arXiv:2601.07484v1 [cs.GR] 12 Jan 2026

$\mathbb{R}^3$-RECON：基于可渲染性的无辐射场主动重建

Xiaofeng Jin Matteo Frosi Yiran Guo Matteo Matteucci
米兰理工大学，意大利
{xiaofeng. jin，matteo.frosi，yiran.guo，matteo.matteucci}@polimi.it

代码链接：github.com/jkff00/r3recon

摘要

在主动重建中，具身智能体必须决定接下来该观察何处，以便高效获取有助于高质量新视角渲染的视图。近期关于神经渲染主动视角规划的研究，主要通过对辐射场进行反向传播，或估计三维高斯基元上的信息熵来推导下一最佳视角（NBV）准则。尽管这些策略效果显著，却将视角选择与计算开销大且依赖特定表征的机制紧密耦合，也未考虑轻量级在线部署所需的计算与资源约束。本文从以可渲染性为中心的视角重新审视主动重建。我们提出 $\mathbb{R}^3$-RECON，这是一种无辐射场主动重建框架，可根据轻量级体素地图，在 SE(3) 上构建隐式的姿态条件可渲染性场。我们的方法将逐体素在线观测统计汇总为统一的标量可渲染性评分，该评分更新开销低，并且无需梯度或辐射场训练，即可在数毫秒内以闭式形式查询任意候选视点。该可渲染性场与图像空间重建误差高度相关，能够自然地引导下一最佳视角选择。我们还提出一种全景扩展，用于估计全向（360°）视图效用，从而加快候选视图评估。在标准室内 Replica 数据集上，在视角与时间预算相同的条件下，$\mathbb{R}^3$-RECON 相较于近期主动高斯泼溅基线，实现了更均匀的新视角质量和更高的 3D 高斯泼溅（3DGS）重建精度。

关键词：主动视角规划；信息增益；3D 高斯泼溅

1 引言

近年来，三维高斯泼溅（3DGS） [1] 能够以实时帧率实现高保真新视角合成和稠密重建 [1, 2], ，使基于高斯的地图成为在线机器人感知的理想选择。然而，3DGS 的质量高度依赖视角密度和覆盖范围：输入稀疏或分布不佳会导致某些区域约束不足，进而在新视角中产生伪影。因此，在探索过程中进行主动视点选择对于减少冗余数据、同时实现高质量重建至关重要 [3–5]。我们研究在线下一最佳视角（NBV）选择

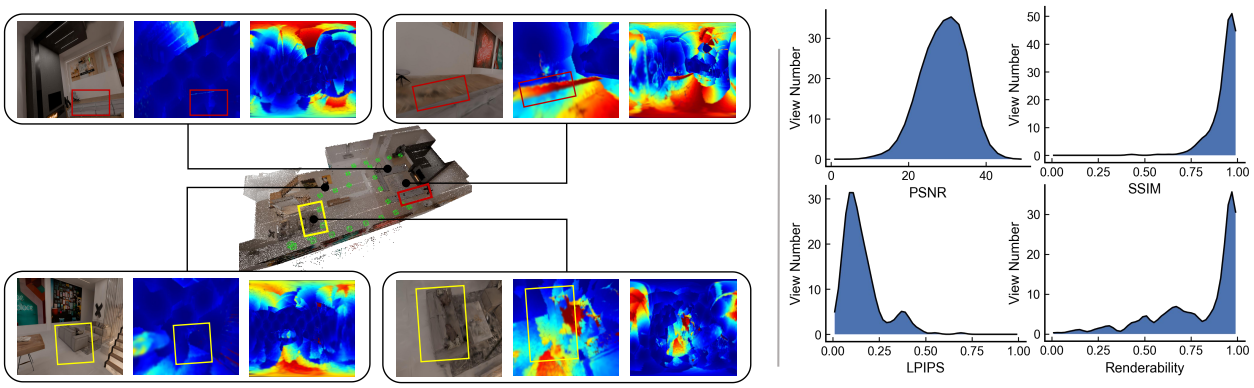


图 1：可渲染性概览。左：绿色相机表示训练视图，黑点表示新视角测试视图；我们展示 RGB 渲染结果、逐视图可渲染性和全景可渲染性。彩色方框突出显示局部区域，这些区域的新视角伪影会随视点变化，并能被可渲染性准确捕捉。右：227 个新视角下学习感知图像块相似度、结构相似性指数、峰值信噪比和我们的可渲染性的分布，显示出它们与图像空间重建质量高度一致。

问题，针对配备机载 RGB-D 相机的 3DGS，智能体会逐步收集观测，并必须决定接下来在哪里观察，以最大限度地提升后续渲染质量。

在线主动重建的一项关键挑战，是在严格的延迟和内存预算下为候选视图评分，同时确保重建质量得到可靠提升。大多数现有方法根据在线优化的辐射表示的内部状态估算信息增益，包括 NeRF 风格的场和 3D高斯泼溅：它们根据模型参数、梯度或训练损失计算基于不确定性或信息论的评分，并通过最大化这些评分来选择下一个视图 [6–8]。这种对模型状态的依赖在在线场景中存在问题。在资源受限的平台上，维护并反复查询大型辐射表示代价高昂，而且在收敛之前，优化状态信号可能并不可靠。近期针对 3D高斯泼溅的敏感性分析 [9]表明，基于损失的二阶信号只有在表示充分收敛后才会趋于稳定，因此早期阶段的视图效用估计会带有噪声，并可能导致短视的选择。

我们将视图效用与瞬态优化状态解耦，并根据历史观测对视角相关辐射亮度的约束程度来表示信息增益。如图1 所示，我们通过一种隐式可渲染性场来表示这种约束，该场直接由在线聚合的轻量级方向观测统计量构建。基于这一思路，我们提出了 $\mathbb{R}^3$-RECON，这是一种无辐射场主动重建框架，可维护体素统计图，并在毫秒级时间内查询任意候选视点的闭式可渲染性评分。我们进一步提出一种全景位姿级公式，用于估计全向（360°）视图效用，以加速候选视点评估。在 Replica 数据集上，在视角与时间预算相同的条件下，所提出的方法在主动 3DGS 重建任务中优于近期基线方法。

我们的贡献如下：(1) 一种不依赖辐射场的闭式视图效用公式，基于观测统计驱动的可渲染性，无需依赖瞬态模型优化状态即可实现稳定的在线评分。(2) 一种轻量级体素统计图，可在任意候选视点实现毫秒级的下一最佳视点评估。(3) 一种 360° 全景式、位姿级效用，以及大量 Replica 实验，证明在预算相同的情况下能够提升主动式 3DGS 重建效果。

2 相关工作

2.1 几何驱动的主动重建。

下一最佳视点（NBV）规划是在线探索和主动建图中的经典课题：机器人逐步构建概率占据地图，并选择视点以降低几何不确定性 [10]。代表性策略包括基于前沿的探索 [11] 以及根据信息熵预期下降为动作评分的信息增益目标 [12, 13]。这些思路也被扩展到主动式三维体积重建，通过在体素或八叉树地图上计算信息增益 [14, 15]。尽管这些准则在几何覆盖和未知空间发现方面有效，但它们与光度保真度仅有间接关联；当神经渲染成为重建目标时，这促使人们采用具备渲染感知能力的下一最佳视点方法。

