GOLD-SLAM：
高斯泼溅全向激光雷达深度同步定位与建图

Weicong Zhang[1,2] **Yuan Huang**[1] **Yu Zheng**[1] **Lin Chenxiao**[1] **Yingying Li**[2] **Lilong Fan**[3] **Liangdong Zhuang**[2,†] **YufanChen**[4,5] **Ruiping Liu**[4] **Junwei Zheng**[4] **Jiaming Zhang**[1,†]

[1] 湖南大学 [2]清华大学[3]中国铁建大桥工程局集团有限公司[4] 卡尔斯鲁厄理工学院 [5]牛津大学

† 通讯作者

arXiv:2611.32335v2 [cs.CV] 25 Nov 2026

摘要

全景相机能够在单帧中捕获几乎整个球形视场。然而，这种全景覆盖范围与激光雷达视场之间的不匹配，为依赖激光雷达深度先验的全景高斯泼溅同步定位与建图带来了重大挑战。我们提出了 **GOLD-SLAM**，一种高斯泼溅全向激光雷达深度同步定位与建图系统，将高质量渲染与准确的几何重建相结合。该系统在球面域中执行跟踪与建图，并利用无激光雷达覆盖区域内跟踪到的视觉特征来提升位姿估计。为解决增量建图过程中新观测表面覆盖不足和漂浮伪影问题，我们提出了深度感知地图扩展与剪枝（DAEP）。为实现更广泛的评估，我们使用无人机平台采集了一个新数据集 **WHOLE**，涵盖弱几何、高动态、全向、低照度和简单场景。在公开基准和 WHOLE 上进行的大量实验表明，与透视系统相比，该系统具备具有竞争力的渲染质量和更高的几何精度。与现有全景基线相比，GOLD-SLAM 将平均峰值信噪比提升了 5.56–9.10 分贝，并将平均深度 L1 误差降低了 71.9–83.9%。为评估其在无人机等资源受限平台上的部署可行性，我们在 Jetson Orin NX（16 GB）上评估了 GOLD-SLAM 的嵌入式配置，结果表明该系统能够在线运行，但需要以牺牲重建质量为代价。

GitHub https://github.com/InSAI-Lab/GOLD-SLAM
Hugging Face https://huggingface.co/datasets/InSAI-Lab/WHOLE
项目页面 https://insailab.org/GOLD-SLAM
InSAI实验室 https://insailab.org/

日期 2026年11月

1 引言

地面与空中机器人需要可靠的定位、稠密几何和照片级真实感渲染来实现交互，且往往受限于传感器预算 [20, 23]。激光雷达提供具有度量尺度的稀疏几何测量，惯性测量单元提供高频惯性约束，而相机则捕捉丰富的颜色与纹理。

它们各自互补的作用构成了多传感器里程计与建图系统的基础 [15, 21, 33–35]。三维高斯泼溅（3DGS） [11] 使用显式的各向异性高斯表示场景。其高效的光栅化使渲染速度快于 NeRF [19], , 推动了基于高斯的同步定位与建图 [9, 12, 13, 18, 22]的发展。

然而，透视相机的视场（FoV）较窄，限制了场景覆盖范围和帧间对应关系，使位姿估计更容易受到重复或稀疏纹理以及快速运动的影响 [7]。全景相机将覆盖范围扩展至 $360° \times 180°$ 球形视场，但其投影会引入更强的非线性畸变 [8]。ODGS-SLAM [22] 将增量式高斯SLAM扩展至等距柱状投影（ERP）图像。然而，该方法在室外会出现尺度漂移，因此有必要利用激光雷达几何约束进行大规模建图。

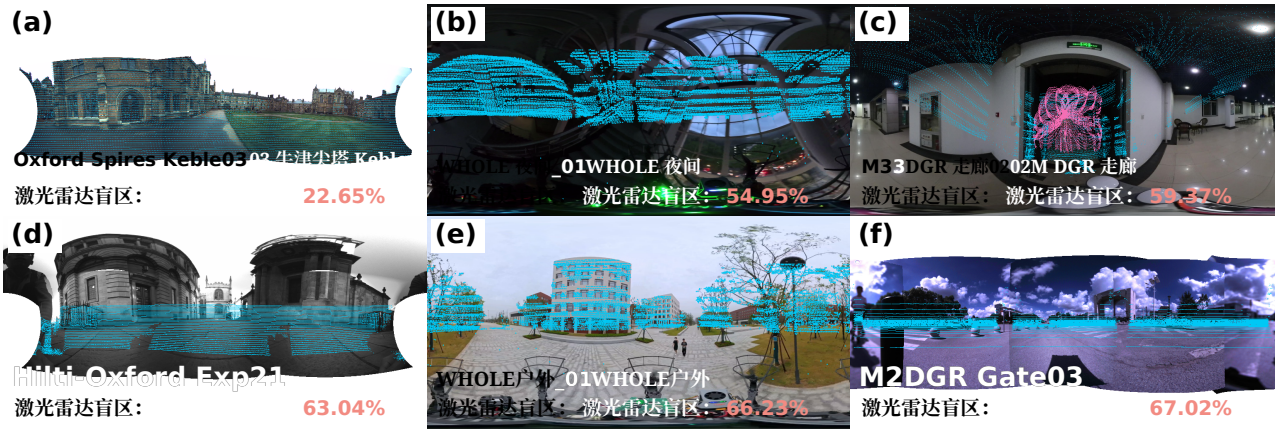


图1. 这 些有限的激光雷达覆盖范围（彩色点）在360°全景图像中会形成大面积激光雷达盲区 ne（百分比 按立体角加权）以及多模态融合方面的挑战。

全景激光雷达–惯性–视觉高斯SLAM面临若干技术挑战：

1. **ERP采样不均匀且存在极区畸变**。 ERP投影会导致球面上的采样不均匀，并使极区附近产生显著拉伸 [17, 22]。2. **激光雷达盲区的影响更大**。 激光雷达–相机视场不匹配会导致某些区域缺乏深度支持 [2, 13]。全景覆盖会暴露出更多此类区域，使高斯初始化变得复杂（图 1）。3. **干扰新表面重建**。在增量式高斯 SLAM 中，已有高斯的高累积不透明度可能会阻碍在新观测到的邻近表面上进行仅基于不透明度的插入 [10, 27]。

除了这些技术挑战之外，常用的激光雷达-视觉基准主要关注地面机器人或手持式/背包式采集 [24, 28, 30, 31], , 对无人机运动的覆盖不足。在此背景下，我们的贡献如下：

• 我们提出了 GOLD-SLAM，这是一种高斯泼溅同步定位与建图框架，将激光雷达、惯性测量单元和全景相机与共享球面模型相融合，用于跟踪和建图。
• 我们结合盲区视觉约束与图像块亮度不确定性（PBU），以提升在激光雷达覆盖范围有限及光照变化条件下的位姿估计性能。
• 我们将实测激光雷达深度与全景深度预测相结合，用于高斯初始化，并设计了深度感知地图扩展与剪枝（DAEP），以提升增量式表面重建效果。
• 我们采集了 WHOLE 数据集，其中包含使用定制无人机平台在室内、室外和退化环境中获取的激光雷达与 IMU 测量值以及全景图像。在 WHOLE 和公开基准上的结果表明，该方法具有具有竞争力的定位性能，并在广泛场景覆盖范围下实现了更优的几何重建。

2 相关工作

高斯泼溅SLAM。3DGS [11] 通过可微高斯光栅化实现高效的照片级真实感渲染。MonoGS [18] 采用可微泼溅进行单目SLAM，并使用解析位姿雅可比矩阵和可选的深度监督。Photo-SLAM [9] 将几何特征与光度特征相结合，并通过渐进式高斯金字塔学习和几何致密化支持单目、立体和RGB-D SLAM。SplaTAM [10] 和GS-SLAM [27] 采用深度引导的高斯插入，GS-SLAM还会抑制漂浮伪影。与这些透视系统不同，GOLD-SLAM利用激光雷达几何监督和预测深度进行全景重建，并通过盲区高斯初始化处理盲区。

全向SLAM。全向SLAM通过多相机系统或鱼眼相机扩大视觉覆盖范围。高斯投影和渲染必须处理ERP图像中的极向拉伸和水平周期性。360ORB-SLAM [3] 结合全景特征、三角测量和深度补全，实现定位和点云建图。全景激光雷达DSO [7] 将全景图像与激光雷达

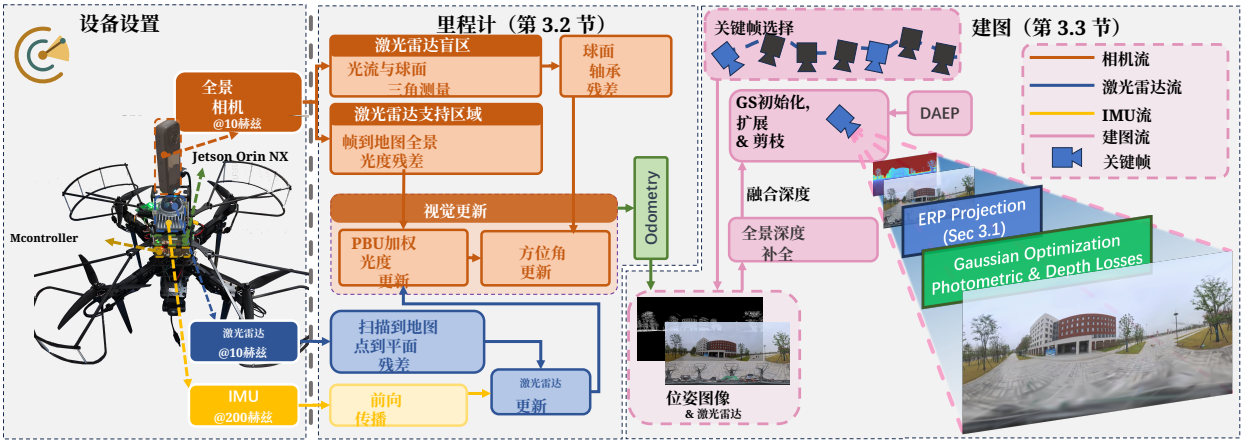


图 2。 GOLD-SLAM 概览。无人机平台采集激光雷达、IMU 测量值和全景图像。受 FAST-LIVO2 误差状态滤波框架的启发 [34],前端将全景视觉更新与 IMU传播和激光雷达扫描到地图约束相结合。帧间约束利用激光雷达盲区中的特征，而图像块亮度不确定性（PBU）则对激光雷达支持区域中的共享亮度变化进行建模。为支持全景重建，我们将 Gaussian-LIC2 后端 [13] 适配到球面投影和渲染。高斯利用测得的激光雷达深度和预测的全景深度进行初始化，并通过光度损失和激光雷达深度损失进行优化。深度感知地图扩展与剪枝（DAEP）结合深度感知插入、深度引导的自适应更新和深度冲突剪枝，从而改善新观测表面的重建效果。

观测关联，以实现稳健跟踪。ODGS [14] 根据ERP图像重建高斯场景，而ODGS- SLAM [22] 则将增量式高斯SLAM扩展至ERP输入。

多模态融合。 在大规模场景中，多模态融合能够在视觉退化或激光雷达退化的情况下支持定位和重建。LIV-GaussMap [6] 利用激光雷达–IMU位姿和初始几何，然后通过光度信息细化高斯。Gaussian-LIC [12] 将紧耦合的激光雷达–IMU–相机定位与基于激光雷达点和盲区视觉三角测量的高斯初始化相结合。GS-LIVM [25]通过高斯过程回归补全激光雷达表面，并利用预测协方差进行高斯初始化。Gaussian-LIC2 [13] 将用于盲区高斯初始化的轻量级深度补全与激光雷达深度监督相结合。在此基础上，GOLD-SLAM 支持全景重建，并利用盲区视觉约束和 PBU 进行位姿估计。

