误差达到 20–54 cm。因此，在设备上进行开放式房间尺度扫描是可行的，但其准确度尚未达到 ARKit 位姿地图的水平；而后者本身也只是初步达到了 USIBD 最粗略的准确度等级（LOA10）。

1 引言

建筑物几何数字孪生的价值会随着资产变化而逐渐降低，除非重新采集其几何信息（Osadcha 等，2023；Davletshina 等，2026）。地面激光扫描能够达到扫描到建筑信息模型所需的精度，但每次更新的成本都很高。在一项针对约 350 m^2 室内空间的对比研究中，地面激光扫描与人工建模工作流程耗时 20 h，而基于平板电脑的工作流程仅需 32 min，但几何偏差更大（Ersoz 和 Bosché，2026）。因此，低成本、可重复的采集方式颇具吸引力。

地球科学、地球科学和土木工程领域的研究评估了配备激光雷达的 iPhone 和 iPad 在室内制图与建筑文档（Teo 和 Yang，2023；Askar 和 Sternberg，2023；Abdel-Majeed 等，2024；Treccani 等，2024）以及地球科学测量（Luetzenburg 等，2021）中的输出。这些输出背后的位姿跟踪和地图更新由专有的 ARKit 框架实现，该框架提供深度、置信度和相机位姿信息，但不公开其内部决策过程。如果导出结果显示物体边缘处的深度模糊不清，或走廊环路无法闭合（Jang 和 Yu，2026），用户便无法追溯误差产生于哪个阶段。

因此，应用层面的精度数值综合反映了深度传感器的误差、估计位姿的误差以及融合软件的选择。开源激光雷达-惯性里程计（LIO）根据范围和运动传感器估计移动设备的位姿，提供了可供检查的估计结果，但具有代表性的系统都是为专用扫描仪设计的（Xu et al. 2022；Chen et al. 2023）。智能手机深度信息属于不同的观测模式： 256×192 深度图像带有三级置信度图，且没有逐点计时。

本文探讨了通过开放式处理流程实现房间尺度扫描是否可行。所谓开放式处理流程，是指每个阶段都有明确说明且可供检查。本文还衡量了这种处理流程与

封闭式处理流程之间的差距，并将相关因素逐一分离。本文有五项贡献：（1）测量录制的 iPhone 深度相对于激光扫描网格的误差随范围变化的情况；（2）针对相同录制内容制定误差预算，给出将相同深度分别与网格对齐的 ARKit 位姿及 Mobile-LIO 位姿融合时的表面误差。Mobile-LIO 是一种可检查的深度惯性处理流程，其设置列于附录；此外，还测量了 Mobile-LIO 相对于 ARKit 派生参考的偏差；（3）借助可检查的处理流程开展受控实验，固定估计器，仅改变深度输入；（4）在相同输入上与六种开放式激光雷达-惯性估计器进行比较；（5）开展辅助检查：代码修正前后的设备端回放，以及由手机应用导出的定性扫描结果。研究结果与建筑文档编制所采用的精度等级相关。本文的研究范围限于单个室内场景和无回环闭合的局部估计（即路线返回已知地点时对漂移进行修正）；设备端证据来自同一部手机上回放同一房间的三段短录制内容（评估几秒内的耗时和与 ARKit 的一致性）、由手机应用导出的定性扫描结果，以及与扫描仪地图进行比较的一段走廊；这些内容独立于留出评估。

2 相关工作

建成环境中的智能手机激光雷达。Teo 和 Yang（2023）以及 Abdel-Majeed 等人（2024）评估了 Apple 设备在室内制图中的应用；Askar 和 Sternberg（2023）比较了用于建筑文档的扫描应用程序；Luetzenburg 等人（2021）则对用于地球科学测量的 iPhone 12 Pro 传感器进行了表征。Treccani 等人（2024）在一条走廊和一条街道上，将五款 iOS 应用程序与地面扫描仪进行了测试，发现误差约为 1 cm 至约 10 cm。为更新竣工BIM，Yang 等人（2024）使用低成本移动式激光扫描仪和惯性单元取代了地面扫描。Díaz-Vilariño 等人（2022）评估了 iPad Pro 对两个房间的扫描结果在精密度、正确性和覆盖率方面的表现；Teppati Losè 等人（2022）和 Abdelghany 等人（2026）则测试了该传感器在遗产记录中的应用。常见的流程是将导出的点云与地面扫描结果配准，通常采用迭代最近点（ICP）算法，并报告点云间距离，有时也报告尺寸。这类比较反映的是最终产品精度，无法区分深度误差与跟踪误差；本研究将深度误差与开放式深度-惯性估计器的跟踪误差区分开来，但无法将其与 ARKit 自身的跟踪误差区分。

激光雷达-惯性里程计与建筑施工SLAM。FAST-LIO2 在迭代卡尔曼滤波器中将惯性传播与直接点到地图配准相结合（Xu et al. 2022）；Faster-LIO、Point-LIO 和 DLIO 是其他开放式估计器（Bai et al. 2022；He et al. 2023；Chen et al. 2023）。在建筑施工领域，ConSLAM数据集用于评测手持式 SLAM（Trzeciak et al. 2023），Keitaanniemi et al.（2023）则分析了室内SLAM激光扫描中的漂移；两者均使用专用扫描仪。在手机上，Jin 和 Ye（2023）展示了在 iPhone 12 Pro 上实现的视觉-激光雷达-惯性里程计。Jang 和 Yu（2026）提出了 Mobile-LIO，其中采用置信度感知选择和预算滤波器处理 iPhone 深度，并包含速度反馈和闭环校正组件。本文仅评估其局部估计器，不包含这些组件；此外，还加入附录所列的设置、针对激光参考的留出评估以及误差预算。

3 方法

图 1 展示了处理流程。一个 C++ 估计器同时服务于 iPhone 应用和桌面回放工具，后者会重新运行已录制的序列；两者都不使用 ARKit 的相机位姿。估计器根据手机的运动传感器预测其运动，并根据每幅新深度图像与已映射表面的对齐程度来校正预测结果。

输入缩减。 每个 256×192 深度帧均以四个像素为步长进行反向投影，最多可得到 3,072 个点。置信度为 0、坐标为非有限值或范围超出限制（10 米或

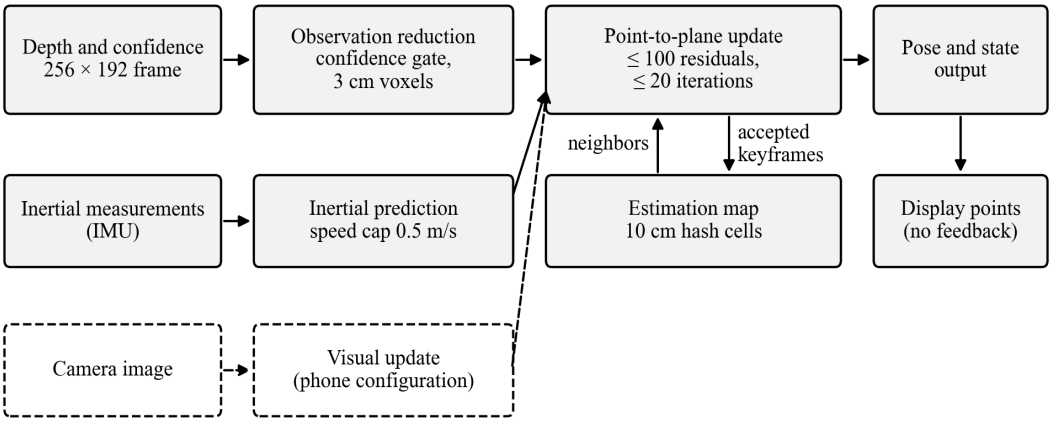

图 1. Mobile-LIO 处理流程。实线框构成深度-惯性配置；虚线框为手机配置中新增的部分。

3 米，具体取决于配置）的样本将被丢弃。保留的点被分组到 3 厘米体素中，并将每个体素替换为其质心和平均置信度。

位姿估计。 运动传感器采样在深度帧之间传播姿态方向、位置、速度、传感器偏置和重力。传播后的速度会被限制在 0.5 m/s；由于深度无法直接观测速度，该上限可使速度估计保持有界，但也可能截断实际的较快运动。随后，将每个降采样点与一个平面进行比较，该平面由地图中 20 厘米半径内的点拟合而成（至少六个点；地图每个 10 厘米网格最多保留 20 个点），该点到平面的距离即为残差。大于 10 厘米的残差会被剔除，其余残差根据置信度加权，噪声水平设为 2 厘米；最多有 100 个残差进入卡尔曼更新，该更新最多重复 20 次，每次都重新匹配平面；增益始终使用预测协方差，而协方差仅在最后一次迭代后更新一次。与 FAST-LIO2 中的迭代滤波器不同，重复更新不会将每次迭代结果拉回预测状态，因此这是重复的卡尔曼校正，而非迭代滤波器。结果章节中的所有结果均使用此更新；其余设置见附录。总校正量超过 1 m 或 0.5 rad 时，将予以丢弃。