2.2 在线光度建图。

神经渲染使新视角质量取决于所采集视图的密度和方向分布 [16]。近期的在线重建骨干网络使得从流式帧中增量优化辐射表示切实可行：它们不再针对所有历史数据重新训练，而是通过更新滑动窗口和/或一小组活跃参数来细化模型，通常辅以关键帧缓冲或回放 [17–21]。与此同时，3D高斯泼溅提供了一种基于显式基元的辐射表示，可实现实时、高保真渲染 [1], ，从而推动了 SplaTAM 和 MonoGS 等在线 3D高斯泼溅建图流程 [22, 23] ，以及面向更具挑战性场景的其他变体 [24–26]。

2.3 渲染驱动的主动重建。

基于在线光度建图，渲染驱动的 NBV 使用在线光度重建状态取代几何地图，并根据不断演进的模型上的信息增益或不确定性降低程度对候选位姿进行评分。因此，大多数渲染驱动的 NBV 方法都从不断演进的光度模型的内部优化信号中推导视图效用。对于辐射场，具有代表性的方法会估计参数空间中的信息增益或不确定性，例如费 Fisher 信息或相关准则 [6, 27, 28]；对于 3D高斯泼溅，效用通常依据基元级置信度、概率不确定性或互信息类目标计算 [7,8, 29, 30]。其他流程则将视角选择与在线建图/规划相结合 [31–33] ，或使用学习得到的排序信号快速筛选候选项 [34]。尽管公式各异，但这些方法有一个共同点：依赖反复维护并查询复杂的优化状态（参数/梯度/损失），这在在线延迟和内存受限时成本高昂，而且在收敛前可能并不可靠；值得注意的是，PUP-3DGS [9] 表明，基于损失的二阶敏感性主要要在经过充分优化后才有意义。

2.4 以可渲染性为中心的主动重建。

为缓解模型状态驱动的信息增益的局限性，我们采用以可渲染性为中心的视角，根据给定历史观测时某个候选视点能否被可靠渲染来刻画视图效用。Yi 等人借助粗略几何，在基于图像的渲染中引入了可渲染性

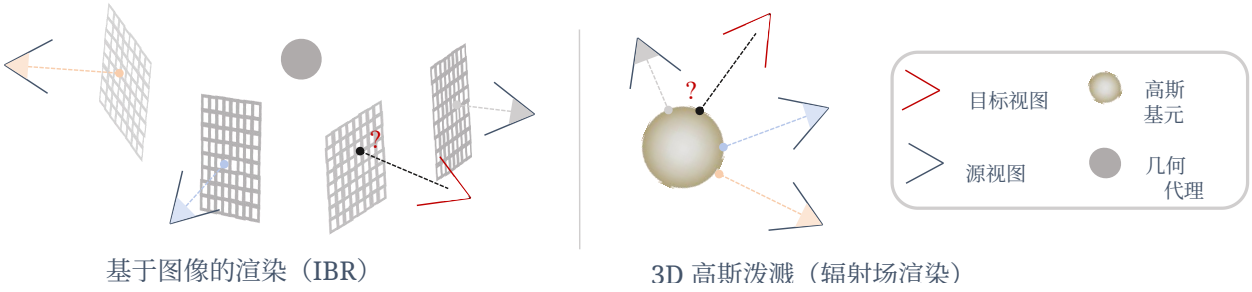


图 2：IBR 与辐射场渲染中的可渲染性。IBR 根据粗略的几何代理和源视图重叠度估计视图可行性，而 3D高斯泼溅 则取决于基元外观在不同观察方向上的约束程度。

几何代理和观测统计来估计新视角合成的可靠性，以进行视点/轨迹选择 [35]。Jin 等人将这一思想应用于 3D高斯泼溅，以在稀疏输入下指导伪视图采样 [36]。然而，现有的可渲染性模型大多建立在基于图像的渲染之上，无法直接刻画辐射表示中的不确定性。受对 3D高斯泼溅 中不确定性来源的分析启发 [37], 我们从不确定性的角度重新诠释可渲染性，并将其扩展为一种在线形式，从而实现高效的视角选择，最终得到轻量级的主动光度重建。

3 技术方法

3.1 概述

我们提出了一种观测统计驱动的可渲染性效用，用于在不依赖瞬态重建优化状态的情况下进行毫秒级在线 NBV 评分。首先，我们从基于图像的渲染（IBR）视角回顾可渲染性的概念，并将其与视角相关辐射表示中的可渲染性进行对比。随后，我们推导出一种解析的基元级可渲染性公式，并展示如何利用轻量级观测统计在线计算该公式。最后，我们将所得的可渲染性评分整合到主动重建流程中，以实现高效的下一最佳视角选择，其中包括用于快速视线方向规划的全景位姿级评估。

3.2 预备知识

可渲染性由 [35] 提出，作为基于图像的渲染（IBR）的一种预评估指标。如图 2所示，通过将目标视点与历史观测进行比较，可渲染性本质上衡量了用于合成的源像素与目标视图的对齐程度。当场景完全符合朗伯特性时，IBR 对视点变化的敏感性会降低，对视点差异的依赖也会减弱。

对于辐射表示，可渲染性的含义有所不同。辐射场本质上具有视角依赖性，因此即使在弱反射区域，也仍需要足够的方向观测。从这个角度看，给定历史采样光线，可渲染性取决于局部基元在查询视线方向上拟合视角相关外观的能力。

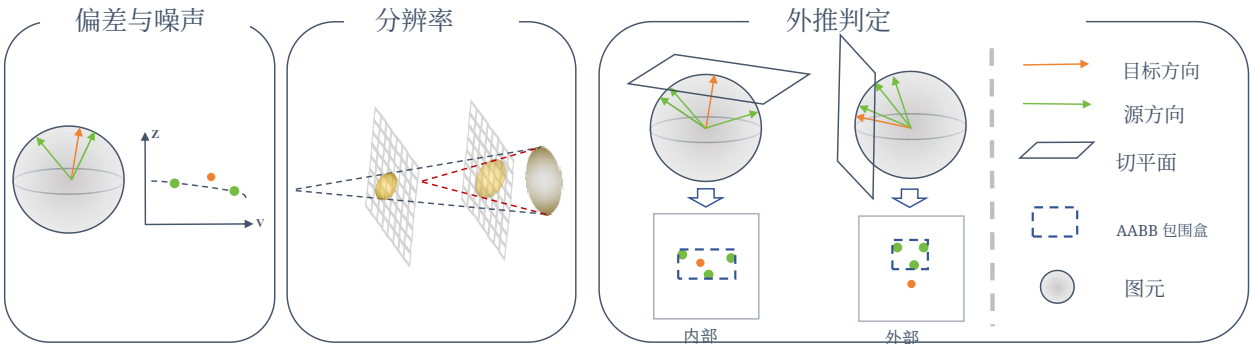


图 3：可渲染性因素。可渲染性由方向偏差、外观噪声和分辨率项共同决定。v 表示观察方向，z 表示观测辐射亮度；橙色点表示新方向上的（未知）辐射亮度。噪声反映视角相关变化，偏差则衡量查询方向偏离历史观测方向支持域的程度。通过将过往方向投影到查询切平面，并检查查询是否位于其支持区域之外，即可检测外推。

3.3 辐射亮度表示中的可渲染性

我们通过分析单个基元的球谐函数（SH）拟合行为来研究其可渲染性。根据 3D高斯泼溅中的不确定性来源，我们用两个因素刻画 SH 拟合质量：观测偏差和噪声（图 3）。观测偏差衡量查询方向与历史观测方向之间的角度差异；偏差为零表示方向相同，拟合损失最小。观测噪声概括历史辐射观测的离散程度，并反映与材质相关的视角变化，非朗伯区域中的这种变化通常更大。因此，对于每个基元 p_i，从视点 s（相机位姿 $T^{(s)}$）获取的观测建模如下：

$$\underbrace{\mathbf{z}_i^{(s)} - \mathbf{z}_i^{\star}}_{\text{rendering deviation}} = \mathcal{F}\left(b_i^{(s)}, \varepsilon_i^{(s)}\right), \tag{1}$$