3 方法

GOLD-SLAM 利用共享的球面几何融合激光雷达、惯性测量单元和全景视觉，以进行位姿估计，并利用激光雷达深度引导高斯建图（图 2）。

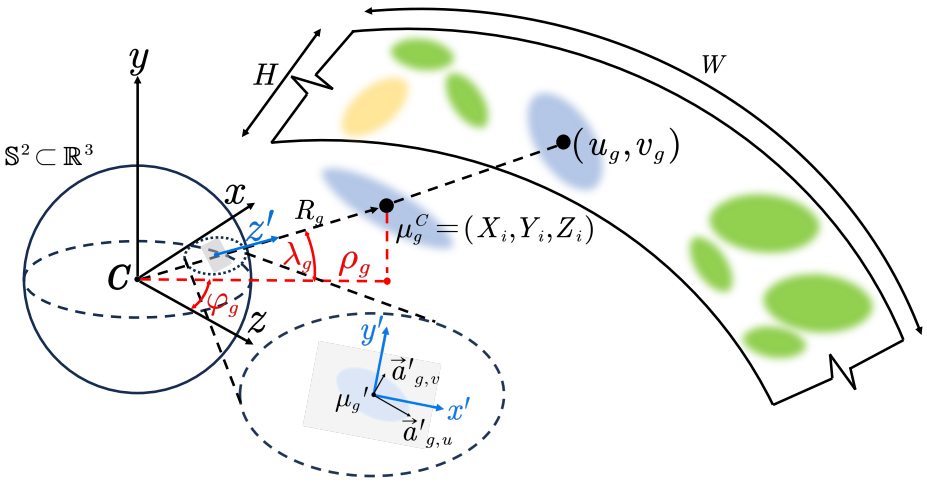


图 3. 统一球面投影几何。

3.1 统一球面几何

球面投影几何如图 3 所示。上标 W、 I 和 C 分别表示世界坐标系、惯性测量单元坐标系和相机坐标系，R_A^B 表示从坐标系 A 到坐标系 B 的旋转。对于点 $\boldsymbol{\mu}^C = (x, y, z)^\mathsf{T}$，我们定义其径向距离 $R = \|\boldsymbol{\mu}^C\|_2$、经度 $\varphi = \mathrm{atan2}(x, z)$ 和纬度 $\lambda = \mathrm{atan2}(y, \rho)$，其中 $\rho = \sqrt{x^2 + z^2}$。对于 $W \times H$ 等距柱状投影图像，投影和反投影为

$$\pi(\boldsymbol{\mu}^C) = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \frac{W}{2\pi}(\varphi + \pi) \\ H(\frac{1}{2} - \frac{\lambda}{\pi}) \end{bmatrix}, \quad \boldsymbol{\mu}^C = R\mathbf{d}(u, v), \tag{1}$$

其中，$\mathbf{d}(u, v) = [\cos\lambda \sin\varphi, \sin\lambda, \cos\lambda\cos\varphi]^\mathsf{T}$ 是相机坐标系中的单位视线向量。对于反向投影，$\varphi = 2\pi u/W - \pi$ 和 $\lambda = \pi(1/2 - v/H)$。

式（1）的**雅可比矩阵** $_\pi = \partial\pi/\partial\boldsymbol{\mu}^C$ [22]用于全景位姿更新和高斯协方差投影。坐标变换和扰动约定见附录 B。

对于高斯渲染，地图 $\mathcal{G}$ 包含 N_{GS} 个图元，每个图元具有一个世界坐标系中心$\boldsymbol{\mu}_g$、协方差 $\boldsymbol{\Sigma}_g$、不透明度o_g 以及球谐函数（SH）系数$\mathbf{f}_g$。在局部切平面近似下 [14, 22],，相机坐标系协方差 $\boldsymbol{\Sigma}_g^C = \mathbf{R}_W^C \boldsymbol{\Sigma}_g (\mathbf{R}_W^C)^\mathsf{T}$ 被投影到 ERP 图像空间，得到

$$\boldsymbol{\Sigma}_g^{\mathrm{img}} = \mathbf{J}_{\pi,g} \boldsymbol{\Sigma}_g^C \mathbf{J}_{\pi,g}^\mathsf{T} + \epsilon\mathbf{I}, \tag{2}$$

其中，$\mathbf{J}_{\pi,g}$ 在高斯中心处求值，低通项 $\epsilon > 0$ 可防止投影结果过小 [11]。

3.2 全景激光雷达-惯性-视觉里程计

激光雷达前端使用基于体素的平面地图 [29], 并在线选择 0.5、1 或 2 米的基础体素尺寸（附录 C）。

激光雷达盲区中的视觉约束。 我们在激光雷达盲区中提取角点，并使用金字塔光流在各帧间跟踪这些角点。前后向光流一致性用于剔除错误的对应关系。当视差足够大时，我们根据历史视线射线对每个特征进行三角测量 i （附录 D）。将三角测量点 $\mathbf{p}_i^W$变换到当前相机坐标系 k 后，得到 $\mathbf{p}_{ik}^C$。预测和测量的视线方向分别为 $\widehat{\mathbf{d}}_{ik} = \mathbf{p}_{ik}^C / \|\mathbf{p}_{ik}^C\|_2$ 和 $\mathbf{d}_{ik} = \mathbf{d}(u_{ik}, v_{ik})$。我们将视线误差投影到 $\mathbf{d}_{ik}$处的切平面上：

$$\mathbf{r}_{ik} = \mathbf{B}_{ik}(\widehat{\mathbf{d}}_{ik} - \mathbf{d}_{ik}), \tag{3}$$

其中，$\mathbf{B}_{ik}$ 的正交行张成该平面。

不确定性加权融合。 我们根据视差和历史位姿不确定性估计每个三角测量点的协方差，并保留各视图间共享的位姿不确定性分量。将该协方差通过式（3）传播，并加入角度观测噪声后，可得到视线残差协方差 $\boldsymbol{\Sigma}_{r,i} \in \mathbb{R}^{2\times2}$。这些残差构成信息矩阵 $\boldsymbol{\Lambda}_V = \sum_i \mathbf{H}_i^\mathsf{T} \boldsymbol{\Sigma}_{r,i}^{-1} \mathbf{H}_i$，其中 $\mathbf{H}_i$ 是残差 $\mathbf{r}_{ik}$相对于误差状态的雅可比矩阵。设 $\mathbf{P}_k = \mathbf{L}\mathbf{L}^\mathsf{T}$ 为激光雷达和直接视觉更新后的状态协方差，其中 $\mathbf{L}$ 为其乔列斯基因子。为限制相对视觉信息，我们对 $\boldsymbol{\Lambda}_V$及相应的加权残差项同时进行缩放，缩放因子为 η:

$$\eta = \min\left(1.4, \frac{\gamma}{\lambda_{\max}(\mathbf{L}^\mathsf{T}\boldsymbol{\Lambda}_V\mathbf{L})}\right), \tag{4}$$

其中， $\lambda_{\max}$ 表示最大特征值， γ 则限定相对视觉信息。

图像块亮度不确定性（PBU）。 在激光雷达支持区域内，图像块范围内的亮度变化可能会在直接视觉更新过程中被误认为是运动。通过采用仿射亮度模型 [4], ，我们将这些变化视为图像块内共享的独立零均值增益 α 和偏移量 b 扰动。残差协方差（附录 E）为

$$\mathbf{C}_{\mathrm{photo}} = \sigma_n^2\mathbf{I} + \sigma_\alpha^2 \mathbf{i}_{\mathrm{ref}}\mathbf{i}_{\mathrm{ref}}^\mathsf{T} + \sigma_b^2 \mathbf{1}\mathbf{1}^\mathsf{T}, \tag{5}$$

其中，$\mathbf{i}_{\mathrm{ref}}$ 包含参考图像块强度，$\mathbf{1}$ 为全一向量， σ_n、 σ_α和 σ_b 分别为独立像素噪声、增益扰动和偏移扰动的标准差。我们使用该协方差对白化残差和雅可比矩阵，以在视觉更新期间考虑共享亮度不确定性。

3.3 高斯建图

高斯渲染。 利用第 3.1 节中的投影 $R_g = \|\boldsymbol{\mu}_g^C\|_2$，我们按照径向深度对高斯进行排序，并通过从前到后的阿尔法合成渲染全景颜色和深度。在像素 $\mathbf{u} = (u,\ v)^\mathsf{T}$处，令 a_g 为按不透明度加权的投影高斯响应，并令 $T_g = \prod_{h<g}(1 - a_h)$ 为透射率。DAEP 使用的累积不透明度和渲染径向深度为

$$A = \sum_g T_g a_g, \qquad \widehat{D} = \frac{\sum_g T_g a_g R_g}{A}, \quad A > 0. \tag{6}$$

初始化。 对于每个建图关键帧，Depth Any Panoramas (DAP) [16]（附录 F）会预测径向深度 D_P。我们通过有效重叠区域内的深度比中位数 s，将其与激光雷达深度 D_L 对齐。融合深度 D_F 为

$$D_F(\mathbf{u}) = \begin{cases} D_L(\mathbf{u}), & D_L(\mathbf{u}) > 0, \\ sD_P(\mathbf{u}), & D_L(\mathbf{u}) = 0,\ D_P(\mathbf{u}) > 0, \\ 0, & \text{otherwise.} \end{cases} \tag{7}$$

采样的激光雷达点直接用于初始化高斯。对于激光雷达盲区内的每个图像块，我们应用深度连续性筛选 D_F，并通过公式（1）将具有最小有效深度的像素反投影，以初始化一个高斯。

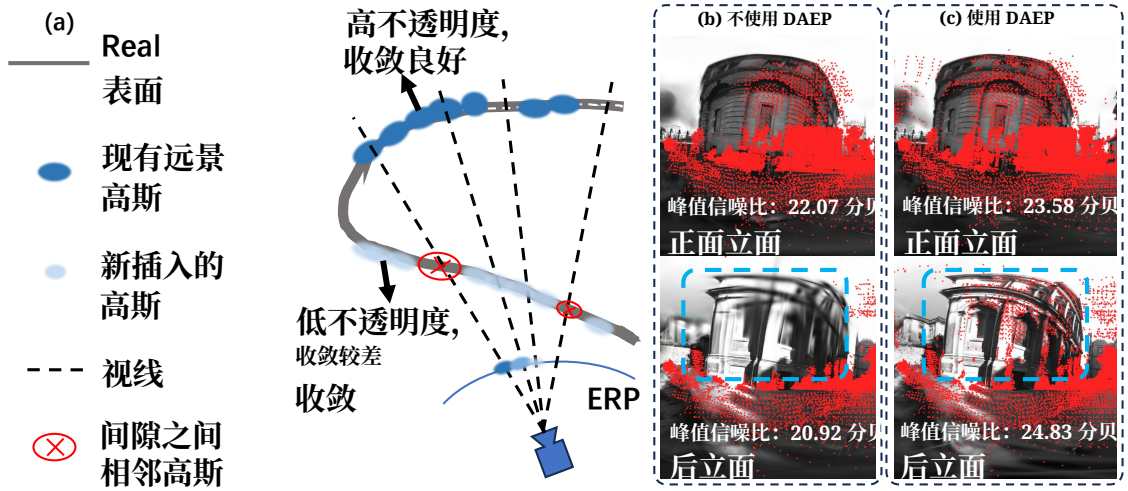


图4。 深度感知地图扩展与剪枝（DAEP）。(a) 仅基于不透明度的插入可能会遗漏新表面。(b、c) 未使用和使用 DAEP 时的渲染结果。红点标记了在五个相邻关键帧中添加的高斯的初始中心。蓝色虚线框标出后立面。

深度感知地图扩展与剪枝（DAEP）。 PanoGS-SLAM 等方法 [17] 会在累积不透明度较低的区域提高高斯插入密度。然而，现有高斯的高累积不透明度可能会妨碍在新观测到的邻近表面上插入高斯（图 4）。新高斯的不透明度较低且覆盖不完整，也可能导致远处表面仍然可见。为解决这些问题，DAEP 将深度感知插入 [10] [27], 与深度冲突剪枝相结合。在像素 $\mathbf{u}$ 处，若深度为 d，则插入有效的初始化候选项

$$A(\mathbf{u}) < \tau_\alpha \quad \text{or} \quad \widehat{D}(\mathbf{u}) - d > \max(\delta_0, \beta d). \tag{8}$$

第一个条件用于填补未覆盖区域，而第二个条件则在渲染表面位于更近观测之后时，即使累积不透明度较高，也允许纳入这些观测。此处，τ_α 是不透明度阈值，而 δ_0 和β 分别控制绝对和相对深度容差。对于深度冲突插入，我们采用更高的初始不透明度和尺度，并根据深度差异自适应地调整 Adam 更新的权重：

$$\Delta\boldsymbol{\theta}_g = w_g \Delta\boldsymbol{\theta}_g^{\text{Adam}}, \qquad w_g \geq 1, \tag{9}$$

其中，$\Delta\boldsymbol{\theta}_g^{\text{Adam}}$是高斯 g的标准 Adam 更新，w_g 是深度引导的自适应权重。

为去除漂浮伪影，每轮建图都会根据当前关键帧检查高斯深度一致性。对于不透明度超过阈值，且中心位于观测表面前方、距离超过深度容差的高斯，若局部深度连续性确认存在冲突，则将其剔除。该检查优先采用激光雷达深度，并在激光雷达盲区使用 DAP 估计值，同时采用更大的冲突容差。随后，在保持前端位姿固定的情况下优化高斯参数：

$$\mathcal{L}_{\text{map}} = (1-\lambda_s)\mathcal{L}_{\text{RGB}} + \lambda_s(1-\text{SSIM}) + \lambda_d\mathcal{L}_{\text{depth}}. \tag{10}$$

其中， $\mathcal{L}_{\text{RGB}}$ 是所选优化视图上的 RGB 平均绝对误差。 $\mathcal{L}_{\text{depth}}$ 是有效激光雷达像素处的平均绝对径向深度误差， λ_s 和 λ_d 是权重。通过多视图光度监督，对激光雷达盲区内的几何进行细化。