地图与恢复。 前五帧在传播得到的位姿处初始化估计地图。初始化后，只有在关键帧（移动 5 厘米或 2° ）且更新被接受时，点才会加入地图。连续十次更新失败后，地图会被清空并重新初始化；显示的点不会反馈到估计过程中。

两种配置。 深度-惯性配置使用量程为 10 米的深度数据，且仅使用运动传感器（惯性测量单元，IMU）数据；在运行留出场景之前，相关设置已固定。手机配置则采用应用程序源代码中指定的输入：3 米量程限制，以及基于跟踪图像特征的视觉更新（也运行了不使用视觉更新的版本）；此处禁用了初始陀螺仪偏差估计，因为录制内容并非从静止状态开始。该配置是在获知首批结果后才用于留出场景的，因此属于事后进行的第二轮考察。回放时，它会将惯性轴映射到点所在的 ARKit相机坐标系；本研究开始前收到的应用程序源代码版本中，这些轴绕光轴旋转了180° ，之后在手机上进行的检查证实了这一点。

4 实验设计

数据。 ScanNet++ 为每个场景提供 iPhone 深度、图像、惯性数据、ARKit位姿、激光扫描网格，以及经过缩放和变换到网格坐标系中的同一组 ARKit位姿（Yeshwanth等，2023）；本文将后者称为数据集位姿，并将其作为所有轨迹指标的参考。它们并非独立的

真实值。在本地可用的 50 个场景中，有 7 个已用于开发或早期评估，因此被排除；评估所用场景均为留出场景。每次运行前记录了一项规则，用于从其余场景标识符中选出最先通过预先声明的文件完整性、时长、帧间隔、惯性一致性和初始运动检查的 3 个场景。查看结果后未替换任何场景。每隔 6 帧取一帧，因此深度数据频率为 10赫兹，惯性数据频率为 60 赫兹。场景 S1 至 S3（ScanNet++ 场景 286b55a2bf、31a2c91c43 和 578511c8a9）分别有 900、1,270 和 2,549 帧，路径长度分别为 22、20 和 92 米，观测到的网格表面分别为 27、34 和 347 米^2 。转换后的数据不含原生置信度，因此每个样本的置信度均设为 2 级。

深度误差。 对于三个留出场景和七个开发场景，我们根据数据集位姿从激光网格渲染深度，并将其与记录深度逐像素进行比较；每个场景涉及 300 至 1,446 帧。记录深度是 ARKit 输出的深度图，而非传感器原始输出。我们按记录范围对误差进行分箱。

位姿来源与轨迹指标。 位姿来源包括数据集位姿、每帧随附的 ARKit 原始位姿（用于一致性检查，因为数据集位姿由其推导而来）、两种 Mobile-LIO 配置，以及使用适配器并采用相同深度点、惯性样本和时间区间运行的 FAST-LIO2、Faster-LIO、Point-LIO 和 DLIO；适配器采用单位外参和零逐点时间偏移。在无噪声的合成序列上，该适配器对这四种估计器产生的误差均低于 1 cm。另有两种开放估计器 iG-LIO 和 LIO-Livox 在部分此类检查中误差较大；表 6 列出了它们的结果，并附有这一说明。绝对轨迹误差（ATE）是经过一次不含尺度的刚性对齐后计算的平移均方根误差。如果某次运行的平移误差在 1 s 内首次超过 0.5 m，则称其丢失跟踪。若某次运行丢失跟踪且最终 ATE 高于 4 m，则称其发散。

表面指标。 对于每个场景，使用数据集位姿和 Mobile-LIO位姿分别融合相同的深度帧，范围为 0.25–4.0 m，每隔一个深度像素取样，体素质心大小为 2 cm，且每个体素至少有两次命中。通过用于计算绝对轨迹误差的刚性对齐，将 Mobile-LIO轨迹置于数据集坐标系中，不进行地图到网格配准，因此距离同时包含残余的全局错位和漂移。计算每个地图点到激光网格的精确点到三角形距离。覆盖率是指：在数据集位姿地图中观测到的网格采样点里，距离待评估地图不超过 5 cm 的采样点所占的比例。

受控深度输入。 在 S3 上，固定深度-惯性估计器，仅改变其深度输入：将记录深度限制在 10、5 或 4 m；记录深度仅保留与网格相距 5 cm 以内的像素（这是一项需要参考信息的诊断）；以及根据网格、按照手机视场和两个更宽视场渲染的无噪声深度。

5 结果

记录深度在 3 m 以内是准确的。 记录深度在近距离时准确，远距离时则不可靠（图2，表1）。在十个场景中，每个场景的绝对误差中位数在 0–1 m 范围内为 0.9–2.2 cm，在 2–3 m 范围内为 1.6–4.8 cm，在 4–5 m 范围内为 2.9–10.3 cm，在 5–7 m 范围内为 19.7–53.5 cm。在 3 m 以内（约占 1420 万个比较像素的 90 %），记录深度的读数最多比实际值低 2 cm，且误差的第90百分位数为 3–32 cm，因此即使是近距离深度也存在长尾误差。S1 和 S2 中几乎没有超过 3 m 的像素，而 S3 中有 16.6% 的像素超过 4 m。

Mobile-LIO 与 ARKit 派生参考的偏差为 0.2 至 2.0 m。 刚性对齐后，原始 ARKit 位姿与数据集位姿相差 3–8 cm（表2）；由于数据集位姿是在这些 ARKit 位姿上应用尺度与变换后得到的，因此此处衡量的是该调整，而非 ARKit 的准确度。深度—惯性 Mobile-LIO 配置完成了全部三次录制，绝对轨迹误差分别为 0.409、0.200 和 1.973 m。采用视觉更新的手机配置达到 0.435、0.170 和 0.590 m；不采用视觉更新时，它

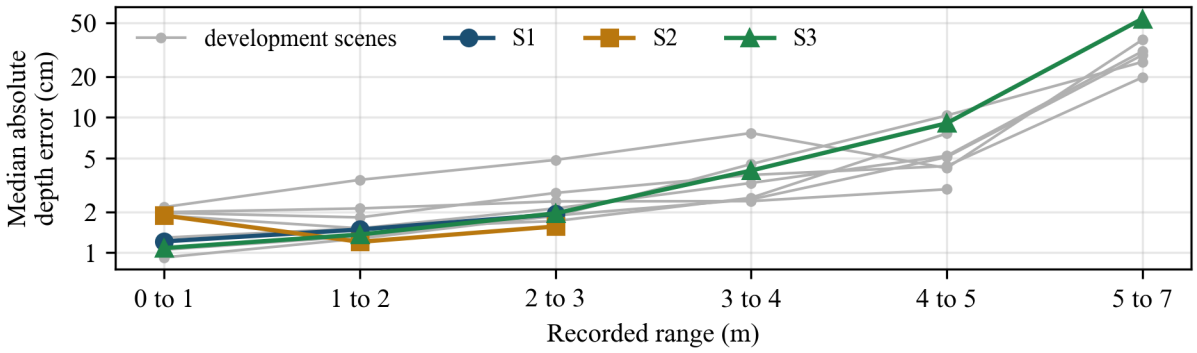

图2。十个场景中，按范围统计的记录 iPhone 深度与激光网格渲染深度之间的绝对误差中位数。

表1。十个场景中按范围统计的记录深度误差。误差列给出了各场景数值的范围；带符号中位数为记录深度减去渲染深度。

范围（米）	像素（%）	带符号中位数 (cm)	绝对值 中位数（厘米）	P90（厘米）	高于5厘米 (%)
0–1	23.9	−1.7 至 0.0	0.9–2.2	2.8–26.5	3–32
1–2	48.6	−1.6 至 −0.5	1.2–3.5	3.7–28.3	5–41
2–3	17.2	−2.0 至 −0.6	1.6–4.8	4.0–32.3	6–49
3–4	5.7	−1.8 至 −0.2	2.4–7.6	8.3–42.2	21–58
4–5	2.0	−3.8 至 −1.1	2.9–10.3	10.3–122.0	30–68
5–7	2.4	−45.1 至 17.0	19.7–53.5	50.9–160.3	82–94

在 S3 中达到 2.920 m，因此，大场景中的改进指向视觉更新，而非量程限制（图3）。根据跟踪丢失准则，深度-惯性配置在 S1 中于 19 s 首次超过阈值，在 S3 中于 35 s 首次超过阈值，在 S2 中则从未超过；它并未发散，但其 1 s 区间中分别有 13%（S1）和 19%（S3）超过 0.5 m。通过适配器检查的四种估计器均发生了发散（六种方法的对比见表 6）：DLIO 在 5 s 内超过阈值，FAST-LIO2 在 23–73 s 后超过，Faster-LIO 在 35–81 s 后超过，Point-LIO 在 32–67 s 后超过；它们最终的绝对轨迹误差均在 7 m 至数十公里之间。该对比并非受控实验：Mobile-LIO 的速度上限为 0.5 m/s，而其他方法没有这一限制；没有任何录制是从静止状态开始的，尽管 DLIO 的标定假设起始状态静止；此外，在开发场景中，移除速度上限后，Mobile-LIO 的七次运行中有三次发生了发散。