其中$\mathbf{z}_i^{\star}$表示真实（未知）的样本值，$b^{(s)}{}_i$ 表示与视点相关的偏差，而$\boldsymbol{\varepsilon}^{(s)}{}_i$ 表示来自对p_i的所有历史观测的噪声。

对于球谐函数拟合，同样的采样偏差会产生显著不同的影响，具体取决于查询方向相对于历史样本的位置。根据最小二乘误差传播以及对病态采样下噪声放大的分析[38], ，当查询落在外推区域时，拟合误差可能迅速增大。为描述这一效应，我们遵循标准球面几何，将历史观察方向的球面凸包视为有效插值域。

如图3所示，我们将历史方向投影到查询方向处的切平面上；如果查询点落在投影样本的二维包围盒之外（这是局部凸包近似的一种实用方法），则将其归类为外推。

对于偏差项，我们用二元因子 $\boldsymbol{\kappa}^{(s)}{}_i \in \{1,2\}$表示外推测试的结果，其中$\boldsymbol{\kappa}^{(s)}{}_i = 1$ 表示插值，$\boldsymbol{\kappa}^{(s)}{}_i = 2$ 表示外推。令 $\{\mathbf{r}^{(t)}{}_i\}_{t=0}^{n_i}$ 表示所有历史观测的单位观察方向

，并令 $\mathbf{r}^{(\mathrm{s})}{}_i$ 表示当前候选视图的单位观察方向。计算当前方向与最近历史方向之间的角度偏差：

$$\cos\theta_i^{(\mathrm{s})} = \max_{t\in\{0,\dots,n_i\}} \langle \mathbf{r}_i^{(t)}, \mathbf{r}_i^{(\mathrm{s})} \rangle \tag{2}$$

$$b_i^{(s)} = \left(\cos\theta_i^{(s)}\right)^{\kappa_i^{(s)}}. \tag{3}$$

对于噪声项，我们通过其历史观测之间的相互差异来量化$\{\mathbf{z}^{(t)}{}_i \quad \}^{n_i}{}_{t=1}$。这反映了$p_i$ 的外观在不同视图间的变化程度。具体而言，平均两两差异为

$$\delta_i = 1 - \frac{1}{n_i(n_i-1)} \sum_{\substack{t,t'=1 \\ t\neq t'}}^{n_i} \left\| \mathbf{z}_i^{(t)} - \mathbf{z}_i^{(t')} \right\|^2, \tag{4}$$

其中，数值越小 δ_i 表示观测噪声越大。

然后将这一差异转换为可渲染性相关项

$$\varepsilon_i^{(s)} = \delta_i^{\kappa_i^{(s)}\left(1-b_i^{(s)}\right)}. \tag{5}$$

这种指数形式具有两个有用的性质。首先，底数 $\delta_i \in [0,1]$ 反映了噪声一致性： δ_i 越大，数值越高，意味着历史观测越可靠。其次，指数控制衰减速率。当角度偏差较小（$b^{(s)}{}_i \approx 1$）时，指数趋近于零， $\delta^{1-}{}_{b^{(s)}{}_i} \approx 1$，表示几乎没有渲染损失。偏差越大，指数越大，衰减越快，从而对存在偏差的视点施加更强的惩罚。

如图所示 3，观测是在像素域中离散化的，因此图像分辨率的变化会导致渲染质量出现额外差异。因此，我们引入辅助分辨率项，以考虑观测质量随距离变化的情况。成像分辨率与观察距离成反比。对于无偏深度为 D 的基元，其投影像素足迹满足

$$\text{pixel size} \propto \frac{1}{D}. \tag{6}$$

We d使用基于深度的分辨率因子定义一个与分辨率相关的辅助项

$$\rho_i^{(s)} = \frac{1}{D_i^{(s)}}, \qquad \rho_i^{\max} = \max_t \frac{1}{D_i^{(t)}}, \tag{7}$$

它反映了基元 p_i 在不同视点下的相对像素覆盖范围。相应的分辨率增益为

$$\gamma_i^{(s)} = \min\left(1, \frac{\rho_i^{\max}}{\rho_i^{(s)}}\right)^{1-\delta_i b_i^{(s)}}. \tag{8}$$

表示当前视图是否需要高于最佳历史观测的有效分辨率。如果 $\rho^{(s)}{}_i \leq \rho_i^{\max}$，则 $\gamma^{(s)}{}_i = 1$；否则， $\gamma^{(s)}{}_i < 1$ 表示由于历史分辨率不足而导致的较低重建可靠性。

算法 1 基于斐波那契分箱方向的在线偏差项

Global (once): choose N bins by θ_{res} and precompute $\{\mathbf{q}_k\}_{k=1}^{N}$
$A \leftarrow 2\pi(1-\cos(\theta_{\text{res}}/2))$; $N \leftarrow \lceil 4\pi/A \rceil$; $\varphi_g \leftarrow (1+\sqrt{5})/2$
for $k = 1..N$ **do**
$z_k \leftarrow 1 - 2\frac{k+0.5}{N}$; $r_k \leftarrow \sqrt{1-z_k^2}$; $\phi_k \leftarrow \frac{2\pi k}{\varphi_g}$; $\mathbf{q}_k \leftarrow (r_k\cos\phi_k,\ r_k\sin\phi_k,\ z_k)$
end for
State (per primitive p_i): visited-bin mask $\mathcal{I}_i \in \{0,1\}^N$
Update with a new view $\mathbf{r}_i^{(t)}$: $k^\star \leftarrow \arg\max_k \langle \mathbf{q}_k, \mathbf{r}_i^{(t)} \rangle$; $\mathcal{I}_i[k^\star] \leftarrow 1$.
Query for direction $\mathbf{r}_i^{(s)}$: $b_i^{(s)} \leftarrow \max_{k:\mathcal{I}_i[k]=1} \langle \mathbf{q}_k, \mathbf{r}_i^{(s)} \rangle$; **return** $b_i^{(s)}$.

算法2 使用韦尔福德更新进行在线噪声估计

状态（每个基元 p_i）： n_i, $\mu_i \in \mathbb{R}^3$, $\mathbf{M}_i \in \mathbb{R}^{3\times3}$（在 其中：$n_i$=0, μ_i=**0**, $\mathbf{M}_i$=**0**)
更新（$\mathbf{z}^{(t)}{}_i \in \mathbb{R}^3$）：$n_i \leftarrow n_i+1$; $\Delta \leftarrow \mathbf{z}^{(t)}{}_i - \mu_i$; $\mu_i \leftarrow \mu_i + \Delta/n_i$; $\mathbf{M}_i \leftarrow$
$\mathbf{M}_i +$（$\mathbf{z}^{(t)}{}_i - \mu_i$）$\Delta^\top$.**查询**（假设$n_i \geq 2$）：$\text{trs}_i \leftarrow \text{tr}$（$\mathbf{M}_i$）/（$n_i-1$）;
$\delta_i \leftarrow 1 - \alpha\sqrt{\text{trs}_i}$, $\alpha = \sqrt{2}/1.5$.