4 实验

4.1 TheWHOLE 数据集

我们构建了**WHOLE** 数据集，其中包含无人机在具有挑战性的几何与光照条件下运动的数据，涵盖弱几何、高动态、全向、低照度和简单场景。数据集包含 20 个序列，记录了校园室内和室外环境中的全景图像、激光雷达扫描和 IMU 测量值。场景包括房间、长廊、车库、玻璃幕墙和开放空间（图 5）。录制涵盖白天和夜间条件、低照度车库，以及伴有图像模糊的快速运动。一些序列将这些视觉挑战与弱几何相结合，可用于评估视觉与几何同时退化时的定位和重建性能。数据由定制的 FanciSwarm Pro [1] 无人机平台（图2）采集。该平台搭载 MID360S 激光雷达（内置 IMU）和 Insta360 X5 双鱼眼全景相机。我们使用 Kalibr [5] 和 AprilGrid 标定两台鱼眼相机的内参，以及相机与 IMU 的时空参数。使用 FAST-Calib 标定激光雷达与相机外参2 [32]。AprilGrid 观测结果提供序列端点处的参考位姿。



(a) 室内_01 (b) 室外_01 (c) 室外_03

(d) 夜间 02 (e) 夜间 03 (f) 车库 02

图 e 5. **WHOLE 场景可视化**：（a）室内，（b、c）室外，（d、e）夜间，（f）车库。

4.2 实验设置

数据集。 在这项工作中，我们在五个数据集上开展实验。表 1 总结了现有四个数据集与 WHOLE 之间的统计信息和比较结果。我们从 Hilti-Oxford 2022 数据集（**Hilti-Oxford**） [31], 中选取了涵盖施工现场（实验04）、地下室（实验14）、走廊和展廊（Exp18）以及建筑外部（实验21）的序列。 **M3DGR** [30] 提供了不同光照条件下的序列（Varying-illu05）、长廊序列（Corridor01/02）以及室外草地区域序列（Grass01/02）。在 M3DGR 上，GOLD-SLAM 和 Gaussian- LIC2 仅使用 MID-360 作为其激光雷达输入。 **M2DGR** [28] 涵盖室内房间和大厅（room_01 和

大厅_02）、室内外过渡区域（门_02）以及室外门区（gate_03）。我们从牛津尖塔数据集（**OxfordSpires**）[24], 中选取了 Keble03、Observatory01、ChristChurch03 和 Blenheim01，涵盖建筑内部和室外空间。

表 1. 公开 SLAM 数据集与我们提出的 WHOLE 数据集的比较。W：弱几何；H：高动态性；O：全向覆盖；L：低照度；E：简单场景。

数据集	年份	序列数	平台	相机			IMU (Hz)	激光雷达	数据多样性	
				Type	分辨率	FPS			整段	
M2DGR[28]	2021	36	地面机器人	7× 鱼眼（RGB） 透视 RGB-D	1280×1024 640×480	15 15	150	VLP-32C，10 赫兹	☑	☑ ☑ ☑ ☑
Hilti-牛津 [31]	2022	16	手持式	5× 鱼眼（灰度）	720×540	40	400	PandarXT-32, 10 Hz	☑	— ☑ ☑ ☑
牛津 Spires [24]	2024	24	背包	3× 鱼眼（RGB）	1440×1080	20	400	禾赛 QT64，10 Hz	☑	— ☑ ☑ ☑
M3DGR[30]	2025	31	地面机器人	全景 X4（RGB） 透视 RGB-D	2880×1440（等距柱状投影） 640×480	15 30	200	Avia，10 赫兹；MID-360，10 赫兹	☑	— ☑ ☑ ☑
WHOLE（我们的方法）	2026	20	手持式；UAV	Panoramic X5（RGB）	（鱼眼）	30	2	MID360S，10 赫兹		

基线方法。 我们评估 FAST-LIO2 [26], FAST-LIVO2 [34], 和 Ground-Fusion++ [30] 在定位任务中的表现，以及 Gaussian-LIC2 [13], ODGS-SLAM [22], 和 PanoGS-SLAM [17] 在定位和重建任务中的表现。视觉输入根据各方法的相机模型和支持的视图配置进行选择（表 2）。Ground-Fusion++ （LIC）使用激光雷达、惯性测量单元和前视图像，并禁用全球导航卫星系统和轮式里程计。为确保深度输入一致，ODGS-SLAM 和 PanoGS-SLAM 使用 GOLD-SLAM 的融合深度 D_F。

表2。 基线输入和评估任务。 L：定位任务；M：建图任务。

方法	视觉输入	L	M	激光雷达	IMU	相机
FAST-LIO2	–	✓	✗	✓	✓	✗
FAST-LIVO2	透视，97°	✓	✗	✓	✓	✓
地面融合++	透视，97°	✓	✗	✓	✓	✓
Gaussian-LIC2	透视，97°	✓	✓	✓	✓	✓
ODGS-SLAM	ERP, 360°	✓	✓	✗	✗	✓
PanoGS-SLAM	ERP, 360°	✓	✓	✗	✗	✓
GOLD-SLAM	ERP, 360°	✓	✓	✓	✓	✓

实现细节。 重建输入以 10 赫兹采样： 1024×512 ERP图像或具有 97° 水平视场角的 512×384透视图像。GOLD-SLAM每 5 帧选取一个建图关键帧，每轮最多优化 70 个视图（附录 H 和 I）。高斯颜色采用三阶球谐函数；损失权重为 $\lambda_s = 0.2$ 和 $\lambda_d = 0.015$。地图扩展参数为 $\tau_\alpha = 0.99$、 $\delta_0 = 0.25$ 米和 $\beta = 0.1$。我们为盲区视觉约束设置$\gamma = 0.28$ ，并为 PBU 设置 $\sigma_\alpha = 0.05$、 $\sigma_b = 2$ 。实验使用配备 64 GB 内存的 Intel Core i7-14700KF 和配备 24 GB 显存的 NVIDIA RTX 5090 D v2 运行。

4.3 定位精度

我们首先使用ATE均方根误差和平移端点漂移评估定位精度。在这些公共数据集上，GOLD-SLAM在Hilti-Oxford和M3DGR上的ATE均值最低，在牛津尖塔数据集上排名第二，并在M2DGR上取得了相当的准确率（表3）。

长廊中的重复几何结构和外观给激光雷达配准与视觉跟踪都带来了挑战。在表 4中，FAST-LIO2 在两条走廊中都累积了较大的漂移，而 FAST-LIVO2 在 Corridor01 上失效。在 Corridor02 上，GOLD-SLAM 将端点漂移从 FAST-LIVO2 的 1.319米降至 0.274米。这种改

表3. 定位精度 (ATE均方根误差，单位为米，↓）。数值为均值（样本标准差）。最优和次优结果分别以粗体（较深底纹）和下划线（较浅底纹）标示。

方法	Hilti-牛津	M3DGR	M2DGR	牛津 Spires
FAST-LIO2	1.865 (3.397)	0.351 (0.227)	0.358 (0.130)	**0.075** (0.047)
FAST-LIVO2	0.064 (0.033)	0.346 (0.238)	0.360 (0.119)	0.100 (0.057)
Gaussian-LIC2	0.196 (0.132)	0.379 (0.183)	0.355 (0.135)	0.563 (0.221)
地面融合++ (LIC)	7.693 (15.070)	0.380 (0.199)	**0.349** (0.137)	0.337 (0.453)
ODGS-SLAM	9.165 (4.223)	3.231 (1.881)	2.316 (2.189)	5.360 (0.489)
PanoGS-SLAM	9.167 (4.356)	29.316 (29.053)	12.627 (17.066)	32.845 (16.454)
GOLD-SLAM（我们的方法）	**0.059** (0.028)	**0.344** (0.241)	0.360 (0.119)	0.077 (0.057)

进的漂移控制可能部分源于全景观测提供的额外视觉约束，这些观测涵盖前向视野之外的区域，包括激光雷达盲区。

WHOLE 将评估扩展到弱几何和快速运动场景。GOLD-SLAM 在 W 组和 H 组中取得了最低的平均端点漂移，相较于 FAST-LIVO2 分别降低了 34% 和 43%（表 4）。广泛的视觉覆盖范围可在观察方向变化时保持重叠，而 PBU 则处理视觉更新中的共享亮度变化。Exp18 消融实验进一步验证了盲区更新和 PBU 的作用：移除其中任一项都会使 ATE 分别增加 61% 和 58%（表7）。

表 4. 平移端点漂移 (米，↓)。M3DGR 仅为 Corridor01/02 提供端点真值；另外三个选定序列的评估结果见表 3。对于 WHOLE，数值为分组均值（样本标准差）。一个 WHOLE 序列可能属于多个组。 * 表示失败。

方法	M3DGR		WHOLE（我们的方法）			
	Corridor01	Corridor02	弱几何	高动态	低光照	Easy
FAST-LIO2	18.317	17.491	3.868 (5.349)	2.025 (3.998)	**1.039** (2.541)	2.746 (6.069)
FAST-LIVO2	34.548*	1.319	2.719 (3.690)	1.806 (3.516)	1.418 (2.456)	0.066 (0.116)
Gaussian-LIC2	14.272	19.512	3.069 (3.927)	1.988 (3.738)	1.712 (2.682)	**0.020** (0.007)
地面融合++ (LIC)	2.700	5.065	135.126* (367.835)	18.405 (60.137)	1.753 (3.303)	245.128* (548.105)
ODGS-SLAM	11.309	9.252	35.309* (31.618)	22.215* (21.939)	27.858* (35.408)	8.044 (6.632)
PanoGS-SLAM	49.698*	10.223	16.895 (16.824)	12.578 (16.043)	16.297 (19.311)	5.709 (2.762)
GOLD-SLAM（我们的方法）	**1.161**	**0.274**	**1.808** (2.276)	**1.021** (2.060)	1.302 (2.041)	0.065 (0.116)

4.4 渲染与几何重建

我们在共享的前视图上评估序列内新视角渲染和深度一致性，各序列的结果见附录 J。表 5中，GOLD-SLAM 在全部四个公共数据集和四个 WHOLE 组中均取得最低的平均深度 L1 误差；在 WHOLE 组中，相较于 Gaussian-LIC2，误差降低了 20–33%。在大多数渲染对比中，GOLD-SLAM 也表现领先，整体优于 Gaussian-LIC2。相较于 ODGS-SLAM 和 PanoGS-SLAM，GOLD-SLAM 的平均峰值信噪比分别提高了 9.10 和 5.56 分贝，平均深度 L1 误差分别降低了 71.9% 和 83.9%。尽管这两个全景基线都使用了 GOLD-SLAM 的融合深度 D_F，这一几何优势依然存在。除深度补全外，GOLD-SLAM 还使用测得的激光雷达深度来监督地图优化。表7 中的仅 DAP 消融实验表明了测得深度的作用：在前端位姿相同的情况下，用预测值替代激光雷达测量值会增大重建误差。

更高的图像保真度并不一定意味着更准确的表面几何。例如，在 Grass02 上，PanoGS-SLAM 的峰值信噪比高于 GOLD-SLAM（22.87 对比 18.92 分贝），但深度 L1 误差也更高（11.312 对比 6.980 米；见附录 J）。

表 5. 序列内新视角渲染与深度一致性结果（基于共享的 512 × 384 前视图（97°水平视场角））。数值为各序列的均值（样本标准差）；深度 L1 误差为相对于激光雷达测得深度的径向深度平均绝对误差。