表 2. 留出场景中与数据集位姿的偏差（ATE，米）。数据集位姿由 ARKit 原始位姿推导得出。

位姿来源	S1	S2	S3
ARKit，原始（与参考相同的来源）	0.030	0.033	0.082
Mobile-LIO，深度惯性	0.409	0.200	1.973
Mobile-LIO，手机配置，无视觉信息更新	0.441	0.230	2.920
Mobile-LIO，手机配置	0.435	0.170	0.590

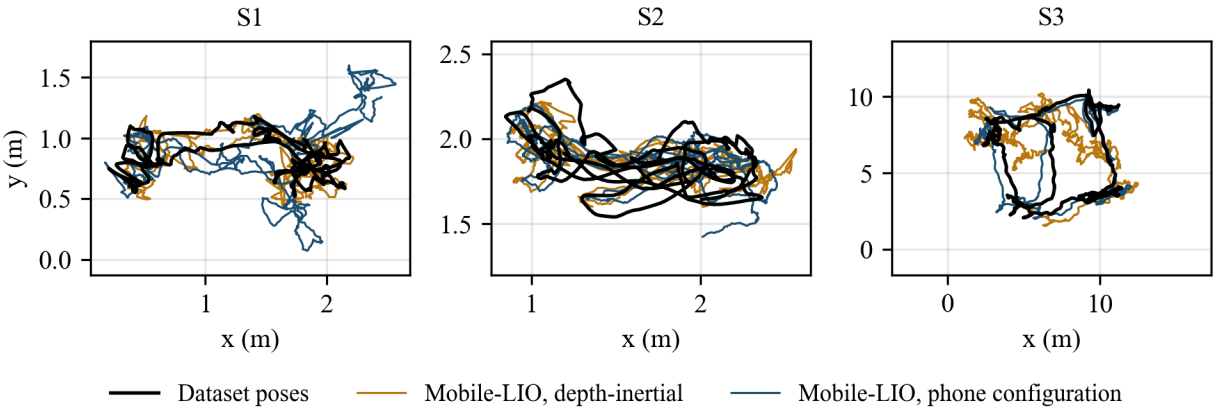

图3。顶部— 在与数据集位姿进行刚性对齐后，留出场景上的向下轨迹 s.

ARKit 位姿地图与激光网格的距离约为 2 cm；Mobile-LIO 地图则为 8 至 17 cm。使用 Mobile-LIO位姿时，位姿误差而非深度决定了大部分中位表面误差（表3、图4）。使用数据集位姿（即网格坐标系中的 ARKit 位姿）时，中位距离为 1.84–2.29 cm；这就是 ARKit 跟踪结合数据集的尺度校正以及该深度和融合规则所达到的效果。USIBD 精度等级区间以 95% 置信水平规定：LOA20 覆盖 15 mm 至 5 cm，LOA10 则粗于 5 cm（USIBD 2016年）。在数据集位姿地图中，17–27% 的点位于 5 cm 之外，因此，与对齐后的 ARKit 位姿融合的地图可指示性地对应于 LOA10；使用 Mobile-LIO位姿时，68–83% 的点位于 5 cm 之外。所有地图点的最近表面距离并非正式 LOA 验证，因此这一归类仅具有指示性。使用深度-惯性 Mobile-LIO 位姿时，该距离上升至 8.35–17.48 cm，且覆盖率在

5 cm 处降至 34–58%。手机配置改善了 S3（中位距离由 17.48 cm 降至 12.52 cm，第90百分位数由 63.99 cm 降至 34.89 cm）以及 S2 中的覆盖率，但在 S1 中表现更差，因此不能称其总体上更优。S2 数据集位姿地图的第90百分位数为 42 cm；图4 显示有些点穿过了一个开口，而网格可能未延伸至该处，因此该尾部可能反映的是网格覆盖率，而非深度误差。该区域未进行掩膜处理。

表3. 对于数据集位姿（网格对齐的 ARKit）和两种 Mobile-LIO 配置，融合地图到激光扫描网格的距离（P90：第90百分位数）。根据定义，数据集位姿的覆盖率为 100%。

场景	位姿来源	中位数（厘米）	P90（厘米）	5以内的点 厘米（%）	5 厘米处的覆盖率 厘米（%）
S1	数据集	1.84	8.44	82.8	100.0
S1	深度惯性	8.35	25.49	31.7	57.7
S1	手机配置	13.75	52.90	21.3	59.6
S2	数据集	2.29	42.44	73.1	100.0
S2	深度惯性	9.28	36.95	25.6	51.0
S2	手机配置	9.32	27.65	28.0	62.6
S3	数据集	2.15	8.55	78.3	100.0
S3	深度惯性	17.48	63.99	16.9	33.7
S3	手机配置	12.52	34.89	21.2	38.6

理想化深度消除了 S3 中的大部分误差；而对记录深度进行滤波并未奏效。在修正估计器后，将记录深度限制在 5 米或 4 米以内并没有改善效果（图5），因此，仅凭范围无法解释这一较大误差。仅保留与网格相距不超过 5 厘米的像素后，绝对轨迹误差从 1.973 米降至 0.594 米；而在手机视场下使用无噪声的渲染深度后，绝对轨迹误差降至 0.042 米。使用两个更宽的视场时，该数值分别为 0.026 米和 0.016 米。渲染深度同时消除了边缘误差、缺失误差、反射表面误差以及与范围相关的误差；此外，它是稠密的，并且与数据集位姿完全一致，因此这是一个理想化的界限，不能据此判断这些因素中哪一个最为重要；该实验仅涵盖一个场景。在逐帧日志中，从起始位姿测得的位置误差在最初 60 秒内保持在 0.13–0.20 米之间（首次超过 1 秒判据是在 35 秒时），而当远距离像素的占比跃升至 30% 时，误差有所增长；这与远距离深度可能发挥作用相一致，但无法将其单独分离出来。

处理成本。 对于搭载 Ryzen 5 5600H 处理器的笔记本电脑上的深度-惯性配置，在 50 至 200 的残差预算范围内，对应关系搜索占平均帧时间的 95.0–99.6%；默认预算下的帧时间中位数为 188–489 毫秒，慢于 10赫兹 输入。手机上的计时结果将在下一节报告。

6 设备端回放

上述结果来自使用方法部分中的估计器对数据集录制进行桌面端回放。本节报告该估计器在手机上的运行情况，以及其手机代码经过校正后发生的变化。这是针对短录制内容进行的事后检查，与留出评估相互独立；此处所述的校正构建版本未在数据集上运行。

录制与协议。 使用 iPhone 15 Pro Max 录制了同一房间的三次手持扫描（6.9、8.4 和 12.7 秒；设备运动范围为 0.4–0.9 米；量程限制为 3 米；其中两次以 60 赫兹、一次以 30 赫兹提供深度）。一次录制恰好存储应用程序传递给估计器的内容：每