指数 $1-\delta_i b^{(s)}{}_i$ 控制该可靠性评分的衰减速率：外观稳定（δ_i 较大）且视角支持充分（$b^{(s)}{}_i$ 较大）的图元所需的额外分辨率较少，因此衰减较弱；而外观变化较大或视角存在偏差的图元则需要更高的分辨率，从而导致更强的衰减。

最后，我们将可渲染性定义为三个归一化因子的乘积：视角偏差$b^{(s)}{}_i$、外观稳定性 $\varepsilon^{(s)}{}_i$ 和分辨率增益 $\gamma^{(s)}{}_i$，

$$R_i^{(s)} \;=\; \mathcal{F}\Big(b_i^{(s)}, \varepsilon_i^{(s)}\Big)\,\gamma_i^{(s)} \;=\; b_i^{(s)}\,\varepsilon_i^{(s)}\,\gamma_i^{(s)}, \tag{9}$$

其中 $b^{(s)}{}_i\,\varepsilon^{(s)}{}_i\,\gamma^{(s)}{}_i \in [0,1]$。这种乘法形式得到 $R^{(s)}{}_i \in [0,1]$，并且只有当基元在方向上得到充分支持、外观一致且观测分辨率足够时，才会赋予其较高的可渲染性。

3.4 在线可渲染性

朴素地计算可渲染性，需要汇总所有历史观测的统计信息。由于两个原因，这种做法不适用于在线计算：（i）偏差项必须在连续球面 $\mathbb{S}^2$ 上保留所有历史观测方向，这会占用难以承受的内存，并增加查询成本；（ii）不同基元累积的观测数量各不相同，因而无法高效向量化，并会进一步增加计算量。噪声项也存在同样的历史依赖性，因此在增量探索中，更新成本会随帧数迅速增长。因此，我们采用一种有理论依据的近似方法，使可渲染性的计算不受已存储观测数量的影响。

在算法 1中，我们使用斐波那契格 [39] $\{\mathbf{q}_k\}_{k=1}^{N}$对单位球面进行离散化。每个基元都会在这些均匀分布的分箱上维护一个固定大小的已访问分箱掩码，从而以恒定的内存开销和不受观测数量影响、可完全向量化的计算，高效地进行偏差跟踪与评估。

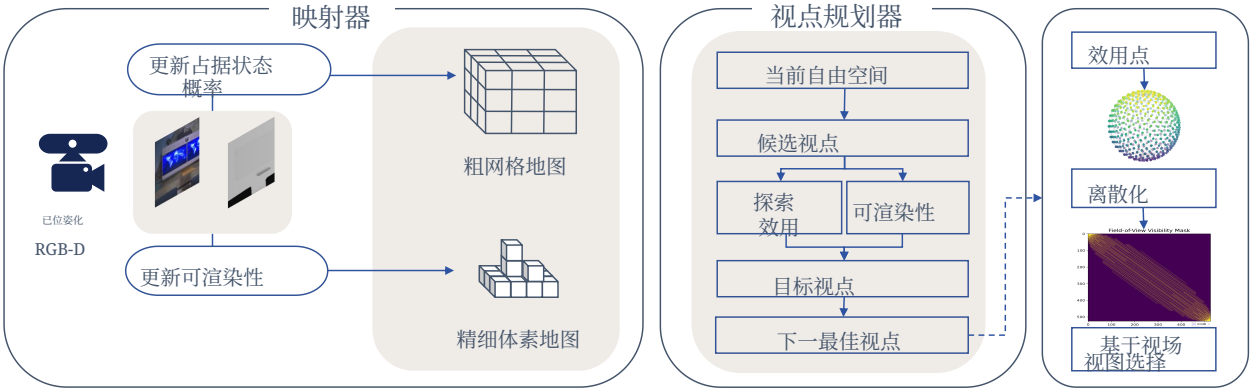


图4：系统概览。带位姿的RGB-D数据会更新一个粗粒度占据栅格，用于探索；同时更新一个精细的体素统计图，用于评估可渲染性。根据探索增益和可渲染性为候选视图评分，并通过球面离散化和视场聚合加速全景方向选择。

对于噪声项，如算法 2所示，每个基元都会维护固定大小的韦尔福德矩，以在线跟踪 RGB 协方差。我们将协方差迹映射为归一化的一致性得分 $\delta_i \in [0,1]$，使用 α。

3.5 基于可渲染性的主动重建

我们基于 Active-GS，将所提出的可渲染性融入在线主动重建[8]，并用基于可渲染性的信息增益替换其视图效用。如图4所示，给定带位姿的 RGB-D 流，我们维护两种地图。第一种是粗粒度占据地图，用于探索和安全性，这与 Active-GS 的做法一致。第二种是轻量级体素统计图，用于计算可渲染性。体素统计图会降采样至 5 cm 的分辨率，每个体素对应一个存储在线统计信息的局部基元。为了高效评估候选视图，我们扩展了 ODGS 中的光栅化器[40]，使其能够针对任意查询位姿快速计算可见体素集合和全景深度代理。

对于位置t处的候选视点，我们计算可见的未探索栅格单元 $\mathcal{G}$，以及用于可渲染性评估的体素统计图上的可见集合 $\mathcal{V}$。探索效用$U_G = |\mathcal{G}|$ 就是栅格单元的数量。而渲染效用为：

$$U_R = \sum_{v \in \mathcal{V}} (1 - R_v). \tag{10}$$

与 Active-GS 中一样，总体位姿得分是加权组合，用于优先考虑早期阶段的探索：

$$U_{\text{view}} = \lambda_1 U_G + U_R - \lambda_2 U_{\text{path}}, \tag{11}$$

其中，λ_1 和 λ_2 分别是 Active-GS 中探索项和路径代价项的权重，而U^{path} 是在探索网格上计算的到达候选视点的行进代价。

我们进一步使用斐波那契球离散化（第 3.4 节）选择观察方向。我们通过合并 $\mathcal{G}$ 和 $\mathcal{V}$ 的中心，构建统一的三维点集，并将每个点投影到一条单位射线

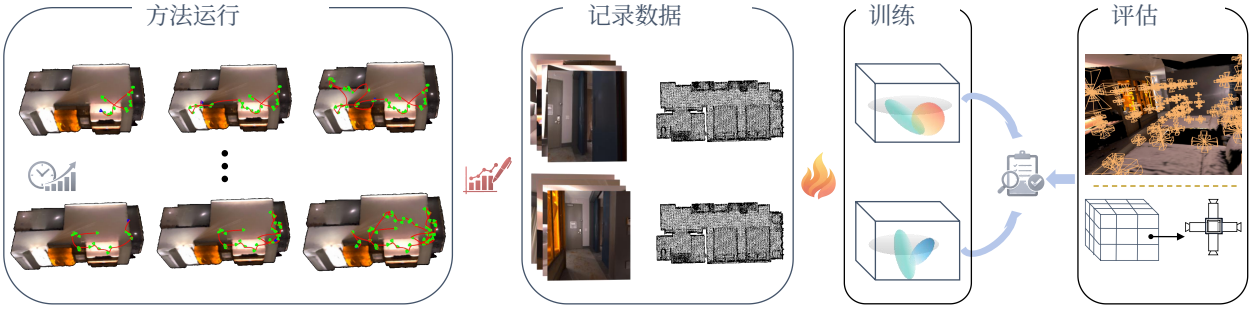


图5：评估流程。每种方法均在线运行以生成轨迹，并收集已位姿化的 RGB 和场景点云。随后，使用记录的数据，在预算匹配的条件下训练 3D高斯泼溅模型。评估时，从自由空间网格位置采样测试视图，每个视图均沿六个规范视角方向进行渲染，以衡量场景级泛化能力（针对留出的新视图）和重建指标。

方向上，该方向位于 $\mathbb{S}^2$，并将其分配到最近的斐波那契分箱k。对于每个分箱 k，当光轴与 k 对齐时，我们预先计算视场覆盖分箱集合 $\mathcal{N}(k)$，并通过对 $\mathcal{N}(k)$ 中各分箱得分求和来计算方向得分。聚合得分最高的分箱将被选为最终观察方向。

