G² S-ICP SLAM: □何感知高斯□点 ICP SLAM

Gyuhyeon Pak¹, Hae Min Cho² and Euntai Kim^{1,*}

Abstract— 在本文中,我□提出了一□新□的□何感知RGB-D 高斯散点SLAM系□,名□G²S-ICP SLAM,所提出的方法通 □使用□束于局部切平面的高斯分布□表示每□□景元素,□□了高保□ 3D重建和□□的□棒相机姿□□踪□正方法有效地□局部表面建模□□ 基□□何□□的二□高斯□□,□□□的具有各向同性不□定性的3D□球□表 示相比,在多□□角下提供了更一致的深度解□。□了□□□表示整合到 SLAM管道中,我□通□引入各向□性□方差先□□表面回的高斯□□ 嵌入到□□ICP□架中,而不改□基□的配准公式。此外,我□提出了 一□□何感知□失□□督光度、一致性和法□一致性。我□的系□在保持 □□和□何保□度的同□,□□了□□操作。□Replica和TUM-RGBD □据集的大量□□表明,G²S-ICP SLAM在定位精度和重建完 整性方面□于□有的SLAM系□,同□保持了渲染□量。

Index Terms—Simultaneous Localization and Mapping (SLAM), 2D Gaussian Splatting, 3D reconstruction.

I. 介□

□□同□定位□建□(vSLAM)已成□机器人和增強□ □□用的基□技□,□相机□踪和□景重建提供了高效和 可□展的解□方案。□□的SLAM流水□ [1]-[6] 主要依 □于稀疏的□何□□表示,如□□点、地□或表面元素□。 □些方法□然在□算效率和□棒性上表□良好,但往往无 法提供密集的、逼□的重建,□在跨□点的□何一致性上 存在困□。□─限制限制了□□在需要感知上有意□和□ □上准□的地□的□景中的适用性。

基于□射□的表示方法的最新□展,特□是3D高斯□射 (3DGS),展示了□□□□景□行□□、高保□渲染的□力。 □些方法通□□□景点建模□□□高斯□合成□□依□ 的RGB□像,□而□□了令人印象深刻的逼□度。然而, □□的各向同性和□□性□通常□在□不同□点□察□ 引入深度不一致。在SLAM□用中,□□多□□□□何一致 性的缺乏可能□降低姿□精度□□致空□□影。

基于 3DGS 的□□,最近的□ 究探□ 了□ 其整 合到 SLAM管道中。□些努力大致可以分□□□。第一□[7], [8] 主要□注使用外部提供或□□算的位姿□行基于3DGS 的映射。□□策略可以□□高保□度的重建和照片□的渲 染,但□以□□的方式□行,□不有助于位姿□□。相比之 下,第二□[9]-[11] □3DGS整合到□踪中,利用可微分渲 染或基于□化的方法□□□相机位姿。□管有□力,□些 方法通常受制于高□算成本、□光照□件的敏感性,以及 在无□理或重□□境中的有限□棒性。

□些限制揭示了一□□□缺口:□有的3DGS-SLAM系 □□先考□□□□量而非□何□定性,使其在□□定位和 建□方面表□不佳。□了解□□□□□,我□提出了一□

* is that the corresponding author.

¹Gyuhyeon Pak and Euntai Kim are with the Department of Electrical and Electronic Engineering, Yonsei University, Seoul 03722, Republic of Korea { gh.pak, etkim } @yonsei.ac.kr

²Hae Min Cho is with the School of Computing, Gachon University, Sujeong-gu, Seongnam-si, Gyeonggi-do 13120, Republic of Korea hmcho9@gachon.ac.kr



Fig. 1: 在Replica□据集中比□渲染的法□□像□量。每 行□示(a) □□□像,以及由(b) GS-ICP SLAM和(c) G² S-ICP SLAM渲染的法□□像