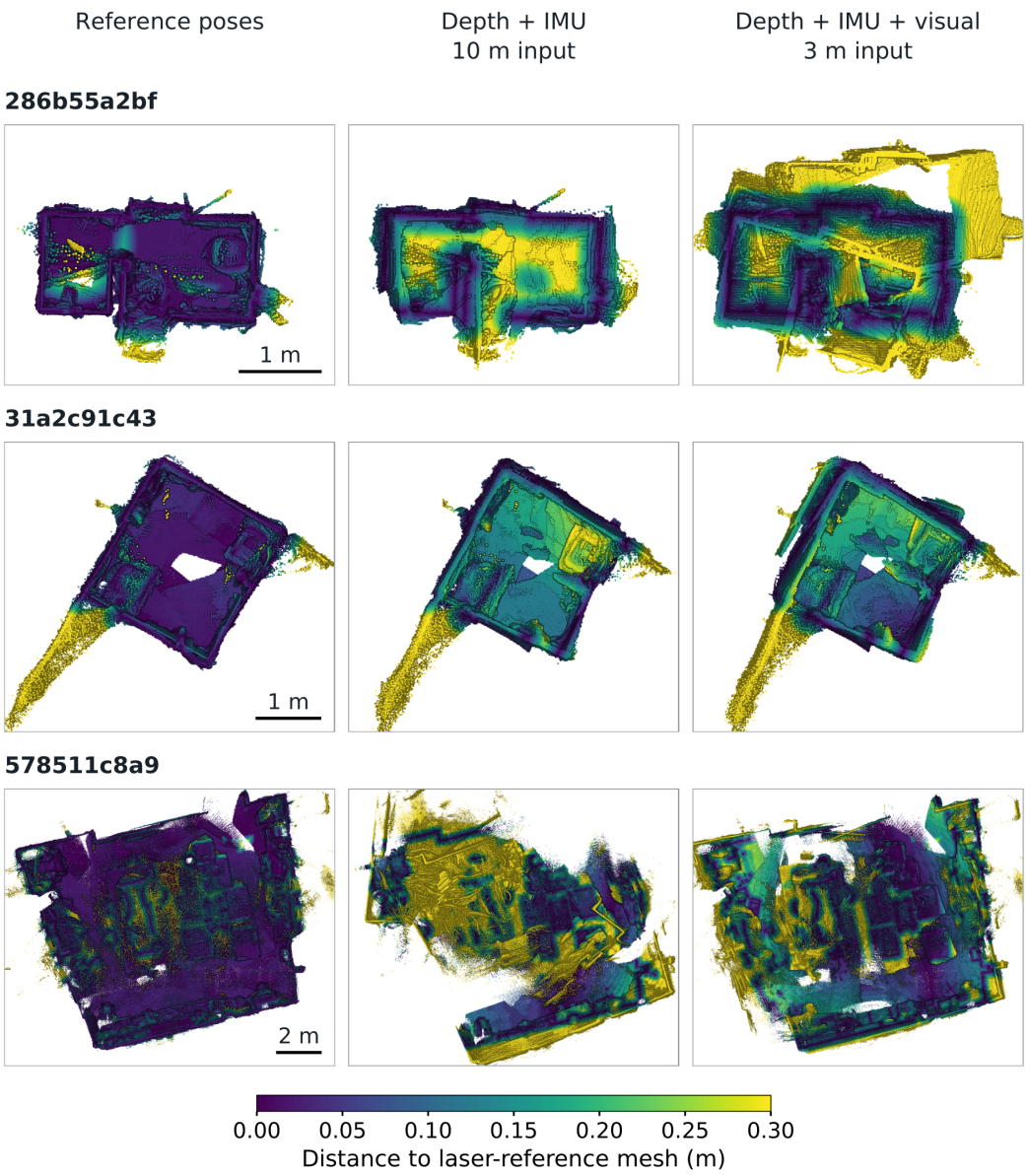

图4. S1至S3（从上到下）的融合地图与激光网格之间的距离，分别采用数据集位姿（网格对齐的 ARKit；标注的参考位姿）、深度-惯性 Mobile-LIO 位姿和手机配置的 Mobile-LIO 位姿（列标题：使用 10 m 输入的深度和 IMU；使用 3 m 输入的深度、IMU 和视觉信息）。行标签为 S1 至 S3 的 ScanNet++ 场景标识符。俯视图；色标在 0.3 m 处饱和。数据来源：Yeshwanth 等（2023）。

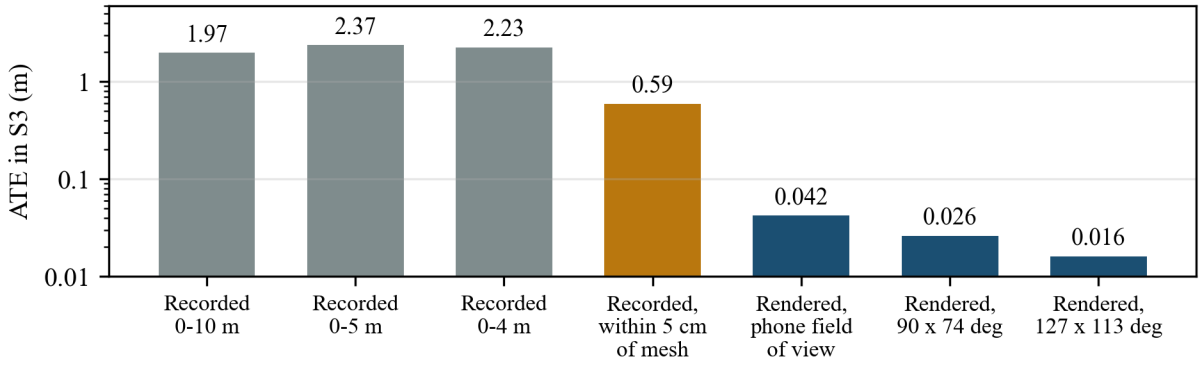

图5。固定估计器、仅改变深度输入时，S3中的轨迹误差。渲染深度是来自激光网格的无噪声深度。

帧、视觉更新所使用的 480 × 360 灰度图像和深度图，以及运动传感器采样。对于计时，手机以录制时的速度，将每段录制数据馈送给一个全新的估计器；每种录制和构建版本各重复五次，并在没有显示负载的发布版本中进行。估计器繁忙时到达的帧会被丢弃，这与实时扫描中的情况相同。延迟从交付一帧开始计算，直到获得其位姿为止，此外还加上实时会话中测得的从采集到交付的33毫秒中位延迟。运行前预先确定的规则规定：如果至少90%的已交付帧得到处理，且第95百分位延迟不超过100毫秒，则认为该次运行能够跟上传感器。位置经过一次刚性对齐后，与每段录制数据中记录的ARKit位置进行比较。ARKit使用相同的摄像头和运动传感器，并非真实值，因此在不足1米的运动范围内与其保持一致只能说明一致性，而非准确性；估计器从不读取ARKit位姿。

两个构建版本。 原始构建版本是本研究开始前的应用程序源代码状态。校正构建版本有三处不同。第一，运动传感器采样会从设备坐标系（纵向时 x 向右、y 向上、z 朝向用户）映射到深度点的 ARKit相机坐标系，映射方式为 (–y, x, z)；原始构建版本使用的是 (y, –x, z)，相当于绕光轴旋转了 180° 。在使用原始构建版本进行的四次实时会话中，估计的向上方向在开始时与 ARKit 的方向相差 165–177° ；在一次使用校正映射的会话中，两者相差 0.2°。第二，重复校正包含迭代卡尔曼滤波器的先验项，将每次迭代结果拉回预测状态。第三，每次迭代会针对约 300 个均匀间隔的候选点尝试进行平面拟合，而不是针对所有降采样点；当步长小于 0.2 mm 和 0.2 mrad 时停止迭代。这两项更改是同时进行的，无法分别区分其影响。

原始构建版本无法跟上处理进度；校正构建版本可以。原始构建版本每个关键帧耗时中位数为 192 ms，仅处理了已交付帧数的 24%，且没有一次运行满足该规则（表 4、图 6）。在设备移动距离小于 1 m 的录制内容中，其与 ARKit 的位置差均方根中位数为 23.5 cm。校正构建版本每个关键帧耗时中位数为 18 ms，处理了 98% 的帧数，第95百分位延迟为 65 ms，并且 15 次运行中有 14 次满足该规则；剩余一次运行处理了 89.9% 的帧数。配准仅在关键帧上运行，根据录制内容的不同，占已交付帧数的 13–35%；其他帧所需时间更短，因此，即使关键帧时间中位数略高于 16.7 ms 的帧间隔，仍能处理几乎所有帧数。校正构建版本与 ARKit 的位置差异中位数为 1.8 cm（各次运行范围为 0.8–4.5 cm）。在两次约 7 s 的实时扫描中，使用相同的估计器设置并存在显示负载时，处理的帧数比例分别为 83% 和 95%，因此，回放对于实时扫描的表现有些过于乐观。

哪项改动产生了什么影响。在台式机上回放相同的录制内容并处理每一帧，可以区分各项改动（表5）。在处理所有帧的情况下，即使是接收时的源数据，其误差也保持在1.6–7.8厘米以内

表 4. iPhone 15 Pro Max 上的设备端回放：15 次运行中每次运行值的中位数（范围）（三段录制内容，每段运行五次）。

测量	原始构建版本	校正构建版本
已处理帧数（%）	24.4（18.8–36.9）	97.5（89.9–100.0）
每个关键帧的耗时，中位数（毫秒）	192（131–215）	18.0（15.4–20.1）
每个关键帧的耗时，第95百分位数 (ms)	303（264–342）	36（32–55）
配准时间，中位数（毫秒）	135（87–151）	4.0（3.5–4.8）
姿态延迟，第95百分位（毫秒）	301（246–399）	65 (55–72)
与 ARKit 的位置差异，均方根（厘米）	23.5（14.1–45.0）	1.8（0.8–4.5）
估计路径 / ARKit 路径	5.75（3.77–13.26）	1.23（1.15–1.32）
符合规则的运行次数	15 次中 0 次	15 次中 14 次