4 实验

4.1 实验设置

在所有场景中，用于体素可渲染性更新的斐波那契球包含 64 个采样点。对于全景方向选择，斐波那契离散化的目标中心角为 $10°$。体素尺寸设为 5 cm，以表示局部基元。所有其他超参数均遵循 Active-GS。

所有方法均在 Habitat-Sim 中进行评估[41]。RGB 图像分辨率为 512×512，相机视场为 $[60°, 60°]$。实验在搭载 NVIDIA RTX A6000 GPU（24 GB）的 Ubuntu 20.04 工作站上运行。

Active-GS [8] 和 ActiveGAMER [29] 被用作强有力的基于3D高斯泼溅的基线方法。两者均支持完整的六自由度相机运动，并在 Replica 上取得了有竞争力的表现。针对地面导航量身定制的 ActiveSplat[31] 未纳入比较，因为其运动假设与我们的自由飞行设置不匹配。在表格中，我们的方法将 Active-GS 的评分模块替换为我们提出的信息增益，而 Ours-Pano 则在此基础上进一步加入全景扩展。

4.2 Replica-Dense测试集

重建的目标是从任意自由空间视点进行高保真渲染。为此，我们在九个 Replica[42] 室内场景上构建了密集测试集 Replica-Dense。对于每个场景，我们使用间距为 1.0 m×1.0 m×0.75 m 的三维网格，在 (x, y, z) 中近似并采样无碰撞自由空间。在每个有效网格位置放置六个测试相机，使其光轴分别与 $\pm x$、$\pm y$ 和 $\pm z$ 对齐。所有方法均在相同的 Replica-Dense 测试集上进行评估。

方法	指标	Hotel0	Off0	Off1	Off2	Off3	Off4	Room0	Room1	Room2
ActiveGAMER	SSIM	0.8338	0.8859	0.8032	0.7468	0.7339	0.7396	0.6846	0.8008	0.7169
	PSNR	20.51	26.16	25.74	15.95	16.45	16.02	14.86	19.75	14.55
	学习感知图像块相似度	0.2738	0.1791	0.2549	0.3658	0.3727	0.3530	0.4369	0.3397	0.3995
Active-GS	SSIM	0.8867	0.9138	0.8957	0.9006	0.8792	0.8982	0.8672	0.8778	0.9060
	PSNR	24.60	29.87	31.39	26.91	25.38	26.45	23.01	24.51	25.07
	学习感知图像块相似度	0.21		3	0.1849	0.1936	0.2209	0.2012	0.2597	0.2508	0.2049
Ours	SSIM	**0.9124**	**0.9493**	**0.9338**	0.9162	0.9140	**0.9269**	0.9053	0.9149	**0.9426**
	PSNR	**26.93**	33.95	**37.14**	29.34	28.93	**31.16**	27.85	28.85	**31.24**
	学习感知									
Ours-Pano	SSIM	0.8975	0.9480	0.9298	**0.9305**	**0.9225**	0.9184	**0.9159**	**0.9201**	0.9314
	PSNR	25.95	**34.17**	36.70	**30.74**	**30.91**	30.34	**29.41**	**29.43**	29.23
	学习感知									

表1：固定时间预算下 Replica-Dense 各场景的结果。峰值信噪比/结构相似性指数越高越好，学习感知图像块相似度越低越好。最佳值以彩色粗体标出，次佳值以下划线标出。

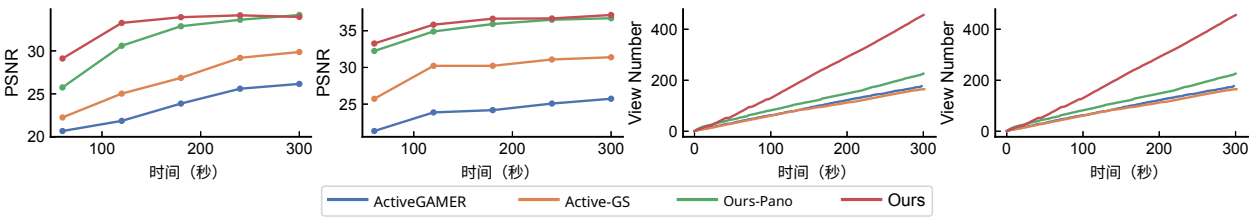


图6：定量评估视图质量与探索效率。每组包含 Hotel0 和 Office0 场景。左侧两幅图展示最终新视角渲染质量随时间的变化。右侧两幅图展示采集图像数量随时间的变化。在相同时间预算下，我们的方法的数据采集效率最高。

如图 5所示，所有方法均被纯粹视为数据收集器。每种方法从相同的初始位姿开始，运行其策略直至达到固定预算，并记录 RGB 图像、深度生成的点云和相机位姿。数据收集完成后，丢弃规划器状态，并使用所有方法相同的 3DGS 实现和超参数，从收集到的视图开始从头训练一个原始 3DGS 模型，不采用针对特定方法的正则化或增强。这样便将规划与重建解耦，因此 Replica-Dense 上的差异可以归因于视角选择。

每种方法均在两种预算设置下进行评估：300 秒的时间预算和 300 个关键帧的视图预算。在每种设置中，我们以 $\{60, 120, 180, 240, 300\}$（秒或视图）为间隔对采集到的轨迹进行子采样，从而为每个场景生成 2×5 个采集集。对于每个采集集，我们训练一个原版 3D高斯泼溅模型，并渲染所有 Replica-Dense 测试视图。我们报告 PSNR、SSIM[43], 和学习感知图像块相似度[44] 以及相应的采集成本，同时记录每种方法在

方法	指标	Hotel0	Off0	Off1	Off2	Off3	Off4	Room0	Room1	Room2
ActiveGAMER	SSIM	0.8831	0.8921	0.8816	0.8735	0.8752	0.8726	0.8796	0.8807	0.8796
	PSNR	23.63	26.93	31.32	22.79	24.61	25.20	24.42	24.42	23.26
	学习感知图像块相似度	0.2102	0.1639	0.1851	0.2250	0.2226	0.2122	0.2389	0.2428	0.2313
Active-GS	SSIM	0.9085	0.9314	0.8782	0.9142	0		.9179	0.9037	0.9139	0.9298
	PSNR	26.36	31.65	29.65	27.06	29.57	28.92	26.68	27.64	27.40
	学习感知图像块相似度	0.1854	0.1244	0.1925	0.1		3	0.1680	0.2173	0.2080	0.1752
Ours	SSIM	0.9084	**0.9532**	**0.9263**	0.9243	0.9054	**0.9282**	0.9200	0.9227	**0.9357**
	PSNR	26.72	**34.69**	36.01	29.93	27.69	**30.95**	29.60	29.84	29.69
	学习感知					0.1627	0			
Ours-Pano	SSIM	**0.9119**	0.9510	0.9248	**0.9277**	**0.9240**	0.9239	**0.9251**	**0.9249**	0.9345
	PSNR	**27.18**	34.34	**36.50**	**30.70**	**31.26**	30.77	**30.41**	**30.39**	**29.92**
	学习感知									

表2：固定视图预算下 Replica-Dense 各场景的结果。峰值信噪比/结构相似性指数越高越好，学习感知图像块相似度越低越好。最佳值以彩色粗体标出，次佳值用下划线标出。