□何感知的SLAM□架,□□架在通□基于表面□□的2D 高斯建模□保空□一致性和深度精度的同□,保留了GS的 □□表□力 [12]。

□此,G²S-ICP SLAM 引入了三□□□□新。首先,我 □采用□局部表面□何□□的二□高斯□□表示,□解□ 了由□□3D高斯引起的多□□深度不一致□□。其次,我 □通□明□□行□局部切平面□□的各向□性□方差□ □,□□表面□□的二□高斯□□□入□□ICP (GICP) [13] □程中。□然底□配准算法保持不□,但我□的公式 通□尺度正□化引入二□先□,引□□化□向表面一致的 □□。第三,我□引入了一□□何感知□化□架,□光度、 深度和基于法□的□何感知□失□合在一起,□采用一□ 保持映射期□表面各向□性的尺度正□化策略。

□些□□使我□的□架能□在保持□何一致性和照片 □逼□渲染□量的同□□□□□性能。在TUM-RGBD和 Replica□据集上的□□表明,G²S-ICP SLAM在定位精 度和深度保□度方面□于之前的3DGS SLAM系□。

基于□射□的SLAM方法,包括NeRF-SLAM和基于GS的SLAM,旨在通□□□式或□式渲染集成到SLAM流水□中□重建照片□逼□的□景。NeRF-SLAM利用□式□□□□建模外□,而基于GS的方法采用□式的3D高斯点用于高效的渲染和地□融合。最近的□究如SplaTAM、MonoGS和GS-SLAM表明,3D高斯点□(3DGS)在□□ 保□度和□存效率上□于□□的稠密或稀疏地□表示,使其非常适合用于SLAM□用。

□管有□些□点,基于3DGS的SLAM方法由于其□□ 性和各向同性特性,□常□致深度不一致性。具□而言,3D □球□中沿法□□的无□束支撑在多□□□置下□□致□ 何失准。MonoGS [10] 通□在□目SLAM中引入□何正□ 化解□了其中的一些□□,而SplaTAM□引入了自适□地 □□化以提高空□一致性。同□,Photo-SLAM [7] 和GS-ICP SLAM [8] 用□典□□里程□ [2] 或 ICP□踪 [13], [14] 替□了基于光度的位姿□□,□而□□了更快更□定 的操作。然而,□些方法仍然□承了各向同性3D表示的局 限性。相□之下,我□的方法引入了一□□表面□□的2D 高斯公式和□何感知□失□解□□一根本□□。

最近在基于□射□映射方面的□展,例如3D高斯点□ (3DGS)[15],使得高度□□和□□感的□景重建成□可 能。然而,大多□□些方法使用各向同性或□□表示—— 如3D高斯□球□——□□缺乏□式的表面□□,常常□ 致□□□的深度不一致和□□□影。□一限制□致了后□ 工作的出□,如2DGS[12]、SuGaR[16]和PGSR[17] ,□□引入了表面法□□督或切平□束,以提高神□渲染 中的□何精□性。

□此同□,□何感知□□□架已□表明,□合表面法□ 作□□督信□可以□著提高重建□量。然而,□些□解□ 未完全整合到SLAM□向的系□中。我□的方法通□提出 一□表面□□的二□高斯表示□□□□□一空白,此表示 □束每□高斯的支持范□在局部切平面上,□感□源于 2DGS [12]。此外,我□引入了一□□何感知□失,利用 □□表面法□□指□渲染□程。□合□些□件,我□的系 □能□在不□牲□行□□或姿□□踪性能的情□下保留 精□的□何□□,□□更好的重建保□度。

II. 方法□

本□□述了G²S-ICP SLAM的核心□件,□是一□□ 何感知SLAM□架,旨在保持□何一致性,同□保留逼□ 的渲染□量。提出的□架包含四□主要□件:(1)表面□□ 的2D高斯表示,(2)□有表面□□先□的□□ICP (GICP) □踪,(3)表面□□尺度正□化和(4)□何感知□化。□些 □件□合了稠密逼□渲染□空□一致的SLAM之□的差 距。整□系□的□述如□2所示,□□的□□方式□在后 □的小□中介□。