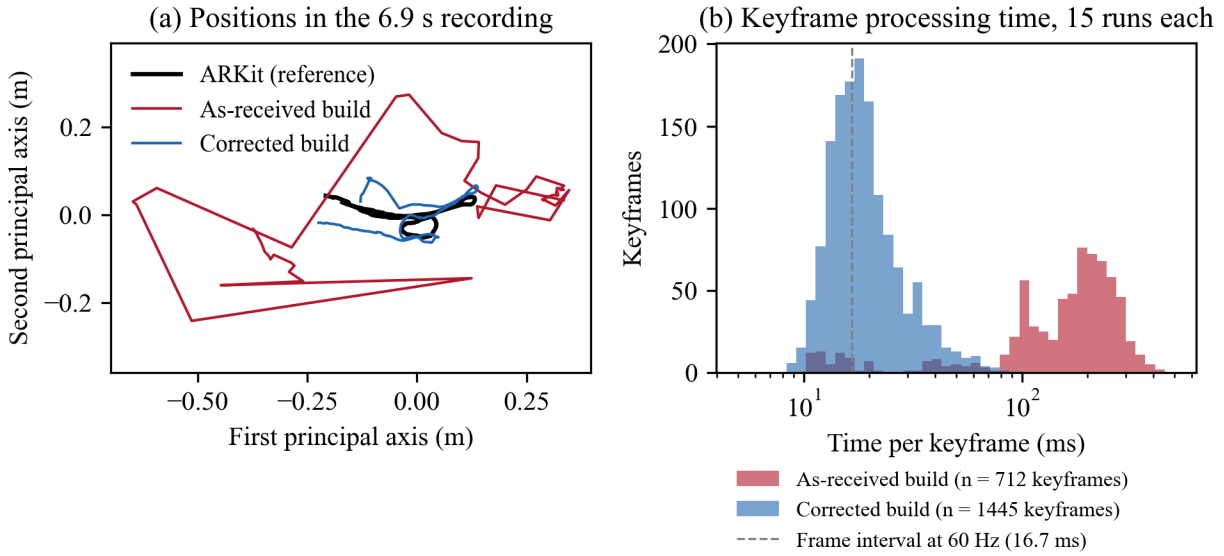

图 6. 设备端回放。（a）经过刚性对齐后，6.9 秒录制数据在 ARKit 路径的两个主轴所在平面上的位置（黑色：ARKit；红色：原始构建版本；蓝色：校正构建版本）。（b）每个关键帧所用的时间，汇总了每个构建版本 15 次运行中的所有关键帧；虚线表示两组 60 Hz 录制数据的 16.7 ms 帧间隔。此处仅对关键帧计时。表 4 给出了每次运行数值的中位数，与汇总中位数略有不同。

ARKit，因此其在手机上的大部分误差都来自丢帧。不过，其路径长度是 ARKit 路径的 3.7–7.8 倍：位姿在相邻关键帧之间发生振荡。轴映射是对正确性的修正，在此次回放中的效果不一：两段录制内容的位置差异有所下降，而另一段基本未变（分别为 3.9 和 4.1 cm），且路径比并未持续改善。先验项消除了振荡（路径比为 1.2–1.3），并使三段录制内容中的两段的迭代次数中位数降低。随后，结合停止规则的候选选择在位置差异不变的情况下，将配准时间缩短了 19 至 30 倍。

表5。对三段手机录制内容进行桌面端回放，处理每一帧。每个单元格依次列出时长为 12.7 s、6.9 s 和 8.4 s 的录制内容；各行以累积方式添加更改。

设置	与……的差异 ARKit，均方根（厘米）	估计路径 / ARKit 路径	迭代次数 （中位数）	配准时间 (ms)
接收时的源数据	1.6, 3.9, 7.8	7.8, 3.7, 5.2	20, 20, 20	100, 144, 176
+ 轴映射	1.1、4.1、2.3	4.9, 4.1, 5.2	20, 20, 20	84, 143, 169
+ 先验项	0.8, 3.2, 1.8	1.2, 1.2, 1.3	11, 8, 20	46、68、95
+ 候选项与 停止规则	0.8, 3.2, 1.8	1.2, 1.2, 1.2	2, 2, 2	2.4, 3.1, 3.2

本文未展示的内容。 这些录制内容时间较短，仅覆盖一个房间和一台设备，且唯一的参考是 ARKit。结果表明，校正构建版本在手机上回放时能够跟上传感器，并在几秒内与 ARKit 保持一致；但这些结果并未展示其在房间尺度扫描中的精度，而且校正构建版本尚未在留出的场景上进行评估。

7 与开放估计器及手机扫描结果的比较

本节报告了在七个场景中与开放的激光雷达-惯性估计器的比较、输入条件的影响、桌面处理成本、默认 ARKit 跟踪器，以及由手机应用导出的重建结果。

六种开放式估计器在 42 次运行中的 41 次里 ATE 至少为 4.46 m；Mobile-LIO 始终保持有界。表 6 给出了六种开放式激光雷达-惯性估计器以及深度惯性 Mobile-LIO 在三个留出场景和四个开发场景上的 ATE。所有估计器均获得相同的深度点和惯性样本，并根据数据集位姿进行评分。Mobile-LIO 在每个场景上的 ATE 都最低，为 0.200–1.973 m。其他结果中最优的是 0.806 m（iG-LIO 在 S1 上）；其余 41 个分数均至少为 4.46 m。该比较并非受控比较：Mobile-LIO 将速度限制在 0.5 m/s，而其他估计器没有；没有任何估计器针对手机深度进行了重新调参；iG-LIO 和 LIO-Livox 在一些无噪声适配器测试中也显示出较大的误差；此外，DLIO 在两次录制内容结束前停止产生输出。因此，该表明，针对专用扫描仪配置的估计器无法以当前配置迁移到此类输入；它并不对这些算法进行排名。图 7 展示了经过每个估计器一次刚性对齐后，留出场景上的位置误差随时间的变化。单次对齐会将后期发散分摊到整个运行过程，因此发散估计器的曲线无法显示各自何时失效；Mobile-LIO 始终将误差保持在几米以内。

在 3米限制下，视觉更新使七个场景中有四个的误差降低了 0.1 米以上。表 7 给出了三种输入条件下 Mobile-LIO 的绝对轨迹误差。比较共享深度、惯性输入和初始化的两列 3米限制数据可知，视觉更新使四个场景（S3 和三个开发场景）的误差降低了 0.1 米以上，使 S1 和 S2 的误差变化小于 0.1 米，并使一个开发场景的误差上升（从 0.943 米升至 1.656 米）。10 米条件与两种 3米限制条件的输入坐标系约定和初始化也不同，因此前两列并未单独考察范围的影响。

表 6. 在相同的手机深度与惯性输入下，六种开源激光雷达-惯性估计器和深度惯性 Mobile-LIO 相对于数据集位姿的绝对轨迹误差（m）。S1–S3 为留出场景；其余为开发场景。DLIO 的输出覆盖了 S2 和 S3 的 82% 和 83%。

场景	DLIO	FAST-LIO2	Faster-LIO	Point-LIO	iG-LIO	LIO-Livox	移动端-LIO
S1	155.8	1,151.7	158.0	9.09	0.806	11,521.5	0.409
S2	8,069.7	4,642.1	833.2	7.77	944.2	43,628.5	0.200
S3	6,810.7	82,003.7	3,068.4	100.5	6,180.8	52,938.2	1.973
09c1414f1b	2,879.2	4,907.6	30.6	5.37	5,748.1	96,550.9	1.105
0d2ee665be	12.6	129.7	208.9	4.46	375.1	10,998.6	0.871
1ada7a0617	8,988.0	719.7	1,582.7	6.14	244.7	19,797.2	0.578
25f3b7a318	5.10	3,432.1	10,880.9	5.79	3,534.8	42,545.1	0.526