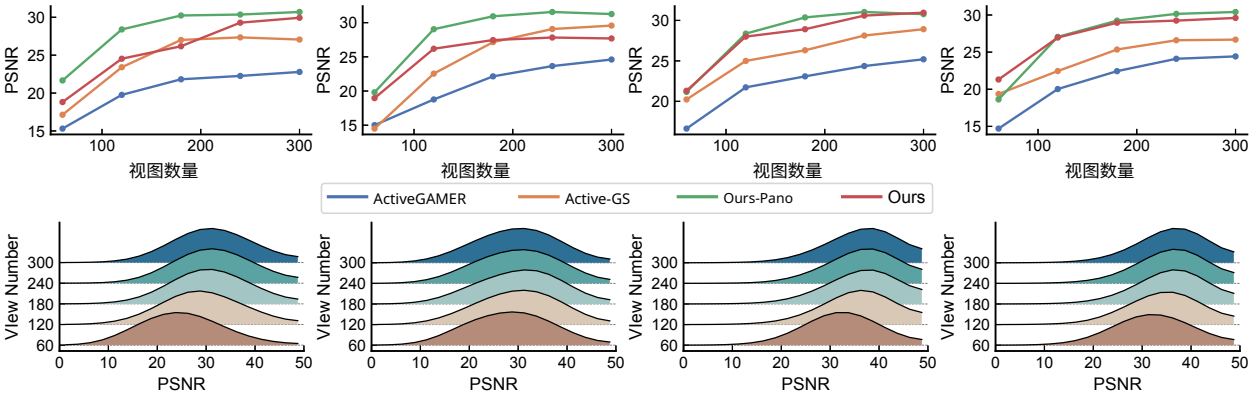


图7：视图预算下的定量结果。上排展示四个场景（Office2、3、4 和 Room0）中峰值信噪比随视图预算的变化。下排展示随着视图预算增加，不同方法（ActiveGAMER、Active-GS、Ours-Pano 和我们的方法）在 Office1 上的峰值信噪比分布变化。

每个场景中的 GPU 峰值内存用量，以量化质量与资源消耗之间的权衡。

图 7总结了不同预算下的性能：在整个预算范围内，我们的方法和 Ours-Pano 的表现始终优于两种基线方法，这反映出它们在低预算下具有更高的样本效率，在高预算下产生的冗余视图更少。峰值信噪比分布也明显更加集中，与基线方法相比，低质量尾部有所缩短。图8 中的定性结果进一步印证了这些趋势：纹理更清晰、边缘更干净，漂浮伪影和重影更少；Ours-Pano 还进一步提升了不同场景下的质量稳定性。



图8：固定视点预算下的定性比较。新视角显示，我们的方法比先前方法生成的细节更清晰、伪影更少；全景变体则在所有场景中带来更稳定的重建质量。红框标出了差异较细微的区域。

4.3 预算约束下的渲染质量

在探索时间固定为 300 秒的条件下，表 1 显示，我们的方法和 Ours-Pano 在几乎所有场景中均取得最佳或次佳的峰值信噪比/结构相似性指数，并获得最低的学习感知图像块相似度；与 Active-GS 相比，峰值信噪比提升约 3–5 dB（相较 ActiveGAMER 的优势更大）。图 6 进一步表明，在 Hotel0 和 Office0 场景中，我们的方法的峰值信噪比上升更快且饱和值更高；在相同的实际时间预算下，我们还收集到更多关键帧，这表明基于可渲染性的评分提供的信息更多，且评估成本更低。

在固定的 300 个关键帧预算下，所有方法均使用相同数量的图像进行训练，因此质量完全取决于视点选择。在这一设置下，表 2显示，我们的方法在峰值信噪比/结构相似性指数和学习感知图像块相似度方面始终优于 Active-GS 和 ActiveGAMER，验证了基于可渲染性的信息增益的有效性。Ours-Pano 通常还能进一步提升表现；而在 Office0 这类较小场景中，由于覆盖范围已趋于饱和，其表现与基础版本相当。

4.4 在线性能分析

图 9 展示了四个场景中峰值 GPU 内存随视图数量的变化。Active-GS 的增长最为显著，因为它会累积大量辐射场；而我们的方法和 Ours-Pano 在所有场景中的内存占用均低于 300MB，因为规划器只维护轻量级的体素统计图。

如图 10所示，逐帧可渲染性更新/查询均处于毫秒范围内，并且随着关键帧增加到数百帧仍几乎保持不变，体现了对观测数量的

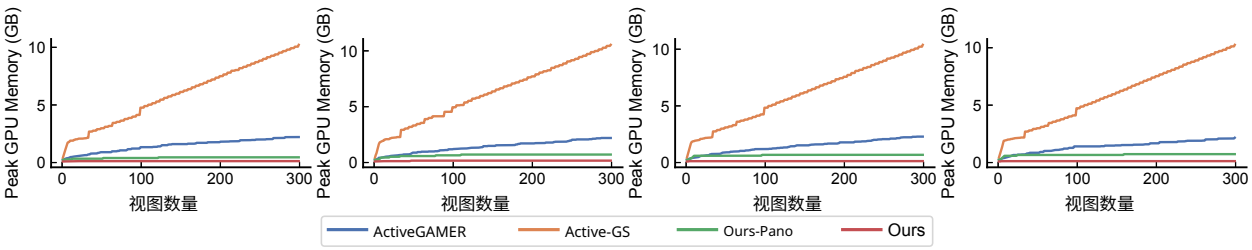


图9：四个场景（Office2、3、4 和 Room0）的 GPU 内存使用情况。每个图表展示了随着采集视图数量增加，不同方法的峰值GPU内存消耗。

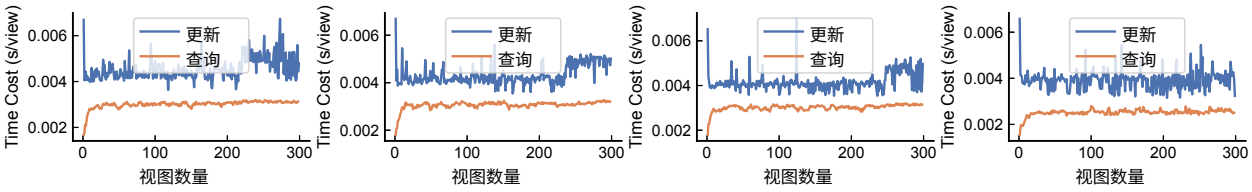


图10：Ours-Pano 的时间成本。该图展示了在 Office2、Office3、Office4 和 Room0 中，随着关键帧数量增加，可渲染性更新和查询的时间成本变化。