□管□有的3DGS SLAM方法采用3D高斯□球□□表 示□景,但□□的□□特性在多□角□置下固有地引入了 深度不一致。因□3D□球□在三□空□□上分布不□定 性,□□缺乏□其□局部表面□何□□的明□□束。因此, 同一点在不同□角下可能□投影不一致,□致□何失□。 □一□□在□3中有□明,其中同□的高斯□球□□□色 和□色□角□察。每□□角□□球□相交于不同的切片平 面,□致□角□深度□□不一致,□在重建中引□深度□ 影。

□了解□□□□□,我□的方法采用了一□基于二□高 斯□□[12]的表面□□表示,□□表示□每□高斯的空□ 支持限制在局部切平面上。□□公式□少了□□依□的□ □性,改善了□□的深度一致性,□而□□更□定的□踪 和□何感知的SLAM□化。

每□二□高斯□由中心位置 $p_k \in \mathbb{R}^3$ 、□方差矩□ $C_k \in \mathbb{R}^{3\times 3}$ 、□色 c_k 和不透明度 σ_k 定□。□方差□算□:

$$\boldsymbol{C}_k = \boldsymbol{R}_k \boldsymbol{S}_k \boldsymbol{S}_k^T \boldsymbol{R}_k^T, \qquad (1)$$

,其中 $\mathbf{R}_k = [\mathbf{t}_1, \mathbf{t}_2, \mathbf{n}_k] \in \mathbb{R}^{3 \times 3}$ 是由□□正交切向量和表 面法□□成的旋□矩□, $\mathbf{S}_k = \text{diag}(\mathbf{s}_1, \mathbf{s}_2, \mathbf{0}) \in \mathbb{R}^{3 \times 3}$ 是□ 切平面□□的□放矩□。 □定在第 k □高斯□□的局部切平面上的一□点 p(u,v)

$$p(u,v) = p_k + s_1 t_1 u + s_2 t_2 v = H(u,v,1,1), \quad (2)$$

,其中

:

$$\boldsymbol{H} = \begin{bmatrix} s_1 t_1 & s_2 t_2 & 0 & p_k \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{RS} & p_k \\ \boldsymbol{0} & 1 \end{bmatrix} .$$
(3)

矩□ $H_k \in \mathbb{R}^{4\times4}$ 定□了一□局部表面□□□□],□2D□ □坐□映射到3D全局坐□系中。□于2D高斯□□上的点 q(u,v),公式□:

$$\mathcal{G}(\boldsymbol{q}) = \exp\left[\frac{u^2 + v^2}{2}\right] \tag{4}$$

我□在二□高斯□□表示中遵循一□3DGS光□化□ 程。□色和深度渲染如下:

$$\boldsymbol{c}(\boldsymbol{x}) = \sum_{i=1}^{n} \boldsymbol{c}_i \alpha_i \mathcal{G}_i(\boldsymbol{x}) \prod_{j=1}^{i-1} (1 - \alpha_j \mathcal{G}_j(\boldsymbol{x})), \quad (5)$$

$$d(\boldsymbol{x}) = \sum_{i=1}^{i} d_i \alpha_i \mathcal{G}_i(\boldsymbol{x}) \prod_{j=1}^{i-1} (1 - \alpha \mathcal{G}_j(\boldsymbol{x})).$$
(6)

□□表示本身就捕捉到了表面□何形□,使模型在□理多 □□□中的深度不一致□更加□健。

在□踪□程中,我□采用基于□□ICP(GICP)配准 算法的□□模型方法。□定第 k □的 RGB-D□像□ (I_k, D_k),我□□深度□ D_k 中提取一□3D点云 $\mathcal{X}_k =$ { x_m }_{m=1,...,N_k},其中 N_k 是有效深度像素的□量。□于每 □点 x_i ,我□通□分析其□近点的空□分布□算一□局部 □方差矩□ $C_m \in \mathbb{R}^{3\times 3}$ 。利用□一点,我□□□ k 定□高 斯集□: 令 $G_k^{src} = \{\mathcal{X}_k^{src}, \mathcal{\Sigma}_k^{src}\}$ □□□前□提取的高斯 集, $G_k^{tgt} = \{\mathcal{X}_k^{tgt}, \mathcal{\Sigma}_k^{tgt}\}$ □□地□中□索到的□□目□高 斯集。通□最小化□□高斯之□的Mahalanobis距□□□ 相□□□ $T_k \in SE(3)$ 。