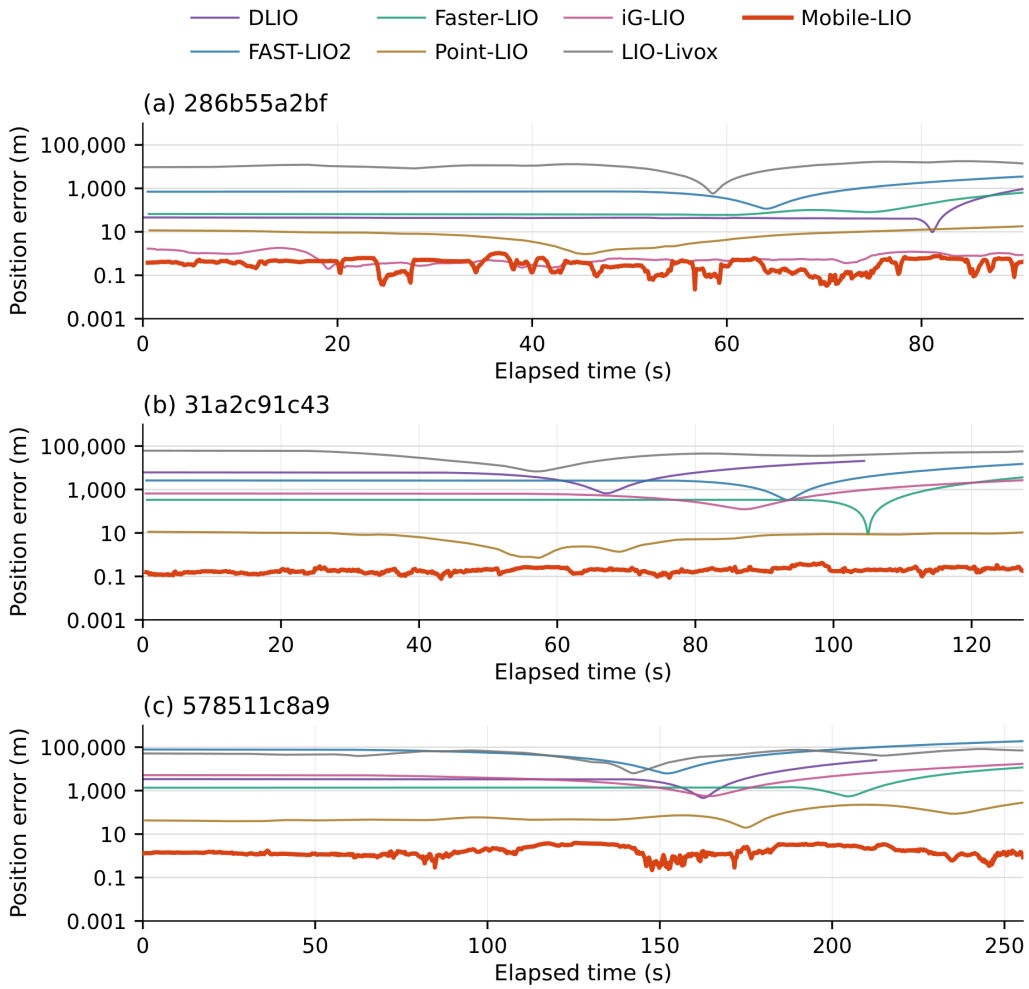

图 7. 在留出场景上，经过每个估计器一次刚性对齐后，位置误差随时间相对于数据集位姿的变化（从上到下为 S1 至 S3），采用统一的对数尺度。面板标签 286b55a2bf、31a2c91c43 和 578511c8a 分别对应 S1、S2 和 S3；theDLIO 曲线提前结束，因为其输出已停止。

表 7. 按输入条件统计的 Mobile-LIO 与数据集位姿之间的绝对轨迹误差（m）。

场景	深度和IMU，10米限制	深度和IMU，3米限制	图像、深度和IMU，3米限制
S1	0.409	0.441	0.435
S2	0.200	0.230	0.170
S3	1.973	2.920	0.590
09c1414f1b	1.105	0.413	0.190
0d2ee665be	0.871	1.185	0.207
1ada7a0617	0.578	0.943	1.656
25f3b7a318	0.526	0.704	0.245

台式机上的成本。 表8给出了在负载较低的主机上进行桌面端回放时每帧的处理成本，以及同一运行过程中的误差。在采用3米限制的手机录制内容中，深度-惯性重放在S1和S2上的每帧耗时分别为188毫秒和272毫秒，而加入视觉更新的回放分别耗时314毫秒和366毫秒，均慢于10赫兹输入。减少迭代次数可以缩短耗时，但会增大误差（S1：迭代20次时为0.435米，迭代5次时为0.734米；S2：分别为0.170米和0.338米）。在一段MID-360录制内容中，迭代3次时每次扫描耗时69毫秒，与迭代20次相比没有精度损失（分别为0.152米和0.162米）。峰值内存为204–493 MiB。这些是离线重放的全进程耗时，包括诊断信息和导出，不是上一节中的设备端计时。

ARKit，默认跟踪器。ARKit 是手机扫描应用程序默认使用的跟踪器，因此它是实践者已经拥有的基线，且未应用数据集的尺度校正。表9将其原始位姿与十个场景中的数据集位姿进行了比较。由于数据集位姿就是经过尺度和变换处理后的同一组 ARKit位姿，因此这并不是针对独立参考的精度测量，而是展示了该调整的幅度。经过一次刚性对齐后，在十个场景中，原始位姿与数据集位姿相差 3–17 厘米（均方根）（在三个留出场景中为 3–8 厘米），原始路径则短了 1–6%。ARKit位姿的基于网格的证据体现在表面结果上：在留出场景中，使用数据集位姿融合得到的地图与激光网格的中位距离为 1.8–2.3 厘米，而使用 Mobile-LIO位姿时为 8–17 厘米（表3）。这些地图展示了在应用数据集的尺度校正和网格对齐后，ARKit位姿结合该深度数据所能达到的效果。Mobile-LIO 和开放估计器仅接受了刚性对齐；未对开放估计器进行表面对比，这里也未测量未经校正的ARKit跟踪的精度。

重建结果。 图8将每个留出场景初始时段内记录的相同深度数据，分别与数据集位姿、FAST-LIO2位姿和Mobile-LIO位姿融合。使用FAST-LIO2位姿时，S2中的房间明显变形，墙壁发生了位移；使用Mobile-LIO位姿时，墙壁仍是彼此分离的表面，但S1中可见对齐误差。

手机应用导出的扫描结果。 图 9 展示了手机应用（DepthViz）于2026年10月15日采集并导出的点云，涵盖走廊环路、楼梯、房间以及带有镜子的场景。这些点云表明，该应用能够在设备上重建这些空间。结果仅供定性参考：当天安装的版本未作记录，楼梯和房间也未采集参考数据。对于走廊，推车上的Livox MID-360扫描仪记录了同一路线（图 10）；经过一次刚性对齐后，在整个范围内，手机导出结果与扫描仪地图之间的中位距离为0.77米，在墙壁带范围内则为0.26米，因此虽能恢复环路形状，但精度未达到测量级别。扫描仪地图本身是未进行回环闭合的里程计结果。

表8。Mobile-LIO 每个保存帧的桌面端回放开销、峰值内存和误差。手机录制使用3米限制，并根据数据集位姿计算绝对轨迹误差；MID-360录制（KI运行_2）使用与所提供对比轨迹的差异，每次扫描最多使用3,200个点。

录制	配置	迭代次数	CPU（毫秒）	耗时 (ms)	误差（米）	Peak 内存 (MiB)
S1	深度和 IMU	20	191.5	188.1	0.441	204
S1	图像、深度和IMU	20	310.0	313.7	0.435	215
S1	图像、深度和IMU	10	215.5	213.6	0.588	212
S1	图像、深度和IMU	5	171.7	169.4	0.734	210
S2	深度和 IMU	20	277.4	271.8	0.230	266
S2	图像、深度和IMU	20	365.0	366.4	0.170	317
S2	图像、深度和IMU	10	223.6	226.6	0.171	306
S2	图像、深度和IMU	5	154.1	152.7	0.338	294
MID-360	300 个残差	20	408.2	399.2	0.195	464
MID-360	1,500 残差	20	389.8	381.3	0.162	493
MID-360	1,500 残差	5	104.1	102.7	0.167	493
MID-360	1,500 残差	3	69.5	68.9	0.152	492

表9。在10赫兹下，经一次不含尺度的刚性对齐后，将ARKit原始位姿与数据集位姿（数据集的尺度与变换处理后的相同位姿）进行比较。路径比为ARKit原始路径长度除以数据集位姿路径长度。S1–S3为留出场景。

场景	帧数	时长（秒）	RMS 差异 (cm)	中位数 差异 (cm)	最大 差异 (cm)	路径比
S1	911	91	3.0	2.7	5.3	0.987
S2	1,275	127	3.3	2.5	13.6	0.945
S3	2,554	255	8.2	7.4	18.2	0.992
09c1414f1b	1,816	182	6.1	5.4	12.6	0.988
0d2ee665be	900	90	10.4	6.6	21.7	0.994
13c3e046d7	1,536	154	5.7	5.3	9.7	0.968
1ada7a0617	919	92	10.1	9.0	20.0	0.981
21d970d8de	1,524	152	12.8	11.7	23.4	0.982
25f3b7a318	1,757	176	9.7	9.2	19.6	0.939
27dd4da69e	4,337	434	17.3	15.5	32.6	0.977