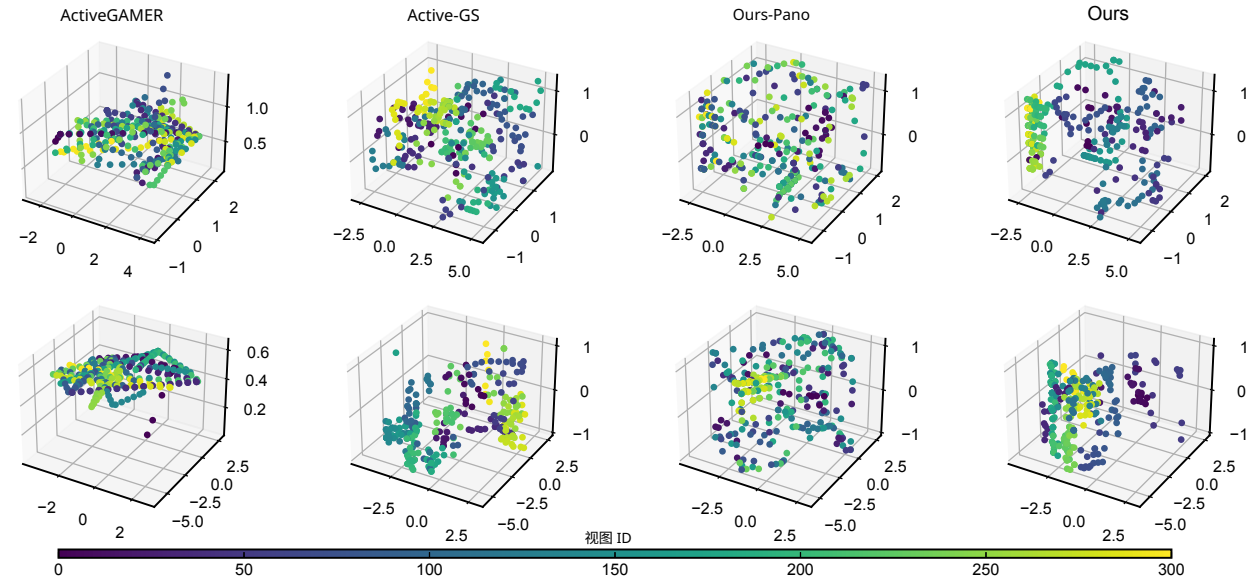


图11：我们的方法的退化分析。我们将四种方法的相机视点绘制为按采集ID着色的三维热力图。顶行展示的是 Hotel0 中的状态；而在 Office3 中，我们的方法退化到一个局部区域，视点反复出现且存在退化。

独立可扩展性，这对于长时域探索至关重要。图 11 可视化了相机分布。与 Hotel0 中的性能相比，我们的非全景变体在随机局部采样下可能由于可渲染性增益与路径成本之间的权衡而退化为局部重复的视点。通过在每个位姿处评估全向效用，Ours-Pano 缓解了这种局部陷阱，实现了更均匀的覆盖范围，从而在不同场景和预算下带来更稳定的性能。

5 结论与局限性

我们提出了 $\mathbb{R}^3$–RECON，这是一种无辐射场主动重建框架，利用观测统计驱动的可渲染性场为候选视图评分。我们将可渲染性分解为方向偏差、外观噪声和分辨率，并通过轻量级体素统计图在线计算这些项，从而无需维护在线辐射场，即可闭式计算 NBV，耗时仅为毫秒级。全景姿态级扩展通过汇总每个位姿的全向效用，加速了视角方向选择。在 Replica-Dense 基准上，在时间和视图预算相同的条件下，R^3-RECON 的 3DGS 重建质量始终优于近期的主动高斯泼溅基线，同时占用更少的 GPU 内存和计算资源，因此有望成为一个即插即用的模块，与现有同步定位与建图系统配合，用于真实场景中的自主数据采集。

目前的一个局限是，候选视点在局部随机采样，在大型或杂乱环境中可能并非最优；未来的工作将把基于可渲染性的评分与学习式或长时域规划策略相结合，以生成信息量更高的候选视点。

参考文献

[1] Bernhard Kerbl、Georgios Kopanas、Thomas Leimk¨uhler 和 George Drettakis。用于实时辐射场渲染的 3D 高斯泼溅。*ACM* 图形学汇刊，42(4):139–1，2023年。[2] Simon Niedermayr、Josef Stumpfegger 和 R¨udiger Westermann。用于加速新视角合成的压缩 3D 高斯泼溅。载于*IEEE/CVF* 计算机视觉与模式识别会议论文集，第 10349–10358 页，2024年。[3] Bashar Alsadik、Hussein Alwan Mahdi 和 Nagham Amer Abdulateef。三维重建下一最佳视角策略的结构化综述与分类。*ISPRS* 摄影测量与遥感开放期刊，第 100098 页，2025年。[4] Xiao Chen、Quanyi Li、Tai Wang、Tianfan Xue 和 Jiangmiao Pang。GenNBV：用于主动三维重建的可泛化下一最佳视角策略。载于*IEEE/CVF* 计算机视觉与模式识别会议论文集，第 16436–16445 页，2024年。[5] Shangjie Xue、Jesse Dill、Pranay Mathur、Frank Dellaert、Panagiotis Tsiotra 和 Danfei Xu。用于不确定性驱动主动建图的神经可见性场。载于*IEEE/CVF* 计算机视觉与模式识别会议论文集，第 18122–18132 页，2024年。[6] Wen Jiang、Boshu Lei 和 Kostas Daniilidis。FisherRF：利用 Fisher 信息进行辐射场主动视角选择与建图。载于欧洲计算机视觉会议（*ECCV*），第 422–440 页，2024年。doi：10.1007/978-3-031-72624-8 24。[7] Yuhan Xie、Yixi Cai、Yinqiang Zhang、Lei Yang 和 Jia Pan。GauSS-MI：用于主动三维重建的高斯泼溅香农互信息。载于机器人学：科学与系统论文集，美国加利福尼亚州洛杉矶，2025年6月。doi：10.15607/RSS.2025.XXI.030。[8] Liren Jin、Xingguang Zhong、Yue Pan、Jens Behley、Cyrill Stachniss 和 Marija Popovi´c。ActiveGS：利用高斯泼溅进行主动场景重建。*IEEE* 机器人学与自动化快报，2025年。[9] Alex Hanson、Allen Tu、Vasu Singla、Mayuka Jayawardhana、Matthias Zwicker 和 Tom Goldstein。PUP 3D-GS：用于 3D 高斯泼溅的原理化不确定性剪枝。载于计算机视觉与模式识别会议论文集，第 5949–5958 页，2025年。[10] Julio A Placed、Jared Strader、Henry Carrillo、Nikolay Atanasov、Vadim Indelman、Luca Carlone 和 Jos´e A Castellanos。主动同步定位与建图综述：研究现状与新前沿。*IEEE* 机器人学汇刊，39(3):1686–1705，2023年。[11] Brian Yamauchi。一种用于自主探索的基于前沿的方法。载于*1997*年 *IEEE* 机器人与自动化计算智能国际研讨会论文集（*CIRA'97*）：“迈向机器人与自动化的新计算原理”，第 146–151 页。IEEE，1997年。[12] Cyrill Stachniss、Giorgio Grisetti 和 Wolfram Burgard。使用 Rao-Blackwell化粒子滤波器进行基于信息增益的探索。载于机器人学：科学与系统，第 2 卷，第 65–72 页，2005年。