每□点 \boldsymbol{x}_m^{src} 被建模□一□高斯分布 $\mathcal{N}(\hat{\boldsymbol{x}}_m^{src}, \boldsymbol{C}_m^{src})$,其 □□的目□□ $\mathcal{N}(\hat{\boldsymbol{x}}_m^{tgt}, \boldsymbol{C}_m^{tgt})$ 。源点和目□点之□的□差定 □□:

$$d_m = \boldsymbol{x}_m^{tgt} - \boldsymbol{T}_k \boldsymbol{x}_m^{src}.$$
 (7)

假□□差服□高斯分布,我□有:

$$d_m \sim \mathcal{N}(\hat{d}_m, \boldsymbol{C}_m^{tgt} + \boldsymbol{T}_k \boldsymbol{C}_m^{src} (\boldsymbol{T}_k)^T)$$
(8)

$$= \mathcal{N}(\hat{x}_m^{tgt} - \boldsymbol{T}_k \hat{x}_m^{src}, \boldsymbol{C}_m^{tgt} + \boldsymbol{T}_k \boldsymbol{C}_m^{src} (\boldsymbol{T}_k)^T).$$
(9)

□了□得最佳的□□ T_k^* ,我□在所有点□□上□用最 大似然□□:

$$\boldsymbol{T}_{k}^{*} = \operatorname*{argmax}_{\boldsymbol{T}_{k}} \prod_{m=1}^{N} p(d_{m}) = \operatorname*{argmax}_{\boldsymbol{T}_{k}} \sum_{m=1}^{N} \log p(d_{m}) \quad (10)$$
$$= \operatorname*{argmax}_{\boldsymbol{T}_{k}} \sum_{m=1}^{N} d_{m}^{T} (\boldsymbol{C}_{m}^{tgt} + \boldsymbol{T}_{k} \boldsymbol{C}_{m}^{src} (\boldsymbol{T}_{k})^{T})^{-1} d_{m}. \quad (11)$$

□了□表面□何□入到□□配准□程中,我□□第??□ 中描述的各向□性二□□方差□□嵌入到每□*C_m^{src}*中。 通□在局部切平面上施加□束──沿表面法□方向的方 差□零──我□的方法在GICP□踪中引入了□式的二□ □束。□□□表面□□的先□□姿□□□偏向于□何一致 的□□,特□是在□□或不明□的□□下。



Fig. 2: □□流程。G²S-ICP SLAM 使用RGB和深度□像作□□入,以□□□建□何一致的3D地□。□了增強多□□ 深度一致性,我□□每□点表示□一□表面□□的二□高斯□□,而不是□□□□球。在□踪□段,□□表示被集成到基 于二□高斯的□□ICP (GICP)中,以□□相机□□。在映射□程中,我□使用□何感知□失□化高斯基元,□些□失 施加光度、一致深度和法□一致性,□而防止方向□位□保持表面□□。



Fig. 3: 多□角□置中三□高斯□球的深度不一致。□□角 相□的三□高斯□球□□致同一□表面点在多□□角中 由不同的截面平面表示,□而□致深度□□不一致和□何 不□□。

如在第??□□□的,我□的□架□每□高斯基元表示 □□局部曲面切平面□□的二□□□。此公式要求□空□ 不□定性限制在□平面□,消除沿曲面法□方向的方差。 □了□足□一□何要求,我□明□□用于□踪和映射的所 有高斯基元的尺度矩□ *S* = *diag*(*s*₁, *s*₂, *s*₃)的第三□分 量□□零。此配置□保每□高斯在法□□上具有零不□定 性,□而強制□行平坦的二□表示。

□之前依□完整3D高斯□球的方法不同,我□的系□在 □建高斯□直接施加□□□□□束,避免了□外正□化□ 或□化□□。□□□保□了□踪模□和地□表示之□的一 致□何行□,□□共享同一□2D高斯。

在我□的映射模□中,全局□景是以□□□□中心逐步 □建的。根据不同的□准定□了□□□型的□□□:□踪 □□□和映射□□□。□踪□□□是基于□GICP□踪□ 程中□得的□何□□比例□□的,而映射□□□□在固定 的□□隔□入,以□保足□的空□覆盖。