数据集	方法	渲染质量			几何精度
		峰值信噪比↑(分贝)	结构相似性指数 ↑	学习感知图像块相似度 ↓	深度-L1 ↓(米)
Hilti-牛津 2022年	Gaussian-LIC2	26.31 (1.90)	0.895 (0.028)	0.236 (0.068)	1.052 (0.692)
	ODGS-SLAM	13.88 (1.73)	0.493 (0.114)	0.610 (0.031)	4.653 (3.503)
	PanoGS-SLAM	19.12 (1.35)	0.776 (0.050)	0.528 (0.046)	5.512 (4.149)
	GOLD-SLAM（我们的方法）	**27.13** (1.61)	**0.916** (0.024)	**0.186** (0.070)	**0.869** (0.718)
M3DGR	Gaussian-LIC2	22.76 (3.69)	0.726 (0.172)	0.332 (0.086)	3.687 (3.045)
	ODGS-SLAM	19.96 (3.03)	0.621 (0.118)	0.364 (0.043)	6.132 (5.465)
	PanoGS-SLAM	22.38 (1.89)	**0.732** (0.131)	0.438 (0.046)	7.602 (3.824)
	GOLD-SLAM（我们的方法）	**23.39** (4.07)	0.730 (0.193)	**0.310** (0.112)	**3.247** (2.982)
M2DGR	Gaussian-LIC2	**18.94** (1.07)	**0.683** (0.034)	**0.342** (0.043)	2.058 (1.941)
	ODGS-SLAM	13.21 (1.32)	0.365 (0.061)	0.570 (0.056)	3.735 (2.157)
	PanoGS-SLAM	15.95 (0.89)	0.570 (0.057)	0.509 (0.031)	11.659 (9.770)
	GOLD-SLAM（我们的方法）	18.72 (2.38)	0.670 (0.119)	0.347 (0.088)	**1.653** (1.649)
牛津尖塔	Gaussian-LIC2	24.94 (1.27)	0.841 (0.038)	0.256 (0.047)	1.227 (0.352)
	ODGS-SLAM	15.58 (5.77)	0.589 (0.179)	0.584 (0.071)	2.534 (0.368)
	PanoGS-SLAM	20.26 (1.10)	0.732 (0.051)	0.624 (0.065)	8.356 (1.546)
	GOLD-SLAM（我们的方法）	**26.15** (1.58)	**0.880** (0.029)	**0.164** (0.035)	**0.514** (0.158)
WHOLE-W	Gaussian-LIC2	24.70 (1.54)	0.855 (0.042)	0.192 (0.057)	1.652 (0.910)
	ODGS-SLAM	15.73 (1.74)	0.550 (0.061)	0.526 (0.084)	5.870 (4.182)
	PanoGS-SLAM	18.54 (1.47)	0.714 (0.060)	0.524 (0.093)	9.169 (4.389)
	GOLD-SLAM（我们的方法）	**25.14** (0.91)	**0.862** (0.040)	**0.158** (0.034)	**1.312** (0.842)
WHOLE-H	Gaussian-LIC2	25.66 (3.68)	0.853 (0.058)	0.194 (0.066)	1.625 (0.892)
	ODGS-SLAM	16.11 (2.74)	0.521 (0.072)	0.523 (0.093)	5.945 (3.524)
	PanoGS-SLAM	20.08 (3.53)	0.734 (0.081)	0.478 (0.115)	9.647 (4.128)
	GOLD-SLAM（我们的方法）	**26.48** (3.24)	**0.867** (0.051)	**0.156** (0.040)	**1.294** (0.802)
WHOLE-L	Gaussian-LIC2	**27.58** (3.77)	0.896 (0.034)	0.138 (0.034)	1.097 (0.576)
	ODGS-SLAM	17.31 (2.53)	0.531 (0.060)	0.514 (0.084)	4.068 (2.839)
	PanoGS-SLAM	20.68 (4.19)	0.764 (0.087)	0.457 (0.131)	7.059 (2.097)
	GOLD-SLAM（我们的方法）	27.55 (3.85)	**0.900** (0.032)	**0.128** (0.027)	**0.875** (0.593)
WHOLE-E	Gaussian-LIC2	26.19 (0.47)	0.875 (0.033)	0.173 (0.021)	1.067 (0.462)
	ODGS-SLAM	17.66 (1.28)	0.646 (0.059)	0.443 (0.081)	2.685 (1.395)
	PanoGS-SLAM	21.24 (1.57)	0.761 (0.063)	0.428 (0.060)	7.937 (4.448)
	GOLD-SLAM（我们的方法）	**27.51** (0.54)	**0.894** (0.033)	**0.132** (0.021)	**0.718** (0.400)

定性评估。 在图 6中，Gaussian-LIC2 在中央视图中保留了细节，但其重建结果朝向外围区域逐渐变得不够清晰。ODGS-SLAM 和 PanoGS-SLAM 呈现出模糊纹理和定义不清的结构。相比之下，GOLD-SLAM 在更宽的视场范围内保持了更清晰的立面边界和路面纹理。图 7 中的全景图展示了超出前向视图范围的这些改进。

计算效率。在表6中，GOLD-SLAM 耗时 95.4–201.6 秒，在两个序列上均短于 ODGS-SLAM 和 PanoGS-SLAM。Gaussian-LIC2 速度更快，但在室外_01 和 Keble03 上生成的地图更小。在 Keble03 上，GOLD-SLAM 使用了 1.407 百万个高斯，而 Gaussian-LIC2 使用了 0.407 百万个，这体现了扩大场景覆盖范围所需的额外建图成本。此外，GOLD-SLAM 支持在 Jetson Orin NX（16 GB）上在线部署（附录 A），同时保持了相当的定位精度

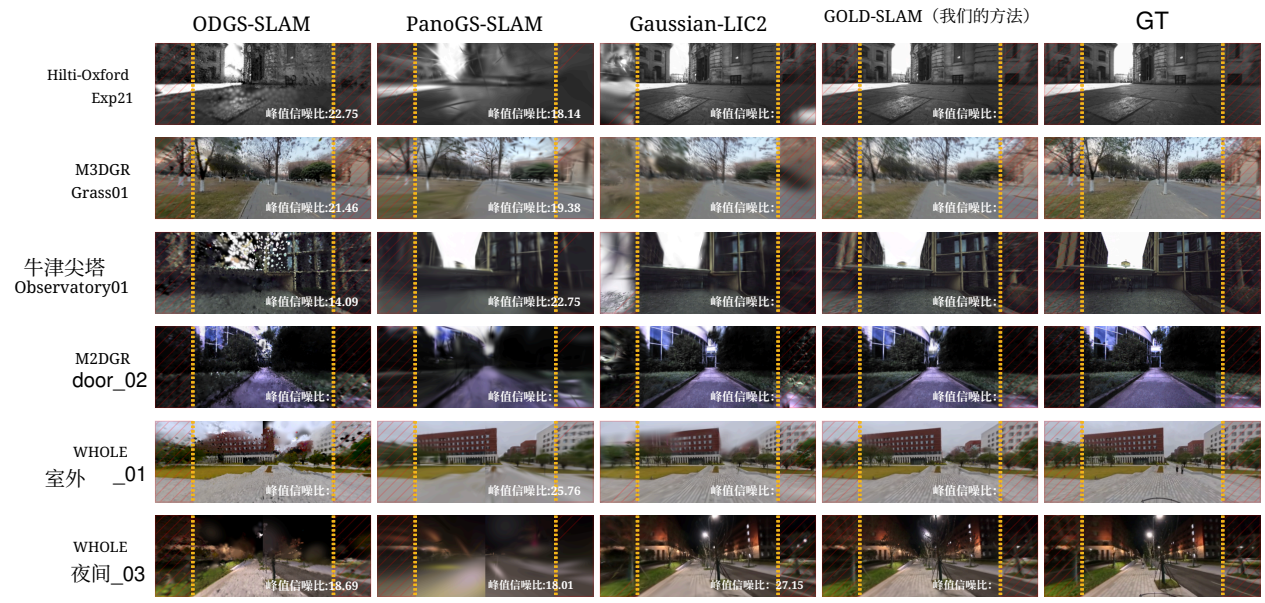


图6。 场景重建的定性比较。 所有面板均显示 120° 水平视场角。全景渲染被投影到相同的透视视图中。黄色虚线标示中央 97° 水平视场角。



图7。 全景重建比较 在 WHOLE户外_01 （顶部）和 WHOLE户外_02 （底部）。

在已完成的 WHOLE序列上，与主要实验配置相比，整体平均峰值信噪比降低了 5.80 分贝。

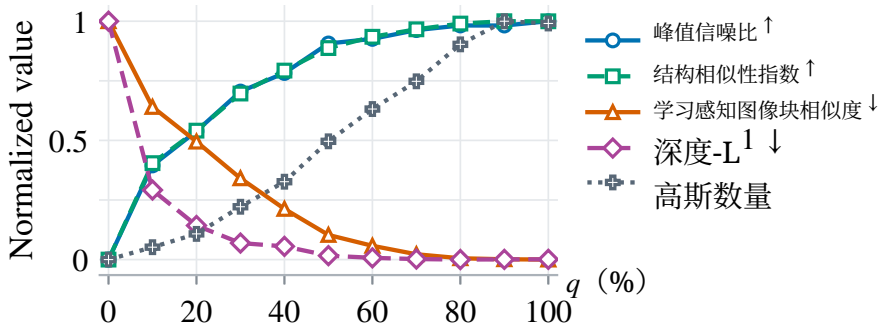
表 6。 计算成本比较 在每个序列的前 600 个输入帧上进行。D、FE 和 BE 分别表示深度估计、里程计和建图。N_{GS} 表示最终的高斯数量。

方法	RAM [吉比字节]	VRAM [吉比字节]	T_D [s]	T_{FE} [s]	T_{BE} [s]	总计 [s]	N_{GS} [10^6]
室外_*01* 从 WHOLE 中随机选择							
Gaussian-LIC2	4.60	**1.38**	**0.7**	**18.8**	**9.3**	**20.6**	0.049
ODGS-SLAM	**4.42**	10.98	58.9	718.0	765.6	946.6	0.908
PanoGS-SLAM	4.47	4.29	4.7	154.6	290.1	318.3	0.042
GOLD-SLAM	7.56	3.52	11.6	20.6	81.9	95.4	0.500
***Keble03* 从牛津尖塔中随机选取**							
Gaussian-LIC2	5.18	**3.67**	**0.8**	**41.8**	**26.4**	**45.3**	0.407
ODGS-SLAM	4.68	7.22	58.5	546.8	641.5	753.1	0.653
PanoGS-SLAM	**4.56**	4.09	6.8	114.4	229.2	247.0	0.025
GOLD-SLAM	7.89	4.28	12.3	46.7	179.6	201.6	1.407

4.5 消融研究

激光雷达盲区分析。在 Oxford Spires Keble03 数据集上，我们逐步裁剪建图后端使用的激光雷达垂直视场，同时保持其水平视场角和前端不变。保留的立体角比例 q 以 10 个百分点为步长，从 100% 降至 0%。在 $q = 0$ 时，高斯初始化完全依赖 DAP，不使用激光雷达尺度对齐或深度监督。

如图所示。8，将 q 从 100% 降至 50% 后，峰值信噪比仅下降 0.33 分贝，而深度 L1 误差则从 0.242 米升至 0.390 米。进一步降低 q 会导致渲染质量和几何精度显著下降。若后端不使用激光雷达，深度 L1 误差将达到 9.797 米。



Figu重新 8. 激光雷达视场对 GOLD-SLAM 的分析。每个指标均独立进行最小–最大归一化 ed.

组件分析。 表 7 评估了 Hilti Exp18 上的系统组件。禁用盲区视觉更新或 PBU（$\sigma_\alpha = \sigma_b = 0$）分别使 ATE 从 0.100 m 增加到 0.161 m 或 0.158 m，并使 PSNR 从 25.42 dB 降低到 23.08 dB 或 24.92 dB。三种后端消融实验均共享完整模型的前端位姿。No DAP 仅使用激光雷达初始化高斯，使 PSNR 降至 24.85 dB。仅使用DAP的建图仅使用 DAP 进行后端初始化和深度监督，其深度 L1 误差从 0.450 m 上升至 1.223 m。No DAEP 保留基于不透明度的插入，使 PSNR 降至 25.21 dB，并使学习感知图像块相似度从 0.189 增加到 0.214。这表明，当远处高斯具有高不透明度时，DAEP 能够重建基于不透明度的插入可能遗漏的较近表面（图 4）。完整模型在 PSNR、结构相似性指数、学习感知图像块相似度和深度 L1 误差方面均取得最佳表现。

表7. GOLD-SLAM各组件的消融实验。 阴影部分突出显示了与完整模型的关键比较。

指标	Unit	Full	w/o 盲zone	w/o PBU	w/o DAP	DAP only	w/o DAEP
(a) 准确率							
峰值信噪比 ↑	dB	**25.42**	23.08	24.92	24.85	24.61	25.21
结构相似性指数 ↑	–	**0.888**	0.835	0.876	0.881	0.863	0.887
学习感知图像块相似度 ↓	–	**0.189**	0.261	0.213	0.219	0.244	0.214
深度-L1 ↓	米	**0.450**	0.540	0.473	0.451	1.223	0.458
绝对轨迹误差 ↓	m	**0.100**	0.161	0.158	**0.100**	**0.100**	**0.100**
(b) 计算成本							
RAM	GiB	14.76	11.45	12.88	**10.44**	10.80	11.89
VRAM	GiB	3.81	3.81	3.81	**1.58**	3.56	3.77
总计	s	184.9	187.2	182.8	127.9	**117.4**	161.8

5 结论

我们提出了 GOLD-SLAM，将全向视觉、激光雷达和惯性测量单元与统一的球面跟踪和高斯建图相结合。盲区视觉约束和 PBU 支持稳健的位姿估计。实测激光雷达深度和全景深度预测引导高斯初始化，而 DAEP 改进了增量式表面重建。此外，我们构建了 WHOLE，这是一个涵盖手持式和无人机运动、面向复杂环境的新型多模态数据集。实验表明，该方法具备具有竞争力的定位性能，并在广泛场景覆盖范围下提升了几何准确率和渲染质量。我们希望，这项具有挑战性的任务及其数据集能够为社区带来新的启发。

局限性与未来工作。我们的几何评估仅限于具有激光雷达参考深度的区域。未来，我们将利用具有稠密真实几何的模拟场景来评估盲区重建。我们还将改进全景拼接和时间对齐，以减小多相机视差和同步误差对跟踪与建图的影响。