[13] Frederic Bourgault、Alexei A Makarenko、Stefan B Williams、Ben Grocholsky 和 Hugh F Durrant-Whyte。基于信息的自适应机器人探索。载于*IEEE/RSJ*智能机器人与系统国际会议，第1卷，第540–545页。IEEE，2002年。[14] Stefan Isler、Reza Sabzevari、Jeffrey Delmerico 和 Davide Scaramuzza。主动式三维体积重建的信息增益表述。载于*2016*年*IEEE*机器人与自动化国际会议（*ICRA*），第3477–3484页。IEEE，2016年。[15] Armin Hornung、Kai M Wurm、Maren Bennewitz、Cyrill Stachniss 和 Wolfram Burgard。Octomap：一种基于八叉树的高效概率三维建图框架。《*Autonomous Robots*》，34(3)：189–206，2013年。[16] Ben Mildenhall、Pratul P Srinivasan、Matthew Tancik、Jonathan T Barron、Ravi Ramamoor- thi 和 Ren Ng。NeRF：将场景表示为用于视图合成的神经辐射场。《*ACM*通讯》，65(1)：99–106，2021年。[17] Edgar Sucar、Shikun Liu、Joseph Ortiz 和 Andrew J Davison。imap：实时隐式建图与定位。载于*IEEE/CVF*国际计算机视觉会议论文集，第6229–6238页，2021年。[18] Zihan Zhu、Songyou Peng、Viktor Larsson、Weiwei Xu、Hujun Bao、Zhaopeng Cui、Martin R Oswald 和 Marc Pollefeys。NICE-SLAM：用于同步定位与建图的神经隐式可扩展编码。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第12786–12796页，2022年。[19] Mohammad Mahdi Johari、Camilla Carta 和 Franc¸ois Fleuret。ESLAM：一种基于有符号距离场混合表示的高效稠密同步定位与建图系统。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第17408–17419页，2023年。[20] Hengyi Wang、Jingwen Wang 和 Lourdes Agapito。Co-SLAM：用于神经实时同步定位与建图的联合坐标与稀疏参数编码。载于*IEEE/CVF*计算机视觉与模式识别会议论文集，第13293–13302页，2023年。[21] Antoni Rosinol、John J Leonard 和 Luca Carlone。NeRF-SLAM：基于神经辐射场的实时稠密单目同步定位与建图。载于*2023*年*IEEE/RSJ*智能机器人与系统国际会议（*IROS*），第3437–3444页。IEEE，2023年。[22] Nikhil Keetha、Jay Karhade、Krishna Murthy Jatavallabhula、Gengshan Yang、Sebastian Scherer、Deva Ramanan 和 Jonathon Luiten。SplaTAM：用于稠密RGB-D同步定位与建图的三维高斯泼溅跟踪与建图。载于*IEEE/CVF*计算机视觉与模式识别会议，第21357–21366页，2024年。[23] Hidenobu Matsuki、Riku Murai、Paul HJ 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。[24] Jianhao Zheng、Zihan Zhu、Valentin Bieri、Marc Pollefeys、Songyou Peng 和 Iro Armeni。WildGS-SLAM：动态环境中的单目高斯泼溅同步定位与建图。载于计算机视觉与模式识别会议论文集，第11461–11471页，2025年。

[25] Yongxin Su、Lin Chen、Kaiting Zhang、Zhongliang Zhao、Chenfeng Hou 和 Ziping Yu。Gaus-slam：基于高斯表面元素的稠密RGB-D同步定位与建图。*arXiv*预印本 *arXiv:2505.01934*，2025 年。[26] Yue Pan 、 Xingguang Zhong 、 Liren Jin 、 Louis Wiesmann、Marija Popovic、Jens Behley 和 Cyrill Stachniss。Pings：高斯泼溅与距离场在基于点的隐式神经地图中的结合。2025年6月。[27] Xuran Pan、Zihang Lai、Shiji Song 和 Gao Huang。ActiveNeRF：通过不确定性估计学习在哪里观察。载于欧洲计算机视觉会议，第230–246页。施普林格，2022年。[28] Ziyue Feng、Huangying Zhan、Zheng Chen、Qingan Yan、Xiangyu Xu、Changjiang Cai、Bing Li、Qilun Zhu 和 Yi Xu。Naruto：基于不确定目标观测的神经主动重建。载于*IEEE/CVF*计算机视觉与模式识别会议，第21572–21583页，2024年。[29] Liyan Chen、Huangying Zhan、Kevin Chen、Xiangyu Xu、Qingan Yan、Changjiang Cai 和 Yi Xu。ActiveGAMER：通过高效渲染实现主动高斯建图。载于计算机视觉与模式识别会议论文集，第16486–16497页，2025年。[30] Yan Li、Yingzhao Li 和 Gim Hee Lee。Active3D：通过分层不确定性量化实现高保真三维重建。*arXiv*预印本 *arXiv:2511.20050*，2025年。[31] Yuetao Li、Zijia Kuang、Ting Li、Qun Hao、Zike Yan、Guyue Zhou 和 Shaohui Zhang。ActiveSplat：通过主动高斯泼溅实现高保真场景重建。*IEEE* 机器人学与自动化快报，2025年。[32] Wen Jiang、Boshu Lei、Katrina Ashton 和 Kostas Daniilidis。AG-SLAM：主动高斯泼溅同步定位与建图。2024年。[33] Wen Jiang、Boshu Lei、Katrina Ashton 和 Kostas Daniilidis。多模态大语言模型引导的探索与基于费 Fisher 信息的主动建图。载于*IEEE/CVF*国际计算机视觉会议，第5392–5404页，2025年。[34] Zirui Wang、Yash Bhalgat、Ruining Li 和 Victor Adrian Prisacariu。主动视点选择器：通过交叉参考图像质量评估实现快速、准确的主动视点选择。*arXiv*预印本 *arXiv:2506.19844*，2025年。[35] Zimu Yi、Ke Xie、Jiahui Lyu、Minglun Gong 和 Hui Huang。在哪里渲染：研究大规模场景的基于图像的渲染的可渲染性。载于*2023*年*IEEE*虚拟现实与三维用户界面会议（*VR*），第356–366页，2023年。[36] Xiaofeng Jin、Yan Fang、Matteo Frosi、Jianfei Ge、Jiangjian Xiao 和 Matteo Matteucci。随处渲染所见：由可渲染性场引导的高斯泼溅。载于*2025*年*IEEE/RSJ*智能机器人与系统国际会议，第7850–7857页，2025年。[37] Marcus Klasson、Riccardo Mereu、Juho Kannala 和 Arno Solin。三维场景重建中的不确定性来源。载于欧洲计算机视觉会议，第271–289页，2024年。

[38] Holger Rauhut 和 Rachel Ward。《球谐展开的稀疏恢复》。*arXiv* 预印本 *arXiv:1102.4097*，2011年。[39] Alvaro Gonz´ález。《使用斐波那契格点和经纬度格点测量球面面积》。数学地球科学，42：49–64，2010年。[40] Suyoung Lee、Jaeyoung Chung、Jaeyoo Huh 和 Kyoung Mu Lee。《ODGS：使用 3D 高斯泼溅从全向图像重建 3D 场景》。神经信息处理系统进展（*NeurIPS*），37，2024 年。[41] Manolis Savva 、Abhishek Kadian、Oleksandr Maksymets、Yili Zhao、Erik Wijmans、Bhavana Jain、Julian Straub、Jia Liu、Vladlen Koltun、Jitendra Malik、Devi Parikh 和 Dhruv Batra。《Habitat：具身人工智能研究平台》。载于*IEEE/CVF*国际计算机视觉会议论文集（*ICCV*），第9338–9346页，2019年。[42] Julian Straub、Thomas Whelan、Lingni Ma、Yufan Chen、Erik Wijmans、Simon Green、Jakob J. Engel、Raul Mur-Artal、Carl Yuheng Ren、Shobhit Verma、Anton Clarkson、Mingfei Yan、Brian Budge、Yajie Yan、Xiaqing Pan、June Yon、Yuyang Zou、Kimberly Leon、Nigel Carter、Jesus Briales、Tyler Gillingham、Elias Mueggler、Luis Pesqueira、Manolis Savva、Dhruv Batra、Hauke M. Strasdat、Renzo De Nardi、Michael Goesele、Steven Lovegrove 和 Richard A. Newcombe。《Replica 数据集：室内空间的数字复刻》。*CoRR*，abs/1906.05797，2019年。[43] Zhou Wang、Alan C. Bovik、Hamid R. Sheikh 和 Eero P. Simoncelli。《图像质量评价：从误差可见性到结构相似性》。*IEEE*图像处理汇刊，13(4)：600–612，2004年。[44] Richard Zhang、Phillip Isola、Alexei A. Efros、Eli Shechtman 和 Oliver Wang。《深度特征作为感知度量的惊人有效性》。载于*IEEE*计算机视觉与模式识别会议论文集（*CVPR*），第586–595页，2018年。