由于地□是□深度□像□建的,因此高斯的空□密度由于量程□感器的投影特性本□上是不均□的——在□□的距□上,每□像素代表更大的空□□域。□了□□逼□

的、感知一致的高斯表示,我□根据其□□感器的距□□ 整每□二□高斯□□的初始尺度:

$$s_1, s_2 \propto \frac{1}{z^p},\tag{12}$$

,其中 z 是深度□,p 是一□□□□□的超□□,用于控制 □ h 速率。□□根据距□的□ h 方法可以提高□□上的均 □性,□增強姿□□□的□定性。

在通□2D GICP□踪□行姿□□□和使用具有尺度感 知的高斯初始化□行地□□建之后,我□□行□何感知□ 化,以□化高斯□量,以提高渲染和□何一致性。我□的□ 化□注于更新2D高斯的□性,以更好地描述□入□□。

我□的□□失函□包含三□互□的□:光度相似性、深 度、□何感知法□。□光度和深度□失有助于保持□□□ RGB-D□□的□□相似性□,我□的□化核心在于使用表 面法□作□□何□督信□。

我口口三口点的口口深度空口梯度口出表面法口 $N_{GT}(x,y)$:

$$N_{GT}(x,y) = \frac{\nabla_x x_s \times \nabla_y x_s}{\|\nabla_x x_s \times \nabla_y x_s\|},$$
(13)

,其中 x_s 表示□深度□像生成的三□位置。 渲染的深度到 表面法□ \hat{N}_d 同□是□渲染的深度 \hat{D} □算出□的。 渲染的 法□ \hat{N} 是通□累□的高斯公式□算的。 然后我□制定了 一□□何感知法□(GAN)□失,□□□N、 \hat{N} 和 \hat{N}_d 之□ 的向量差□和角度失配:

$$\mathcal{L}_{GAN} = \|N_{GT} - \hat{N}\|_{1} + \left[1 - \frac{N_{GT} \cdot \hat{N}}{\|N_{GT}\| \| \hat{N}\|}\right] + \left[1 - \frac{N_{GT} \cdot \hat{N}_{d}}{\|N_{GT}\| \| \hat{N}_{d}\|}\right].$$
 (14)

。□□基于法□的□督在□点之□施加強□粒度的表面□ □,□□充附加□失:

$$\mathcal{L}_p = \|I - \hat{I}\|_1, \quad \mathcal{L}_d = \|D - \hat{D}\|_1, \quad (15)$$

,□支持外□保□度和深度回□。完整的目□函□□:

$$\mathcal{L} = \lambda_p \mathcal{L}_p + \lambda_d \mathcal{L}_d. + \lambda_{GAN} \mathcal{L}_{GAN}$$
(16)

TABLE I: Replica上的3D重建口口	我口的方法在重建地口的口量	F 口王所有其他口口宫斯卢云 SLAM 口架
TADLE I. Replica上的D里廷口口。	我口时刀伍住里廷地口时口里-	L凵 J // 有共把凵凵同别点厶SLAM凵木。