致谢

本研究部分得到了中国国家自然科学基金（项目编号：62503166）、湖南省科研开发项目（项目编号：2026QK3018）以及岳麓山工业创新中心项目的资助。本研究还部分得到了德国亥姆霍兹研究中心协会以及巴登-符腾堡州科学、研究与艺术部（MWK）的资助，资助项目为基于人工智能辅助技术的无障碍合作研究生院（KATE）（项目编号：BW6-03）。

参考资料

[1] FanciSwarm Pro。 https://www.fancinnov.com/FanciSwarm-Pro-lidar。访问日期：2026年11月3日。6[2] Mostafa Ahmadi、Amin Alizadeh Naeini、Mohammad Moein Sheikholeslami、Zahra Arjmandi、Yujia Zhang 和 Gunho Sohn。HDPV-SLAM：采用倾斜激光雷达和全景视觉相机的移动测绘系统混合深度增强全景视觉SLAM。在 IEEE自动化科学与工程国际会议（CASE）上，第1–8页，2023年。2[3] Yichen Chen、Yuqi Pan、Ruyu Liu、Haoyu Zhang、Guodao Zhang、Bo Sun 和 Jianhua Zhang。360ORB-SLAM：一种结合深度补全网络的全景图像视觉SLAM系统。在2024年第27届计算机支持的协同工作与设计国际会议（CSCWD）上，第717–722页，2024年。2[4] Jakob Engel、Vladlen Koltun 和 Daniel Cremers。直接稀疏里程计。《IEEE模式分析与机器智能汇刊》，40(3)：611–625，2018年。4[5] Paul Furgale、Joern Rehder 和 Roland Siegwart。多传感器系统的统一时间与空间标定。在 IEEE/RSJ智能机器人与系统国际会议上，第1280–1286页，2013年。6[6] Sheng Hong、Junjie He、Xinhu Zheng 和 Chunran Zheng。LIV-GaussMap：用于实时三维辐射场地图渲染的激光雷达-惯性-视觉融合。《IEEE机器人与自动化快报》，9(11)：9765–9772，2024年。3[7] Qirui Hu、Zikang Yuan、Tianle Xu、Xiaoxiang Wang、Jinni Geng 和 Xin Yang。全景直接激光雷达辅助视觉里程计。arXiv预印本 arXiv:2409.09287，2024年。1，2[8] Huajian Huang 和 Sai-Kit Yeung。360VO：使用单个360相机的视觉里程计。在 IEEE机器人与自动化国际会议上，第5594–5600页，2022年。1[9] Huajian Huang、Longwei Li、Hui Cheng 和 Sai-Kit Yeung。Photo-SLAM：面向单目、立体和 RGB-D 相机的实时同步定位与逼真建图。在 IEEE/CVF计算机视觉与模式识别会议上，第21584–21593页，2024年。1，2[10] Nikhil Keetha、Jay Karhade、Krishna Murthy Jatavallabhula、Gengshan Yang、Sebastian Scherer、Deva Ramanan 和 Jonathon Luiten。SplaTAM：泼溅、跟踪并建图三维高斯，实现稠密 RGB-D SLAM。在 IEEE/CVF计算机视觉与模式识别会议上，第21357–21366页，2024年。2，5[11] Bernhard Kerbl、Georgios Kopanas、Thomas Leimkühler 和 George Drettakis。用于实时辐射场渲染的三维高斯泼溅。《ACM图形学汇刊》，42(4)：139:1–139:14，2023年。1，2，4[12] Xiaolei Lang、Laijian Li、Chenming Wu、Chen Zhao、Lina Liu、Yong Liu、Jiajun Lv 和 Xingxing Zuo。Gaussian-LIC：结合高斯泼溅与激光雷达-惯性-相机融合的实时逼真SLAM。在 IEEE机器人与自动化国际会议上，第8500–8507页，2025年。1，3[13] Xiaolei Lang、Jiajun Lv、Kai Tang、Laijian Li、Jianxin Huang、Lina Liu、Yong Liu 和 Xingxing Zuo。Gaussian-LIC2：激光雷达-惯性-相机高斯泼溅SLAM。《国际机器人研究杂志》，文章编号02783649261479363，2026年。1，2，3，7

[14] Suyoung Lee、Jaeyoung Chung、Jaeyoo Huh 和 Kyoung Mu Lee。ODGS：利用全向图像和三维高斯泼溅进行三维场景重建。在《神经信息处理系统进展》上，第57050–57075页，2024年。3，4[15] Jiarong Lin、Chunran Zheng、Wei Xu 和 Fu Zhang。R^{2}LIVE：一种鲁棒、实时的激光雷达-惯性-视觉紧耦合状态估计器与建图方法。《IEEE机器人与自动化快报》，6(4)：7469–7476，2021年。1[16] Xin Lin、Meixi Song、Dizhe Zhang、Wenxuan Lu、Haodong Li、Bo Du、Ming-Hsuan Yang、Truong Nguyen 和 Lu Qi。Depth Any Panoramas：用于全景深度估计的基础模型。在IEEE/CVF计算机视觉与模式识别会议上，第26835–26844页，2026年。5[17] Yongqi Mao、Hao Shi、Yufan Zhang、Zhonghua Yi、Xiangfei Guo 和 Kaiwei Wang。PanoGS-SLAM：全景三维高斯泼溅SLAM。在 IEEE/RSJ智能机器人与系统国际会议（IROS）上，2026年。2，5，7[18] Hidenobu Matsuki、Riku Murai、Paul H. J. Kelly 和 Andrew J. Davison。高斯泼溅同步定位与建图。在 IEEE/CVF计算机视觉与模式识别会议上，第18039–18048页，2024年。1，2[19] Ben Mildenhall、Pratul P. Srinivasan、Matthew Tancik、Jonathan T. Barron、Ravi Ramamoorthi 和 Ren Ng。NeRF：将场景表示为用于视图合成的神经辐射场。在欧洲计算机视觉会议上，第405–421页。施普林格，2020年。1[20] Richard A. Newcombe、Shahram Izadi、Otmar Hilliges、David Molyneaux、David Kim、Andrew J. Davison、Pushmeet Kohli、Jamie Shotton、Steve Hodges 和 Andrew Fitzgibbon。KinectFusion：实时稠密表面建图与跟踪。在 IEEE混合现实与增强现实国际研讨会上，第127–136页，2011年。1[21] Tixiao Shan、Brendan Englot、Carlo Ratti 和 Daniela Rus。LVI-SAM：通过平滑与建图实现的紧耦合激光雷达-视觉-惯性里程计。在 IEEE机器人与自动化国际会议上，第5692–5698页，2021年。1[22] Stefan Spiss、Joey Hieronimy、Marcel Ritter 和 Matthias Harders。ODGS-SLAM：全向高斯泼溅同步定位与建图。在 IEEE/CVF计算机视觉与模式识别会议上，第26114–26123页，2026年。1，2，3，4，7[23] Edgar Sucar、Shikun Liu、Joseph Ortiz 和 Andrew J. Davison。iMAP：实时隐式建图与定位。在 IEEE/CVF计算机视觉国际会议上，第6229–6238页，2021年。1[24] Yifu Tao、Miguel Ángel Muñoz-Bañón、Lintong Zhang、Jiahao Wang、Lanke Frank Tarimo Fu 和 Maurice Fallon。牛津尖塔数据集：大规模激光雷达-视觉定位、重建与辐射场方法的基准测试。《国际机器人研究杂志》，45(6)：839–857，2026年。2，7[25] Yusen Xie、Zhenmin Huang、Jin Wu 和 Jun Ma。GS-LIVM：结合高斯泼溅的实时逼真激光雷达-惯性-视觉建图。在 IEEE/CVF计算机视觉国际会议上，第26869–26878页，2025年。3[26] Wei Xu、Yixi Cai、Dongjiao He、Jiarong Lin 和 Fu Zhang。FAST-LIO2：快速直接激光雷达-惯性里程计。《IEEE机器人学汇刊》，38(4)：2053–2073，2022年。7[27] Chi Yan、Delin Qu、Dan Xu、Bin Zhao、Zhigang Wang、Dong Wang 和 Xuelong Li。GS-SLAM：采用三维高斯泼溅的稠密视觉SLAM。在 IEEE/CVF计算机视觉与模式识别会议上，第19595–19604页，2024年。2，5[28] Jie Yin、Ang Li、Tao Li、Wenxian Yu 和 Danping Zou。M2DGR：面向地面机器人的多传感器多场景同步定位与建图数据集。《IEEE机器人与自动化快报》，7(2)：2266–2273，2022年。2，6，7[29] Chongjian Yuan、Wei Xu、Xiyuan Liu、Xiaoping Hong 和 Fu Zhang。用于高精度在线激光雷达里程计的高效概率自适应体素建图。《IEEE机器人与自动化快报》，7(3)：8518–8525，2022年。4[30] Deteng Zhang、Junjie Zhang、Yan Sun、Tao Li、Hao Yin、Hongzhao Xie 和 Jie Yin。迈向鲁棒的传感器融合地面SLAM：全面的基准测试与弹性框架。在 IEEE/RSJ智能机器人与系统国际会议上，第8894–8901页，2025年。2，6，7[31] Lintong Zhang、Michael Helmberger、Lanke Frank Tarimo Fu、David Wisth、Marco Camurri、Davide Scaramuzza 和 Maurice Fallon。Hilti-Oxford数据集：毫米级精度的同步定位与建图基准。《IEEE机器人与自动化快报》，8(1)：408–415，2023年。2，6，7[32] Chunran Zheng。FAST-Calib2：利用反光环形靶标进行激光雷达-相机外参标定。GitHub仓库： https://github.com/xuankuzcr/FAST-Calib2，2026年。访问日期：2026年9月15日。6[33] Chunran Zheng、Qingyan Zhu、Wei Xu、Xiyuan Liu、Qizhi Guo 和 Fu Zhang。FAST-LIVO：快速且紧耦合的稀疏直接激光雷达-惯性-视觉里程计。在 IEEE/RSJ智能机器人与系统国际会议上，第4003–4009页，2022年。1[34] Chunran Zheng、Wei Xu、Zuhao Zou、Tong Hua、Chongjian Yuan、Dongjiao He、Bingyang Zhou、Zheng Liu、Jiarong Lin、Fangcheng Zhu、Yunfan Ren、Rong Wang、Fanle Meng 和 Fu Zhang。FAST-LIVO2：快速直接激光雷达–惯性–视觉里程计。《IEEE机器人学汇刊》，41：326–346，2025年。3，7[35] Xingxing Zuo、Patrick Geneva、Woosik Lee、Yong Liu 和 Guoquan Huang。LIC-Fusion：激光雷达-惯性-相机里程计。在 IEEE/RSJ智能机器人与系统国际会议上，第5848–5854页，2019年。1

GOLD-SLAM：高斯泼溅全向激光雷达深度同步定位与建图

补充材料

A 嵌入式部署

为评估其在无人机应用中的计算可行性，我们在配备 16GB 共享内存的 Jetson Orin NX 上，对 GOLD-SLAM 的在线实现进行了测试，覆盖全部 37 个基准序列。

配置。 我们使用 512×256 等距柱状投影/DAP 输入、半精度浮点数 DAP 推理、零阶球谐函数，以及每次更新 70 个优化视图。传感器数据以 1× 倍速回放，同时异步运行跟踪和建图。

评估与结果。 表8 报告了重建质量、处理速率和内存使用量。四个序列因内存不足（OOM）错误而失败。嵌入式配置的平均峰值信噪比为 19.51 分贝，而主实验所用配置为 25.31 分贝。在已完成的 17 个 WHOLE序列中，平均端点漂移为 0.333 米。同时，设备的物理内存使用量峰值为 14.82 吉比字节。这些结果表明，在经过测试的配置下，尽管重建质量有所降低且可用内存余量有限，嵌入式部署仍具备应用于无人机的可行性。

表 8。 Jetson Orin NX（16 GB）上的重建质量和计算成本。 表中数值为均值及样本标准差。

数据集	重建				处理速率		内存
	峰值信噪比 [分贝]	SSIM	学习感知图像块相似度	深度-L1 [米]	前端 [赫兹]	后端 [关键帧/秒]	常驻集大小 [吉比字节]
Hilti-Oxford	19.70 (2.02)	0.781 (0.061)	0.431 (0.044)	1。178 (0.670)	5.62 (1.97)	0.74 (0.12)	7.53 (0.99)
M3DGR	15.17 (2.52)	0.452 (0.045)	0.612 (0.034)	6.284 (0.721)	6.96 (3.37)	0.40 (0.08)	8.36 (1.47)
M2DGR	15.99 (0.91)	0.549 (0.075)	0.496 (0.024)	2.431 (2.901)	2.86 (1.74)	0.64 (0.17)	7.36 (1.95)
牛津尖塔	21.48 (1.77)	0.745 (0.067)	0.404 (0.073)	0.932 (0.356)	2.83 (0.53)	0.37 (0.03)	6.68 (0.67)
WHOLE	20.85 (3.81)	0.753 (0.092)	0.350 (0.090)	1.490 (0.829)	6.40 (1.34)	0.51 (0.13)	7.85 (1.28)