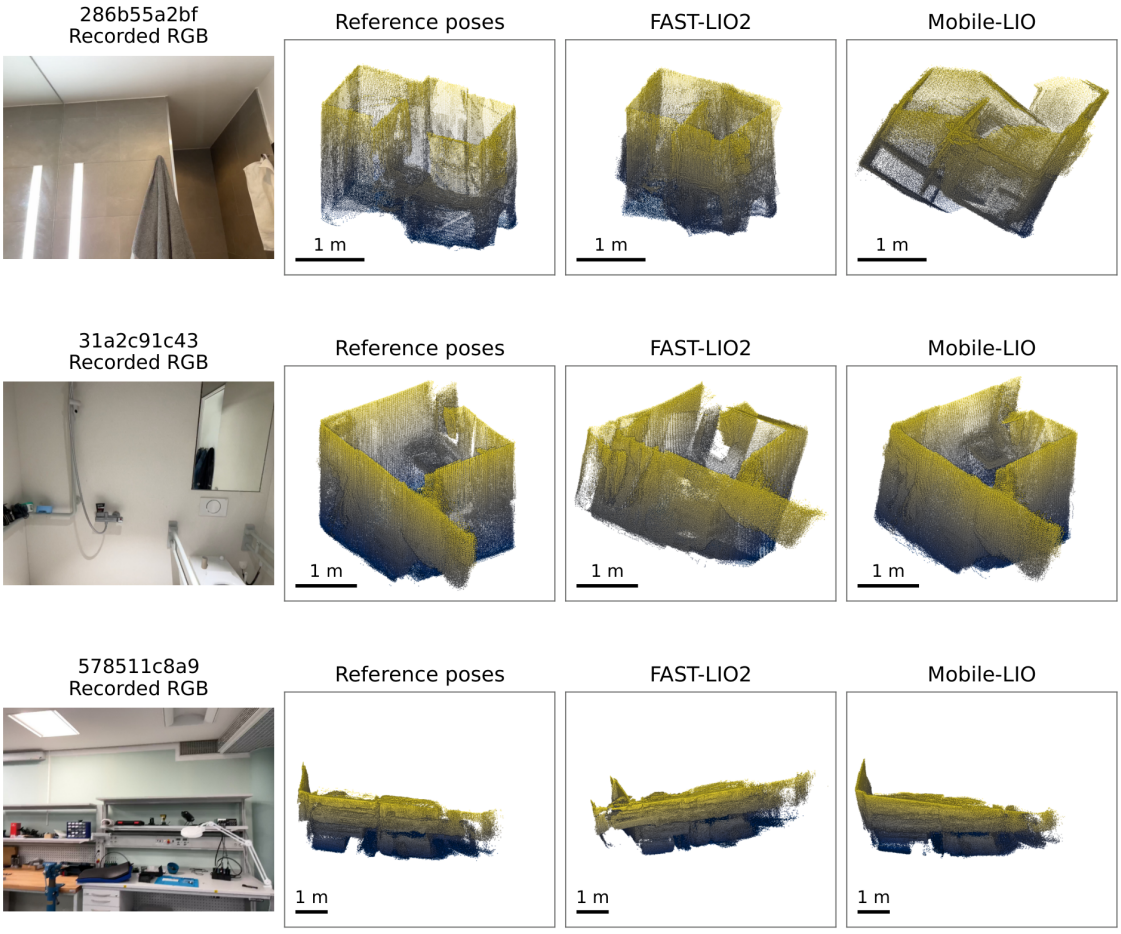

图8. 每个留出场景（从上到下为 S1 至 S3；时长分别为 35.6、71.8 和 21.8 秒）初始时间段内融合的记录图像和记录深度，以及数据集位姿（标记为参考位姿）、FAST-LIO2 位姿和 Mobile-LIO位姿。场景标签 286b55a2bf、31a2c91c43 和 578511c8a9 分别对应 S1、S2 和 S3。颜色表示距地面高度，范围为 0 米（深蓝色）至 2 米（黄色）；每一行使用统一的调色板。

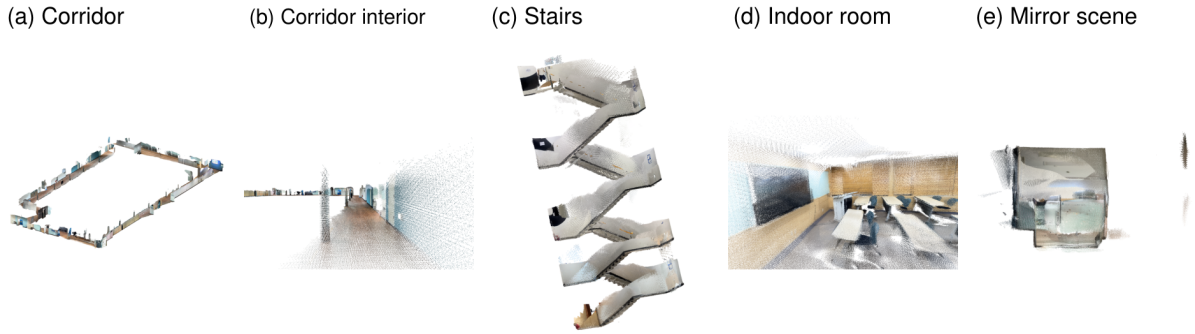

图 9. 由手机应用导出的点云：(a) 走廊环路，(b) 沿走廊方向的视图，(c) 楼梯，(d) 房间，以及 (e) 带有镜子的场景。保留原始颜色；各视图的尺度不统一。

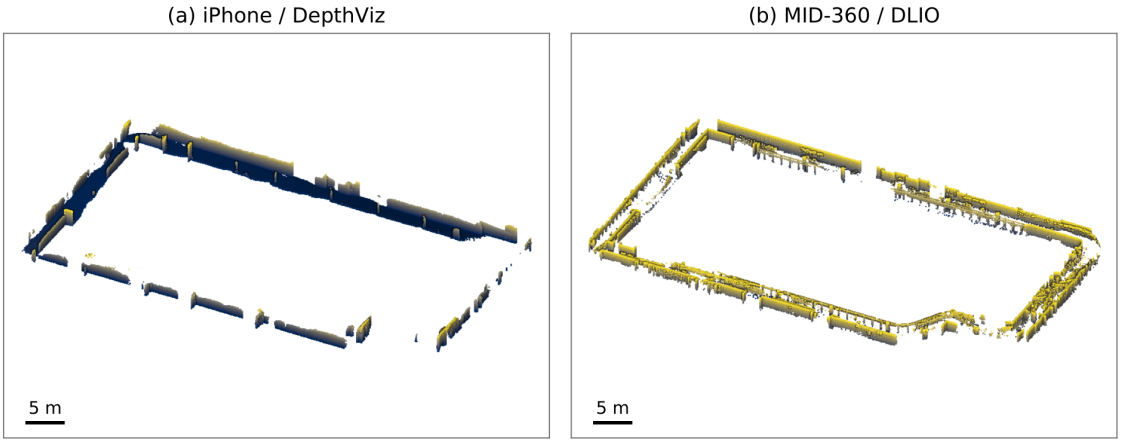

图 10. 同一条走廊的 (a) 手机导出结果（DepthViz 是该手机应用）和 (b) 由 DLIO 里程计融合的 Livox MID-360 扫描仪数据，经过一次刚性对齐后进行比较。颜色表示高度，范围为固定参考地面以下 0.3 m（深蓝色）至以上 2.2 m（黄色）；两幅图使用相同的调色板、摄像头和尺度；显示距离路线 1.5 m 以内的点。高度未沿路径压平，因此两幅图整体颜色的差异表明，对齐后两张地图之间仍存在高度偏移。

8 讨论与局限性

对于实践人员而言，这些预算给出了三点启示。第一，量程。记录深度在 3 m 以内的中位误差为 1–5 cm，在 5–7 m 范围内为 20–54 cm；这些录制内容中约 90% 的像素位于 3 m 以内。S3 是个例外，其 16.6% 的像素超过了 4 m。深度数据表明，扫描表面时距离宜控制在约 3 m 以内，但量程限制对位姿的影响并不一致，且本研究未将其作为一项规程进行测试。第二，跟踪。将 ARKit 位姿与网格对齐后，地图与网格的中位距离为 2 cm；而 Mobile-LIO 位姿使这一距离增至原来的四至八倍。因此，区分两者的是跟踪，而非深度。对齐后的位姿包含现场用户无法获得的尺度校正，因此在使用扫描结果前，应通过几处卷尺测量的尺寸或控制点进行核验。第三，适用性。ARKit 位姿地图初步对应 LOA10（离群值滤波和网格误差可能会改变超过 5 cm 的数据占比，即 17–27%），可用于空间清查、视觉环境记录和粗略的进度记录，但不适用于带尺寸标注的竣工建模或净空检查。Mobile-LIO 地图的中位误差为 8–17 cm，第90百分位数为 25–64 cm，不适用于任何带尺寸标注的用途。

对于研究而言，受控实验表明，在唯一测试场景的干净深度数据上，估计器与参考之间的偏差达到厘米级，但并未说明应优先消除哪种深度缺陷。由于没有独立的真实轨迹可用，因此本文未测量 ARKit 本身的轨迹精度；不过，通过网格可以评估其地图。

可检查性有助于诊断。日志显示，初始陀螺仪偏置状态（九个场景中为 0.019–0.136 rad/s）远高于相对于数据集位姿测得的偏置（低于 0.004 rad/s），原因是没有录制内容从静止状态开始；日志还确定了 S3 中误差增长的起始时间。

本研究存在若干局限。设备端检查仅涵盖一个房间、一部手机，以及时长为 7–13 s 的录制内容，且仅以 ARKit 为参考。导出的扫描结果仅作定性展示，而走廊对比采用会发生漂移的里程计结果作为参考。轨迹指标使用 ARKit 派生参考。深度对比会受到数据集位姿和网格中任何误差的影响。评估涵盖三个留出场景；手机配置和受控实验是在首轮留出场景结果之后进行的，属于事后实验。在七个开发场景中，速度上限导致 13–49% 的传播步骤被截断。最近表面距离不会惩罚靠近另一网格表面的重复结构；此外，地图通过最佳拟合轨迹对齐进行放置，而这种对齐在实际使用中不可用。与该领域的常见做法不同，地图未通过 ICP 与参考配准，且未测量房间尺寸。回放使用了恒定置信度，因此未测试原生数据上的置信度剔除。该估计器不具备回环闭合功能，且仅在单个室内场景上进行了评估。