Methods	Metrics	R0	R1	R2	Of0	Of1	Of2	Of3	Of4	Avg.
GS-SLAM [11] (8.34fps)	Depth L1 [cm]↓ Precision [%]↑ Recall [%]↑ F1-score [%]↑	$ \begin{array}{c c} 1.31 \\ 64.58 \\ 61.29 \\ 62.89 \end{array} $	$0.82 \\ 83.11 \\ 76.83 \\ 79.85$	$1.26 \\ 70.13 \\ 63.84 \\ 66.84$	$\begin{array}{c} 0.81 \\ 83.43 \\ 76.90 \\ 80.03 \end{array}$	$\begin{array}{c} 0.96 \\ 87.77 \\ 76.15 \\ 81.55 \end{array}$	$\begin{array}{c} 1.41 \\ 70.91 \\ 61.63 \\ 65.95 \end{array}$	$\begin{array}{c} 1.53 \\ 63.18 \\ 62.91 \\ 59.17 \end{array}$	$\begin{array}{c} 1.08 \\ 68.88 \\ 61.50 \\ 64.98 \end{array}$	$1.16 \\ 74.00 \\ 67.63 \\ 70.15$
GS-ICP SLAM* (30fps)	Depth L1 [cm]↓ Precision [%]↑ Recall [%]↑ F1-score [%]↑	$\begin{array}{c c} 22.24 \\ 4.77 \\ 15.81 \\ 7.33 \end{array}$	17.88 3.69 15.65 5.97	23.01 3.63 13.83 5.75	$16.76 \\ 4.77 \\ 15.81 \\ 7.33$	$16.23 \\ 5.20 \\ 11.44 \\ 7.15$	23.90 3.54 12.15 5.49	$21.71 \\ 2.50 \\ 8.61 \\ 3.87$	22.59 3.58 10.89 5.39	$20.54 \\ 3.96 \\ 13.02 \\ 6.04$
G ² S-ICP SLAM (30fps)	Depth L1 [cm]↓ Precision [%]↑ Recall [%]↑ F1-score [%]↑	$\begin{array}{c c} 0.61 \\ 85.89 \\ 81.10 \\ 83.43 \end{array}$	$\begin{array}{c} 0.39 \\ 91.47 \\ 84.41 \\ 87.80 \end{array}$	$\begin{array}{c} 0.90 \\ 78.34 \\ 74.24 \\ 76.23 \end{array}$	$\begin{array}{c} 0.34 \\ 90.56 \\ 87.12 \\ 88.81 \end{array}$	$\begin{array}{c} 0.77 \\ 92.35 \\ 83.93 \\ 87.94 \end{array}$	$\begin{array}{c} 0.98 \\ 80.50 \\ 72.29 \\ 76.17 \end{array}$	$1.27 \\76.39 \\71.70 \\73.97$	$\begin{array}{c} 0.66 \\ 82.23 \\ 74.59 \\ 78.22 \end{array}$	$\begin{array}{r} 0.74 \\ 84.72 \\ 78.67 \\ 81.57 \end{array}$

* denotes the reproduced results by running official code.



Fig. 4: Replica□据集的定性□果。(a)、(b)□示了由GS-ICP SLAM和G² S-ICP SLAM重建的room1和office0的□格□。G² S-ICP SLAM□□了干□且准□的3D□格重建。

通□□表面感知□化集成到渲染管道中,我□的方法在 保持照片□□感的同□□保□何合□性,□而□致更加□ 定的SLAM□迹和一致的地□□□。

III.

A. □□□置

我□描述了我□的□□□置,□□G²S-ICP SLAM□ SOTA方法□行比□,以□□我□在□自Replica□据集 [18]的8□合成□景和□自TUM RGB-D□据集 [19]的3 □□景中的性能。

□□□□。所有□□均在一台配置□Intel Core i5-12500 CPU、32 GB RAM和NVIDIA RTX A5000 GPU的台式 机上□行。□了□少□算□源,我□□□失函□的超□□ λ_P 、 λ_D 和 λ_{GAN} 分□□置□1、0.1 和 0.05。我□□□距 □感知□放因子 p□置□ 0.333。□了生成三□□景的□ 格,我□在□□的□迹上渲染□整合每□深度和□色□, □使用□素大小□1 厘米的 TSDF 融合 [20]。

□价指□。□于□格重建,我□使用深度L1 [21],□是用 于□□□景□何的二□指□,精□度、召回率以及以1cm □距□□□的F1分□。□于追踪精度,我□使用□□□迹 □差(ATE)的均方根□差(RMSE)□□□方法 [19]。我 □采用□准的□像□量指□,包括峰□信□比(PSNR)、□ □相似性指□(SSIM) [22],以及□□感知□像□相似性 (LPIPS) [23]□□□渲染。 □据集。Replica [18] 和 TUM-RGBD [19] 是□□□有 NeRF 和基于 GS 的 SLAM 方法最常用的□据集。Replica □据集包含各□室□□景的高□量□□感合成 RGB-D □ 像。TUM-RGBD □据集由于 RGB-D □像□量□低而更 □□□。RGB □像包含□□模糊,深度□像□□大,且有 □多未□充□域。