B 球面几何

状态与位姿雅可比矩阵。前端状态为

$$\mathbf{x}_t = (\mathbf{R}_t, \mathbf{p}_t, \mathbf{v}_t, \mathbf{b}_{g,t}, \mathbf{b}_{a,t}, \mathbf{g}_t). \tag{11}$$

系统位姿由世界坐标系中的 IMU 位姿表示，其中，$\mathbf{R}_t$ 将 IMU 坐标系坐标映射到世界坐标系，$\mathbf{p}_t$ 表示 IMU 位置。其余状态变量$\mathbf{v}_t$、$\mathbf{b}_{g,\,t}$、$\mathbf{b}_{a,\,t}$和$\mathbf{g}_t$ 分别表示速度、陀螺仪偏置、加速度计偏置和重力。

对于世界点$\mathbf{P}^W$，其在 IMU 坐标系和相机坐标系中的坐标分别为$\mathbf{p}^I = \mathbf{R}_t^\mathsf{T}(\mathbf{P}^W - \mathbf{p}_t)$和$\mathbf{p}^C = \mathbf{R}_I^C(\mathbf{p}^I - \mathbf{p}_C^I)$，其中$\mathbf{p}_C^I$是用 IMU 坐标系表示的相机中心。对于右乘旋转扰动和世界坐标系平移增量，$\delta\boldsymbol{\xi} = [\delta\boldsymbol{\theta}^\mathsf{T}, \delta\mathbf{p}^\mathsf{T}]^\mathsf{T}$，可得

$$\frac{\partial \pi(\mathbf{p}^C)}{\partial \delta\boldsymbol{\xi}} = \mathbf{J}_\pi \mathbf{R}_I^C \begin{bmatrix} [\mathbf{p}^I]^\wedge & -\mathbf{R}_t^\mathsf{T} \end{bmatrix}, \tag{12}$$

其中，$[\mathbf{p}^I]^\wedge$是$\mathbf{p}^I$的反对称矩阵。

协方差投影。 式（2）由局部切平面构造 [2, 5] 得出。对于高斯 g，令$R_g = \|\boldsymbol{\mu}_g^C\|_2$ 和 $\mathbf{d}_g = \boldsymbol{\mu}_g^C / R_g$ 分别表示其在相机坐标系中的径向距离和单位方位向量。在图 3 中，z'轴与 $\mathbf{d}_g$对齐，而 x' 和 y' 张成 $\mathbf{d}_g$处单位球面的切平面。令$\mathbf{B}_g \in \mathbb{R}^{2\times 3}$的各行构成该平面的正交归一基，则$\mathbf{B}_g^\mathsf{T}\mathbf{B}_g = \mathbf{I} - \mathbf{d}_g\mathbf{d}_g^\mathsf{T}$。将点一阶投影到单位球面后，切平面协方差以及将切平面坐标映射到 ERP 像素的雅可比映射为

$$\boldsymbol{\Sigma}_g' = \frac{\mathbf{B}_g \boldsymbol{\Sigma}_g^C \mathbf{B}_g^\mathsf{T}}{R_g^2}, \qquad \mathbf{A}_g = R_g \mathbf{J}_{\pi,g} \mathbf{B}_g^\mathsf{T}, \tag{13}$$

其中， $\mathbf{\Sigma}_g^C$是相机坐标系协方差，$\mathbf{J}_{\pi,g}$ 是高斯中心处的投影雅可比矩阵。由于 ERP 投影不受径向缩放影响，$\mathbf{J}_{\pi,g}\mathbf{d}_g = \mathbf{0}$，因此$\mathbf{J}_{\pi,g}\mathbf{B}_g^\top\mathbf{B}_g = \mathbf{J}_{\pi,g}$。代入式（13），并加入低通项，得到

$$\begin{aligned}\mathbf{\Sigma}_g^{\text{img}} &= \mathbf{A}_g\mathbf{\Sigma}'_g\mathbf{A}_g^\top + \epsilon\mathbf{I} \\ &= \mathbf{J}_{\pi,g}\mathbf{\Sigma}_g^C\mathbf{J}_{\pi,g}^\top + \epsilon\mathbf{I}.\end{aligned} \tag{14}$$

C 自适应体素尺度选择

更新三个基础体素尺寸分别为 0.5、1 和 2 m 的激光雷达体素地图，每个地图均保留其八叉树子划分。活动尺度初始化为 0.5 m。系统定期在扫描点的共享子集上使用两折交叉验证评估各尺度。在一个子集上估计的位姿增量会在另一个子集上进行评估，随后交换两者的角色。我们使用留出点到平面残差的负对数似然为每个尺度评分，同时考虑测量不确定性和平面不确定性。未匹配的点会产生惩罚。评分根据近期评估结果取平均值。只有当某个候选尺度持续获得更优评分，并且在每个方向上都至少保留当前尺度后验位姿信息的 80% 时，我们才会切换尺度。

表9。 全部20个WHOLE序列上的体素尺度消融实验。 数值表示平移端点漂移（m，↓）。

序列	固定 0.5 米	自适应
走廊 _01	0.272	0.272
车库_01	4.007	4.007
车库_02	1.350	1.350
车库_03	0.040	0.040
车库_04	0.015	0.015
室内 _01	0.020	0.020
室内 _02	0.008	0.008
运动 _blur_01	0.001	0.001
运动 _blur_02	0.004	0.004
运动 _blur_03	**0.005**	0.006
运动 _blur_04	0.005	0.005
夜间_01	0.004	0.004
夜间_02	0.009	0.009
夜间_03	6.418	**4.983**
夜间_04	0.004	0.004
室外 _01	0.886	**0.057**
室外 _02	**0.008**	0.009
室外 _03	0.016	0.016
室外 _04	6.369	**5.815**
室外 _05	10.292	**3.334**
Mean	1.487	**0.998**

WHOLE 上的消融实验。 与固定尺度基线相比，自适应选择将平均端点漂移从 1.487 米降至 0.998 米（32.9%）（表 9）。在 16 个序列中，前端在大部分运行过程中保留初始尺度，端点漂移的变化不到 1 毫米。所有四项明显改进都出现在室外序列中，在这些序列里，前端在每次运行的大部分时间里都使用更粗尺度。激活时，与初始尺度下的地图相比，更粗尺度的地图平均能够匹配同一批采样激光雷达点中更大比例的点。在开放场景中，较大的体素可以将来自采样稀疏的远处表面的更多激光雷达点聚集起来，以拟合局部平面。

D 方位射线三角测量

对于第3.2节中的盲区约束，令 $\mathcal{H}_i$ 表示在当前帧k之前观测到特征 i 的帧。对于每个 $j \in \mathcal{H}_i$，世界坐标系单位方位向量为 $\mathbf{d}_{ij}^W = \mathbf{R}_{C_j}^W\mathbf{d}(u_{ij}, v_{ij})$，相机中心为 $\mathbf{c}_j^W$。

三角测量点使其到这些视线射线的距离平方和最小：

$$\mathbf{P}_i^W = \arg\min_{\mathbf{P}} \sum_{j\in\mathcal{H}_i} \left\| (\mathbf{I} - \mathbf{d}_{ij}^W \mathbf{d}_{ij}^{W\mathsf{T}})(\mathbf{P} - \mathbf{c}_j^W) \right\|_2^2. \tag{15}$$

每个矩阵 $\mathbf{I}-\mathbf{d}_{ij}^W\mathbf{d}_{ij}^{W\mathsf{T}}$ 都将相机到点位移投影到垂直于观测射线的平面上。在具有足够视差的情况下，我们保留深度为正且方向与历史视线方向一致的点，以形成公式 (3) 中的当前帧残差。

不确定性传播。 为对当前帧视线残差进行加权，我们首先估计三角测量点的协方差，然后将其传播到残差。利用 $n_i = |\mathcal{H}_i|$ 个历史观测以及概括其视线几何的 $\mathbf{A}_i = \sum_{j\in\mathcal{H}_i}(\mathbf{I} - \mathbf{d}_{ij}^W\mathbf{d}_{ij}^{W\mathsf{T}})$，我们将点协方差和残差协方差近似表示为

$$\begin{aligned} \mathbf{\Sigma}_{P,i} &= \big(v_i + (n_i - 1)v_i^{\mathrm{sh}}\big)\mathbf{A}_i^{-1}, \\ \mathbf{\Sigma}_{r,i} &= \mathbf{J}_i\mathbf{\Sigma}_{P,i}\mathbf{J}_i^{\mathsf{T}} + \sigma_\theta^2\mathbf{I}_2. \end{aligned} \tag{16}$$

对于每条历史射线，我们根据距离缩放角噪声和位姿不确定性，估计垂直于该射线方向的方差。v_i 是这些方差中的最大值，v_i^{sh} 是历史视图之间最大的共享位姿贡献。该共享项考虑了历史视图所共有的位姿不确定性；否则，这部分不确定性会被误认为独立噪声并在平均过程中减小。在此，$\mathbf{J}_i = \partial\mathbf{r}_{ik}/\partial\mathbf{P}_i^W$ 包含世界到相机旋转，而 σ_θ 是当前帧角度观测噪声的标准差。

E 图像块亮度不确定性

对于包含 N 个像素的图像块，$\mathbf{i}_{\mathrm{ref}},\mathbf{i}_{\mathrm{cur}} \in \mathbb{R}^N$ 分别表示参考强度和经过几何对齐的当前强度。对齐图像块强度建模为

$$\mathbf{i}_{\mathrm{cur}} = (1+\alpha)\mathbf{i}_{\mathrm{ref}} + b\mathbf{1} + \mathbf{n}, \tag{17}$$

其中，α 是相对增益扰动，b 是加性偏移，$\mathbf{n}$ 是像素噪声。假设增益、偏移和像素噪声相互独立且均值为零，其中 $\mathbb{E}[\alpha^2] = \sigma_\alpha^2$、$\mathbb{E}[b^2] = \sigma_b^2$ 和 $\mathbb{E}[\mathbf{n}\mathbf{n}^{\mathsf{T}}] = \sigma_n^2\mathbf{I}$。给定参考图像块后，残差 $\mathbf{r}^{\mathrm{photo}} = \mathbf{i}_{\mathrm{cur}} - \mathbf{i}_{\mathrm{ref}}$ 的协方差为

$$\begin{aligned} \mathbf{C}_{\mathrm{photo}} &= \mathbb{E}[\mathbf{r}_{\mathrm{photo}}\mathbf{r}_{\mathrm{photo}}^{\mathsf{T}}] \\ &= \sigma_n^2\mathbf{I} + \sigma_\alpha^2\mathbf{i}_{\mathrm{ref}}\mathbf{i}_{\mathrm{ref}}^{\mathsf{T}} + \sigma_b^2\mathbf{1}\mathbf{1}^{\mathsf{T}}, \end{aligned} \tag{18}$$

由于相互独立且均值为零，交叉项消失，因此得到式（5）。

F 全景深度估计器消融实验

我们使用公开的预训练权重比较 DAP [3], BiFuse++ [6]和 UniDAC [1]，且不进行微调。在五个序列中的四个序列上，DAP 的渲染深度 L1 误差最低；在三个序列上，其峰值信噪比和结构相似性指数最高（表 10）。在 gate_03序列上，UniDAC 的原始深度误差最低，而 DAP 在经过激光雷达对齐和深度监督后，渲染误差最低。图 9 显示，DAP 通常能更清晰地呈现前景结构与背景之间的深度边界。这些重建结果表明，DAP 是默认估计器的更佳选择；而在全部五个序列上，BiFuse++ 的运行时间最短，峰值显存使用量最低。

表 10. 全景深度估计器消融实验。 原始深度 L1 衡量激光雷达尺度对齐前的预测误差；渲染深度 L1 衡量最终地图误差。渲染指标使用序列内新视角 ERP。

序列	估计器	PSNR（分贝）↑	SSIM ↑	Raw 深度 L1 误差 [米] ↓	渲染 深度 L1 误差 [米] ↓	T_{total} [秒] ↓	VRAM [吉比字节] ↓	RAM [吉比字节] ↓
Exp21	DAP	**23.52**	**0.910**	**3.053**	**0.531**	270.4	5.46	16.70
	BiFuse++	22.68	0.892	6.949	0.560	**184.3**	**2.96**	**16.40**
	UniDAC	22.64	0.894	4.792	0.544	257.1	7.27	17.32
Corridor02	DAP	25.79	0.879	**0.558**	**0.221**	569.0	7.48	9.59
	BiFuse++	25.85	0.879	0.964	0.227	**532.5**	**5.02**	**9.51**
	UniDAC	**25.87**	**0.881**	0.587	0.230	567.7	8.91	9.89
门控_03	DAP	**16.75**	**0.595**	15.037	**2.022**	374.0	5.88	17.34
	BiFuse++	16.71	0.580	12.970	2.098	**296.2**	**3.34**	**14.35**
	UniDAC	16.45	0.572	**9.406**	2.100	320.1	7.02	14.67
Observatory01	DAP	**25.11**	**0.822**	**2.429**	**0.482**	482.7	7.63	12.56
	BiFuse++	24.80	0.814	4.985	0.525	**390.4**	**4.37**	11.96
	UniDAC	24.93	0.814	8.432	0.516	420.1	8.21	**11.59**
室外 _01	DAP	24.94	0.858	18.743	1.435	250.5	5.46	**11.27**
	BiFuse++	24.65	0.832	12.945	1.454	**194.1**	**3.00**	11.44
	UniDAC	**24.95**	**0.859**	**7.031**	**1.413**	269.8	7.21	11.66