9 结论

ARKit 为应用提供深度和位姿，却不提供检查其生成方式的手段。本文考察了使用不依赖 ARKit 位姿的开放式深度惯性管线进行房间尺度扫描是否可行，并以激光扫描参考为基准，测量其与封闭管线的差距。使用数据集将尺度校正并与网格对齐后的 ARKit 位姿——现场用户无法采取这一步操作——记录深度生成的地图与激光网格的中位距离约为 2 cm，但有 17–27% 的点位于 5 cm 以外。记录的 iPhone 深度在 3 m 以内的中位误差为数厘米，但存在较长误差尾部，在 5 m 以外则不可靠。按照 95% 判据，ARKit 位姿地图可初步对应 USIBD 最粗略的精度等级，而 Mobile-LIO 位姿使中位距离增至 8–17 cm。记录深度在约 3 m 以内最准确，但尚未测试限制范围是否能改善扫描；使用前应根据已知尺寸验证扫描结果。Mobile-LIO 与 ARKit 派生参考的偏差为 0.2–2.0 m；在其误差最大的场景中，无噪声深度将偏差降至厘米级。六种开放式激光雷达-惯性估计器按照专用扫描仪的配置运行，未重新调参，在该输入上的 42 次运行中有 41 次的绝对轨迹误差至少为 4.46 m（这是一次非受控比较；Mobile-LIO 设有它们没有的速度上限），而未记录构建版本的定性手机导出数据显示，设备上重建出了走廊、楼梯和房间（走廊与扫描仪地图的中位距离为 0.77 m）。

经修正手机代码的构建版本在一部手机上、无显示负载时，15 次回放运行中有 14 次能跟上传感器（每个关键帧 18 ms；两次实时扫描中的帧数占比分别为 83% 和 95%），但其房间尺度精度尚未测试，数据集结果使用的是修正前的估计器。未来工作将确定哪些深度缺陷会产生影响，评估未经数据集尺度校正的 ARKit 原始位姿，测量房间尺寸，并扩充留出集。

附录：估计器设置

表10列出了结果章节中每项结果所用估计器的设置。

表10。在数据集上评估的 Mobile-LIO 估计器设置。

Item	值
深度采样	地图中每隔 4 个像素采样一次（最多 3,072 个点）
输入剔除	置信度为 0；非有限值；范围超过 10 米（深度惯性）或 3 米（手机配置）
点云降采样	每个 3 厘米体素保留一个质心
状态	姿态方向、位置、速度、陀螺仪偏置、加速度计偏置、重力（18 个误差状态）
初始化	30 个运动传感器采样；重力大小设为测量均值；初始方差 10^{-4}，重力为 1.0
过程噪声密度	陀螺仪 0.01，加速度计 0.02（rad/s 和 m/s^2 每平方根赫兹）；偏置随机游走为 0.001 和 0.005，单位与对应项相同
速度上限	传播速度为 0.5 m/s
估计地图	10厘米网格哈希，每个网格最多20个点
平面拟合	20厘米半径内的点，至少6个，最小特征值与中间特征值之比不超过0.3
残差门限和噪声	超过10厘米则拒绝；标准差为2厘米；权重为1或按置信度取 0.5
残差预算	每次更新最多 100 个，均匀采样
重复校正	最多 20 次迭代；低于 10^{-6}时停止；无先验项；在最后一次迭代后仅更新一次协方差
更新拒绝	总校正量超过 1 m 或 0.5 rad 时恢复该条目状态
关键帧和地图	每隔 5 cm 或 2° 设置一个关键帧；5 个初始化帧；地图重置连续更新失败 10 次后
视觉更新（手机配置）	最多跟踪 200 个 Shi–Tomasi角点，采用金字塔 Lucas–Kanade光流，像素阈值为 1.5 前向–后向检查；重投影残差标准差为 2 像素；至少需要跟踪 5 个特征
外参	摄像头与运动传感器之间的恒等变换

参考文献

Abdelghany, A. A., D. D. Lichti, P. Dawson, and S. Jabari. 2026. “评估用于遗产记录的基于 iPhone 的三维扫描应用：受控实验与未来方向。”《国际摄影测量、遥感与空间信息科学档案》XLIX-M-1-2026: 1–10. Abdel-Majeed, H. M., I. F. Shaker, A. M. Abdel-Wahab, and A. A. D. I. Awad. 2024. “Apple 设备的激光雷达传感器与地面激光扫描仪的室内制图精度比较。”《*HBRC* 期刊》20 (1): 915–931. Askar, C., and H. Sternberg. 2023. “利用 iPhone 13 Pro 的智能手机激光雷达技术，以低成本进行建筑三维记录：移动扫描应用的比较分析。”《*Geomatics*》3 (4): 563–579. Bai, C., T. Xiao, Y. Chen, H. Wang, F. Zhang, and X. Gao. 2022. “Faster-LIO：采用并行稀疏增量体素的轻量级紧耦合激光雷达惯性里程计。” *IEEE Rob. Autom. Lett.* 7 (2): 4861–4868. Chen, K., R. Nemiroff, and B. T. Lopez. 2023. “直接激光雷达惯性里程计：采用连续时间运动校正的轻量级 LIO。” 收录于《*IEEE*机器人与自动化国际会议论文集》，3983–3989。纽约：IEEE。 Davletshina, D., V. K. Reja, and I. Brilakis. 2026. “道路几何数字孪生的构建与维护。”《土木工程计算杂志》40 (1): 03125002. Díaz-Vilariño, L., H. Tran, E. Frías, J. Balado, and K. Khoshelham. 2022. “使用 Apple 智能设备对室内外环境进行三维制图。”《国际摄影测量、遥感与空间信息科学档案》XLIII-B4-2022: 303–308. Ersoz, A. B., and F. Bosché. 2026. “基于移动设备与地面激光扫描的扫描到BIM工作流比较研究。” 收录于《第*43*届国际建筑自动化与机器人研讨会论文集》，2242–2249。IAARC。 He, D., W. Xu, N. Chen, F. Kong, C. Yuan, and F. Zhang. 2023. “Point-LIO：稳健的高带宽激光雷达惯性里程计。” *Adv. Intell. Syst.* 5 (7): 2200459. Jang, M., and H.-C. Yu. 2026. “Mobile-LIO：面向消费级智能手机激光雷达、具备实时校正功能的轻量级激光雷达惯性里程计。” 收录于*SIGGRAPH Asia 2026*海报，1–3。纽约：ACM。 Jin, L., and C. Ye. 2023. “视觉—激光雷达惯性里程计：一种基于 iPhone 12 Pro 的新型视觉惯性同步定位与建图方法。” 收录于《*IEEE/RSJ*智能机器人与系统国际会议论文集》，1511–1516。纽约：IEEE。 Keitaanniemi, A., P. Rönnholm, A. Kukko, and M. T. Vaaja. 2023. “ 基于室内同步定位与建图（SLAM）的激光扫描数据的漂移分析与分段后处理。”*Autom. Constr.* 147 (March): 104700. Luetzenburg, G., A. Kroon, and A. A. Bjørk. 2021. “Apple iPhone 12 Pro 激光雷达在地球科学中的应用评估。” *Sci. Rep.* 11 (1): 22221. Osadcha, I., A. Jurelionis, and P. Fokaides. 2023. “建成资产数字孪生中的几何参数更新：系统性文献综述。” *J. Build. Eng.* 73 (August): 106704. Teo, T.-A., and C.-C. Yang. 2023. “评估 iPad Pro 内置激光雷达用于三维室内制图的精度与质量。” *Dev. BuiltEnviron.* 14 (April): 100169. Teppati Losè, L., A. Spreafico, F. Chiabrando, and F. Giulio Tonolo. 2022. “Apple 激光雷达传感器用于三维测量：文化遗产领域中的测试与结果。”《遥感》14 (17): 4157. Treccani, D., A. Adami, and L. Fregonese. 2024. “评估基于激光雷达的应用在使用 Apple 设备测量中等规模室内外区域时的有效性。”《国际摄影测量、遥感与空间信息科学档案》XLVIII-2/W8-2024: 431–438. Trzeciak, M., K. Pluta, Y. Fathy, L. Alcalde, S. Chee, A. Bromley, I. Brilakis, and P. Alliez. 2023. “Con-

SLAM：用于同步定位与建图的数据集。”《土木工程计算杂志》37 (3): 04023009. USIBD（美国建筑文档研究所）。2016年。精度等级（*LOA*）规范指南：*C120*号文件，*2.0*版。加利福尼亚州塔斯廷：USIBD。Xu, W., Y. Cai, D. He, J. Lin, and F. Zhang. 2022. “FAST-LIO2：快速直接激光雷达惯性里程计。”《*IEEE*机器人学汇刊》38 (4): 2053–2073. Yang, Y., Y.-T. Chen, C. Hancock, N. A. S. Hamm, and Z. Zhang. 2024. “一种利用惯性测量单元和移动式激光扫描仪进行竣工BIM更新的新方法。”《遥感》16 (15): 2743. Yeshwanth, C., Y.-C. Liu, M. Nießner, and A. Dai. 2023. “ScanNet++：高保真三维室内场景数据集。”载于论文集，*IEEE/CVF*国际计算机视觉会议，12–22。纽约：IEEE。