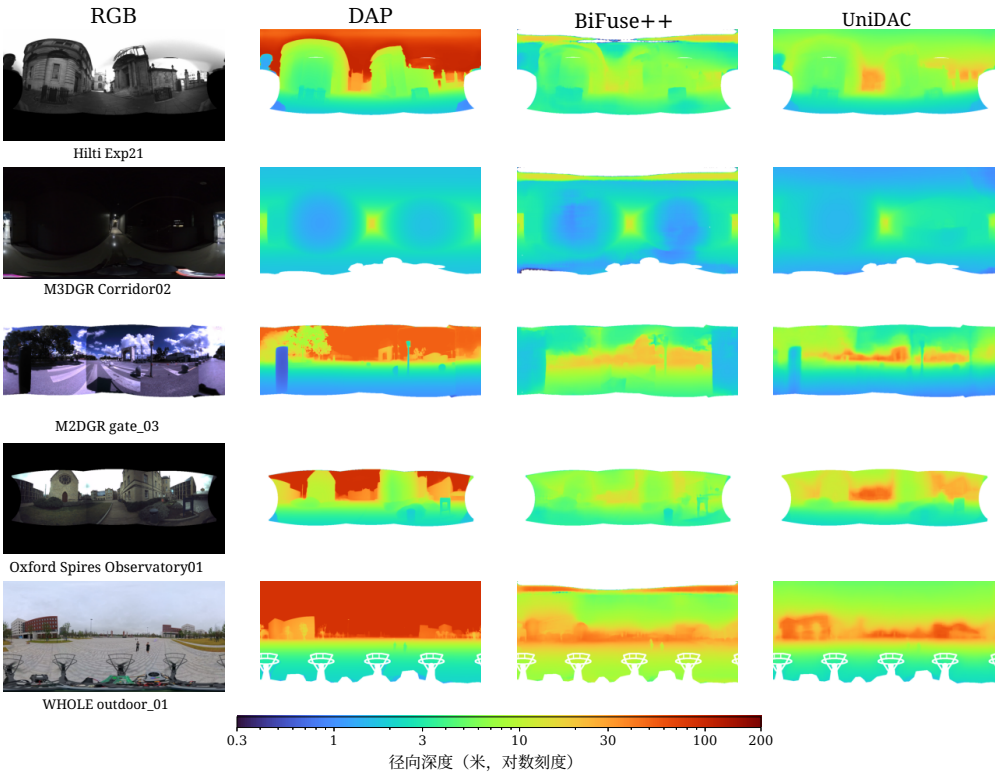


图 9. DAP、BiFuse++和 UniDAC 的激光雷达对齐全景深度预测。

G 激光雷达采样密度消融实验

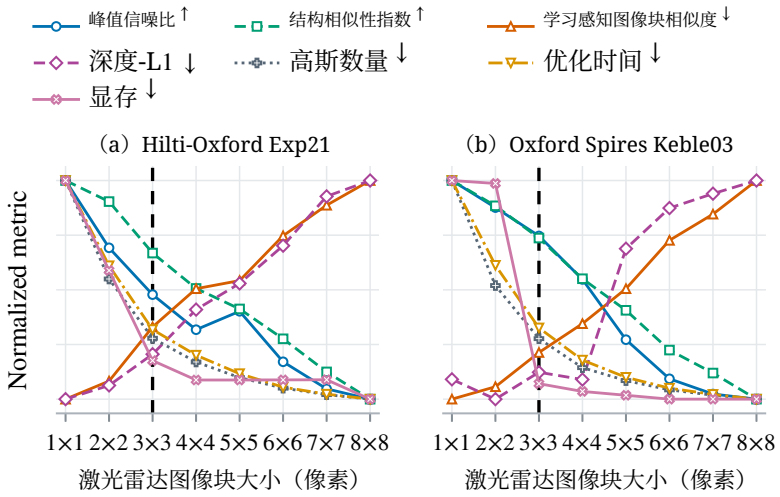


图 10. 激光雷达图像块大小消融实验。 每条曲线均经过最小-最大归一化。竖直虚线标示默认3×3设置。

为从激光雷达初始化高斯，我们保留每个等距柱状投影图像块中径向深度最小的有效点。我们将图像块大小从 1×1 调整为 8×8 像素，使用 Hilti-Oxford Exp21 和 Oxford Spires Keble03 的前 600 个输入帧，并固定 DAP采样。与 1×1 图像块相比， 3×3 图像块使两个序列的最终高斯数量均减少超过 60%，同时降低优化时间和峰值显存使用量。这些节省带来的代价是峰值信噪比下降不到 0.3分贝，深度 L1 误差增加不到 0.02米。当图像块大小超过 3×3 时，可进一步节省优化时间，但会降低结构相似性指数并提高学习感知图像块相似度（图 10）。

H 关键帧选择消融

对于每个序列的前 1,000 个输入帧，我们比较了固定间隔采样（固定）、共可见性与深度归一化平移（C+T） [4], 球面角位移（视差）和位移距离（位移）。在五个序列中，固定采样在其中四个序列上取得了最低的学习感知图像块相似度（表11）。在 Observatory01 上，视差选取了 351 个关键帧，而固定采样选取了 198 个；前者使峰值信噪比提高了 0.43 分贝，但优化时间超过后者的两倍。我们采用固定采样作为默认方案，以保持关键帧率可预测，同时维持有竞争力的重建质量。

表11。 在共享的留出ERP视图上进行的关键帧选择消融实验。

策略	峰值信噪比↑ (dB)	结构相似性指数↑	学习感知图像块相似度↓	深度 L1 误差（米）↓	优化时间（秒）↓
Hilti-Oxford Exp21					
固定	23.51	**0.916**	**0.125**	**0.481**	203.2
C+T	23.25	0.910	0.136	0.512	**125.7**
视差	**23.70**	0.915	0.134	0.494	268.0
位移	23.39	0.913	0.135	0.499	163.5
M3DGR Corridor02					
固定	25.95	0.890	**0.160**	0.220	507.9
C+T	**26.37**	**0.896**	0.172	**0.215**	1158.0
视差	25.61	0.883	0.173	0.236	562.4
位移	25.66	0.884	0.161	0.234	**334.6**

表11（续）

策略	峰值信噪比↑ (dB)	结构相似性指数↑	学习感知图像块相似度↓	深度 L1 误差 (米) ↓	优化时间 (秒) ↓
M2DGR 门限_03					
固定	16.80	**0.607**	**0.359**	**1.982**	298.3
C+T	**16.83**	0.605	0.362	2.077	339.7
视差	16.35	0.588	0.377	2.128	268.7
位移	16.52	0.596	0.372	2.078	**211.4**
Oxford Spires Observatory01					
固定	25.09	0.821	0.120	0.497	**411.7**
C+T	25.44	0.830	**0.113**	0.485	744.3
视差	**25.52**	**0.835**	0.115	**0.478**	965.9
位移	25.35	0.830	0.114	0.491	585.4
WHOLE户外 _01					
固定	25.25	0.860	**0.128**	1.361	200.1
C+T	24.88	0.846	0.136	1.483	**113.4**
视差	**25.31**	**0.864**	0.131	**1.349**	371.9
位移	25.05	0.858	0.133	1.410	238.5

I 优化视图预算消融

在五个序列中，将预算从100 降至70 每轮视图数，平均可将在线总运行时间缩短 20.1%；峰值信噪比下降 0.01–0.25 分贝，深度 L1 误差变化 −0.026 至0.011 米（图11）。这些结果表明，将每轮视图数设为70 ，能够在重建质量与效率之间取得平衡。

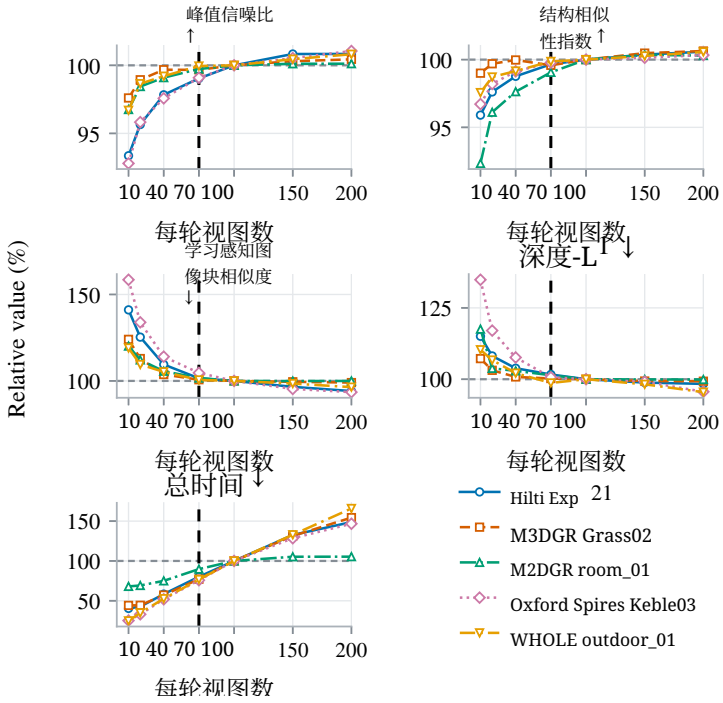


图 11 。**优化视图预算消融**。每项指标均按每轮 100 个视图时的数值进行归一化 d.

J 逐序列评估

表 12 和 13 分别报告了与表 3 和表 4 对应的逐序列定位结果。表 14 和 15 给出了表 5 所依据的渲染和几何结果。

表 12。 逐序列定位精度。 ATE均方根误差（米，↓）。OOM 表示 GPU 内存不足终止。
* 表示失败。

序列（图像数量）	FAST-LIO2	FAST-LIVO2	高斯-LIC2	地面融合++ (LIC)	ODGS-SLAM	PanoGS-SLAM	**GOLD-SLAM** (我们的方法)
Hilti-Oxford 2022年							
实验04（1257）	**0.053**	**0.053**	**0.053**	**0.053**	7.345	7.408	0.054
实验14 (740)	0.318	0.039	0.311	0.367	7.653	7.662	**0.037**
Exp18 (1093)	6.959	0.112	0.305	30.296*	6.229	5.990	**0.100**
Exp21 (1527)	0.132	0.051	0.115	0.054	15.432	15.606	**0.048**
M3DGR							
Grass01 (4318)	**0.530**	0.533	0.531	0.545	4.787	20.180*	0.532
Grass02（2594）	**0.428**	**0.428**	0.429	0.435	3.764	5.930	**0.428**
Varying-illu05（4000）	0.096	0.078	0.177	0.160	1.141	61.840*	**0.073**
M2DGR							
房间_01 （724）	0.400	**0.393**	0.398	0.395	0.441	0.413	**0.393**
门_02 (1256)	0.340	0.330	0.340	**0.316**	4.721	10.280	0.328
大厅_02 (1279)	**0.501**	**0.501**	0.502	0.505	1.786	2.399	**0.501**
大门_03 (2771)	0.190	0.217	**0.180**	**0.180**	OOM	37.414*	0.218
牛津尖塔							
Keble03 (5601)	0.071	0.097	0.245	**0.043**	OOM	31.541*	0.059
Observatory01 (5632)	**0.064**	0.122	0.632	0.269	5.014	28.133*	0.086
ChristChurch03 (6077)	0.027	0.023	0.757	0.037	5.705	16.230	**0.014**
Blenheim01 (7894)	**0.139**	0.158	0.618	0.997	OOM	55.477*	0.150

表13. 各序列平移端点漂移（米，↓）。

序列（图像数量）	FAST-LIO2	FAST-LIVO2	高斯-LIC2	Ground-Fusion++ (LIC)	ODGS-SLAM	PanoGS-SLAM	GOLD-SLAM (我们的方法)
M3DGR							
Corridor01（6058）	18.317	34.548*	14.272	2.700	11.309	49.698*	**1.161**
Corridor02（4400）	17.491	1.319	19.512	5.065	9.252	10.223	**0.274**
WHOLE							
走廊_01 （3395）	13.602	0.273	**0.011**	1225.608*	18.978	7.863	0.272
车库_01 (2900)	**0.055**	3.026	4.136	2.150	23.219*	21.788*	4.007
车库_02 (4712)	**0.719**	1.381	1.964	1.005	106.074*	7.434	1.350
车库_03 (2330)	0.086	**0.035**	0.191	0.723	2.123	9.815	0.040
车库_04 (2810)	0.042	0.011	0.022	**0.008**	4.673	22.782*	0.015
室内_01 (1765)	0.040	0.023	0.021	**0.004**	2.655	5.081	0.020
室内_02 (1528)	0.034	0.005	0.018	**0.002**	9.327	3.251	0.008
运动_模糊_01 (1916)	0.024	0.006	0.014	0.006	32.252*	8.572	**0.001**
运动_模糊_02 (1547)	0.019	0.008	0.019	0.005	21.534*	2.622	**0.004**
运动_模糊_03 (596)	0.010	**0.003**	0.011	**0.003**	3.196	2.354	0.006
运动_模糊_04 (1409)	0.025	0.006	0.033	226.881*	8.056	2.821	**0.005**
夜间_01 (1602)	0.048	0.004	0.052	**0.003**	8.786	2.435	0.004
夜间_02 (2474)	0.038	0.010	0.050	**0.009**	1.579	2.087	**0.009**
夜间_03 (3843)	7.300	6.873	7.266	9.729	47.383*	59.671*	**4.983**
夜间_04 (1734)	0.021	0.005	0.014	0.394	29.028*	4.365	**0.004**
室外_01 (2881)	0.399	0.888	1.473	0.906	24.638*	0.994	**0.057**
室外_02 (1138)	0.025	**0.008**	0.021	0.009	5.721	3.093	0.009
室外_03 (1663)	0.028	0.020	0.030	0.019	3.541	9.256	**0.016**
室外_04 (4146)	8.474	6.868	7.327	8.005	56.730*	22.207*	**5.815**
室外_05 (3578)	11.829	10.544	11.323	10.978	67.498*	26.106*	**3.334**

表 14. 逐序列新视角渲染评估。 峰值信噪比以分贝为单位。

序列（图像数量）	Gaussian-LIC2 峰值信噪比↑	结构相似性指数↑	学习感知图像块相似度↓	ODGS-SLAM 峰值信噪比↑	结构相似性指数↑	学习感知图像块相似度↓	PanoGS-SLAM 峰值信噪比↑	结构相似性指数↑	学习感知图像块相似度↓	**GOLD-SLAM** 峰值信噪比↑	结构相似性指数↑	学习感知图像块相似度↓
Hilti-Oxford 2022												
实验04 (1257)	**26.39**	0.895	0.320	15.80	0.552	0.582	19.51	0.800	0.571	26.36	**0.902**	**0.283**
实验14 (740)	27.78	0.905	0.199	14.54	0.619	0.615	19.45	0.820	0.539	**29.14**	**0.938**	**0.147**
Exp18 (1093)	23.61	0.856	0.258	13.49	0.438	0.590	17.18	0.705	0.538	**25.42**	**0.888**	**0.189**
Exp21 (1527)	27.46	0.922	0.165	11.71	0.365	0.652	20.33	0.779	0.462	**27.59**	**0.935**	**0.124**
M3DGR												
Grass01 (4318)	20.22	0.555	0.419	15.28	0.479	0.430	**20.62**	**0.562**	0.515	19.51	0.531	**0.417**
Grass02 (2594)	19.59	0.524	0.433	19.37	0.507	**0.378**	**22.87**	**0.626**	0.436	18.92	0.509	0.440
Varying-illu05 (4000)	**28.89**	0.889	**0.269**	23.53	0.683	0.319	25.38	0.866	0.424	28.49	**0.891**	0.275
Corridor01 (6058)	23.06	0.835	0.261	21.19	0.708	0.353	21.04	0.791	0.421	**25.14**	**0.861**	**0.209**
Corridor02 (4400)	22.06	0.827	0.279	20.41	0.726	0.338	21.98	0.813	0.395	**24.87**	**0.861**	**0.210**
M2DGR												
房间_01 (724)	**19.81**	**0.639**	**0.283**	13.89	0.304	0.557	16.67	0.489	0.470	16.77	0.504	0.464
门_02 (1256)	19.38	0.715	0.353	14.04	0.426	0.521	15.94	0.586	0.540	**20.52**	**0.759**	**0.275**
大厅_02 (1279)	**17.39**	**0.676**	**0.348**	11.69	0.366	0.631	14.69	0.582	0.498	16.56	0.661	0.367
大门_03 (2771)	19.19	0.703	0.385		OOM		16.50	0.621	0.528	**21.02**	**0.755**	**0.283**
牛津尖塔												
Keble03 (5601)	25.59	0.832	0.265		OOM		20.98	0.721	0.578	**27.94**	**0.884**	**0.145**
Observatory01 (5632)	23.50	0.809	0.288	11.50	0.462	0.634	20.60	0.710	0.587	**25.15**	**0.858**	**0.166**
ChristChurch03 (6077)	26.35	0.896	0.188	19.66	0.715	0.534	20.84	0.806	0.613	**26.96**	**0.920**	**0.133**
Blenheim01（7894）	24.32	0.828	0.284		OOM		18.62	0.691	0.720	**24.55**	**0.860**	**0.212**
WHOLE												
室外_01 （2881）	24.09	**0.823**	0.223	15.60	0.551	0.451	20.51	0.700	0.426	**24.64**	0.819	**0.167**
室内_01 （1765）	26.04	0.894	0.160	19.25	0.722	0.431	21.44	0.809	0.418	**27.96**	**0.916**	**0.116**
室内_02 （1528）	26.77	0.900	0.149	18.18	0.687	0.447	21.55	0.801	0.407	**28.12**	**0.920**	**0.108**
运动_模糊_01 （1916）	24.70	0.868	0.192	15.76	0.558	0.579	18.14	0.763	0.538	**26.97**	**0.904**	**0.127**
室外_02 (1138)	25.91	0.825	0.204	17.90	0.583	0.393	21.49	0.662	0.460	**27.43**	**0.845**	**0.161**
室外_03 (1663)	26.59	0.858	0.173	17.19	0.598	0.366	23.02	0.733	0.348	**27.26**	**0.874**	**0.144**
运动_模糊_02 (1547)	23.70	0.781	0.250	11.55	0.392	0.572	19.43	0.652	0.558	**24.73**	**0.809**	**0.192**
运动_模糊_03 (596)	21.76	0.776	0.276	14.14	0.434	0.571	15.01	0.643	0.639	**25.44**	**0.829**	**0.182**
夜间_01 (1602)	32.83	0.941	0.102	19.20	0.451	0.614	26.15	0.870	0.318	**33.38**	**0.947**	**0.090**
夜间_02 (2474)	**34.14**	0.951	0.097	21.68	0.516	0.464	27.88	0.891	0.297	34.13	**0.953**	**0.096**
走廊_01 (3395)	25.66	0.900	0.178	15.80	0.640	0.577	18.68	0.799	0.507	**26.80**	**0.914**	**0.131**
运动_模糊_04 (1409)	22.17	0.821	0.253	12.16	0.444	0.678	16.78	0.730	0.590	**24.70**	**0.857**	**0.169**
车库_01 (2900)	24.45	0.850	0.172	15.90	0.521	0.512	17.22	0.655	0.618	**25.09**	**0.867**	**0.138**
车库_02 (4712)	**26.09**	0.883	0.132	15.20	0.538	0.534	16.59	0.668	0.650	25.82	**0.888**	**0.121**
车库_03 (2330)	24.94	0.875	0.148	18.37	0.656	0.381	21.46	0.792	0.352	**25.07**	**0.887**	**0.135**
车库_04 (2810)	**26.54**	**0.892**	**0.113**	16.58	0.561	0.496	19.06	0.724	0.465	25.57	0.887	0.118
室外_04 (4146)	23.10	0.794	0.267	16.05	0.578	0.534	18.64	0.650	0.582	**23.74**	**0.802**	**0.224**
室外_05 (3578)	23.04	0.786	0.279	15.75	0.552	0.506	17.93	0.624	0.615	**23.79**	**0.793**	**0.209**
夜间_03 (3843)	**27.01**	**0.902**	**0.148**	17.96	0.513	0.472	18.56	0.727	0.478	25.55	0.876	0.168
夜间_04 (1734)	24.65	0.878	0.194	13.62	0.495	0.642	18.51	0.781	0.481	**25.77**	**0.894**	**0.157**

表15。 逐序列几何重建评估。 深度 L1 误差是相对于激光雷达观测的平均绝对径向深度误差（米，↓）。

序列（图像数量）	Gaussian-LIC2	ODGS-SLAM	PanoGS-SLAM	**GOLD-SLAM**
Hilti-Oxford 2022				
实验04（1257）	1.997	4.542	5.462	**1.873**
实验14（740）	0.426	2.278	2.501	**0.267**
Exp18（1093）	0.669	2.149	2.691	**0.450**
Exp21（1527）	1.114	9.643	11.396	**0.887**

续见下页

表15（续）

序列（图像数量）	Gaussian-LIC2	ODGS-SLAM	PanoGS-SLAM	**GOLD-SLAM**
M3DGR				
Grass01 (4318)	6.199	12.338	11.576	**5.992**
Grass02 (2594)	7.738	11.771	11.312	**6.980**
Varying-illu05 (4000)	1.594	3.447	7.513	**0.819**
Corridor01 (6058)	1.435	1.525	3.568	**1.236**
Corridor02 (4400)	1.470	1.579	4.042	**1.206**
M2DGR				
房间_01 (724)	0.507	1.285	1.517	**0.433**
门_02 (1256)	1.334	4.574	9.879	**1.035**
大厅_02 (1279)	1.495	5.346	10.229	**1.058**
大门_03 (2771)	4.896	OOM	25.011	**4.088**
牛津尖塔				
Keble03 (5601)	1.249	OOM	8.583	**0.357**
Observatory01 (5632)	1.681	2.795	9.786	**0.727**
ChristChurch03 (6077)	0.828	2.274	6.167	**0.449**
Blenheim01 (7894)	1.151	OOM	8.887	**0.523**
WHOLE				
室外_01 (2881)	2.774	12.173	15.625	**2.128**
室内_01 (1765)	0.669	1.415	4.637	**0.434**
室内_02 (1528)	0.821	1.924	5.518	**0.452**
运动_模糊_01 (1916)	0.935	2.707	5.078	**0.587**
室外_02 (1138)	1.581	3.766	11.152	**1.188**
室外_03 (1663)	1.557	4.569	14.136	**1.123**
运动_模糊_02 (1547)	1.918	6.689	10.792	**1.488**
运动_模糊_03 (596)	2.284	4.388	10.422	**1.845**
夜间_01 (1602)	0.585	3.965	6.172	**0.441**
夜间_02 (2474)	0.575	3.426	5.796	**0.417**
走廊_01 (3395)	0.705	1.751	4.243	**0.393**
运动_模糊_04 (1409)	1.185	4.446	6.190	**0.767**
车库_01 (2900)	1.470	2.611	6.965	**0.986**
车库_02 (4712)	1.108	2.721	6.992	**0.841**
车库_03 (2330)	0.911	1.849	5.504	**0.692**
车库_04 (2810)	0.783	2.454	6.423	**0.635**
室外_04 (4146)	2.757	10.336	14.712	**2.378**
室外_05 (3578)	3.137	10.711	15.579	**2.624**
夜间_03 (3843)	2.324	10.697	12.087	**2.266**
夜间_04 (1734)	1.020	4.821	6.535	**0.721**

参考资料

[1] Girish Chandar Ganesan、Yuliang Guo、Liu Ren 和 Xiaoming Liu。UniDAC：适用于任意相机的通用度量深度估计。载于 *IEEE/CVF*计算机视觉与模式识别会议，第 26953–26963 页，2026年。 16[2] Suyoung Lee、Jaeyoung Chung、Jaeyoo Huh 和 Kyoung Mu Lee。ODGS：利用三维高斯泼溅从全向图像进行三维场景重建。载于 神经信息处理系统进展，第 57050–57075 页，2024年。 14[3] Xin Lin、Meixi Song、Dizhe Zhang、Wenxuan Lu、Haodong Li、Bo Du、Ming-Hsuan Yang、Truong Nguyen 和 Lu Qi。Depth Any Panoramas：用于全景深度估计的基础模型。载于 *IEEE/CVF*计算机视觉与模式识别会议，第 26835–26844 页，2026年。 16[4] Hidenobu Matsuki、Riku Murai、Paul H. J. Kelly 和 Andrew J. Davison。高斯泼溅同步定位与建图。载于 *IEEE/CVF*计算机视觉与模式识别会议，第 18039–18048 页，2024年。 18[5] Stefan Spiss、Joey Hieronimy、Marcel Ritter 和 Matthias Harders。ODGS-SLAM：全向高斯泼溅同步定位与建图。载于 *IEEE/CVF*计算机视觉与模式识别会议，第 26114–26123 页，2026年。 14[6] Fu-En Wang、Yu-Hsuan Yeh、Yi-Hsuan Tsai、Wei-Chen Chiu 和 Min Sun。BiFuse++：用于 360° 深度估计的自监督高效双投影融合。 *IEEE*模式分析与机器智能汇刊，45(5):5448–5460，2023年。 16