比□方法。我□□ GS-ICP SLAM [8] 作□ 3D 重建、□ 踪和渲染性能的主要基□。我□□□提出的 G² S-ICP SLAM □其他 SOTA 基于 NeRF 的 SLAM 方法 [24], [25] 以及其他基于 GS 的 SLAM [9]-[11] □行比□。

我□在Replica□据集上□□我□方法的三□重建性能, □果如表 I 所示。我□在□何精度和重建完整性方面□ GS-SLAM [11] 和GS-ICP SLAM [8] □行了比□。由于 我□的□架能□以每秒30□的速度□□□行,我□□GS-ICP SLAM□□我□的主要基准,□□外包括GS-SLAM, 其提供高保□度重建但□行□率□著低于8.34 FPS。

G²S-ICP SLAM在所有□景中都□□了最佳的□何精 度,平均深度L1□差□0.74厘米,□GS-SLAM和GS-ICP SLAM相比,分□提高了36%96% 2D 3D Precision

Recall F1 SLAM

□ 4 中的定性□果□一步支持了我□方法的□越性。 □□常出□□□□影和□□缺失的GS-ICP SLAM相比, G²S-ICP SLAM生成了更完整的□格地□。在room1和 office0□景中,我□的方法忠□地重建了□粒度□何,□保

TABLE II: 在Replica和TUM-RGBD上的□踪性能(ATE RMSE ↓ [cm])。最佳□果分□□□□ 第一 、 秒 和 第三 。我□的方法在相机姿□□□方面□到了最新的性能水平。

	Dataset		Replica							TUM-RGBD				
	Method	R0	R1	R2	Of0	Ōf1	Of2	Of3	Of4	Avg.	fr1/desk	fr2/xyz	fr3/office	Avg.
	NICE-SLAM [24]	0.97	1.31	1.07	0.88	1.00	1.06	1.10	1.13	1.06	2.7	1.8	3.0	2.5
Non	Point-SLAM [25]	0.61	0.41	0.37	0.38	0.48	0.54	0.69	0.72	0.52	4.34	1.31	3.48	3.04
Declations	GS-SLAM [11]	0.48	0.53	0.34	0.52	0.41	0.59	0.46	0.70	0.50	3.3	1.3	6.6	3.7
Real-time	SplaTAM [9]	0.31	0.40	0.29	0.47	0.27	0.29	0.32	0.55	0.36	3.35	1.24	5.16	3.25
	$MonoGS^*$ [10]	0.48	0.36	0.34	0.44	0.52	0.23	0.16	2.53	0.58	1.48	1.45	1.50	1.48
Pool time	GS-ICP SLAM* [8]	0.15	0.16	0.10	0.29	0.12	0.16	0.18	0.20	0.17	3.07	1.79	2.46	2.44
Real-time	G 2 S-ICP SLAM	0.14	0.16	0.10	0.19	0.12	0.16	0.17	0.20	0.15	2.74	1.59	2.78	2.37

* denotes the reproduced results by running official code.

持了表面□□性,展示了在□□室□□境中的□棒性。

B. □踪性能

我□通□□G²S-ICP SLAM□基于NeRF的SLAM [24], [25]、非□□的GS SLAM [9]-[11]和GS-ICP SLAM [8]□行比□,□□□其在Replica和TUM-RGBD□据集 上的□踪性能。如表所示. II 我□的方法在Replica□据集 上始□□到最低的□踪性能,□在TUM-RGBD□据集上 □到了非常有□□力的□踪性能。在Replica□据集上,G²S-ICP SLAM□于所有比□方法,包括作□主要基□的 GS-ICP SLAM□于所有比□方法,包括作□主要基□的 GS-ICP SLAM、基于NeRF的方法以及SOTA GS SLAM 系□。在TUM-RGBD上,我□的方法在整□□景中取得 了第三名的平均□果,表□出在□□世界RGB-D□置中 的□比SOTA性能。

□□強□的性能強□了在□□ICP□踪中□入二□表面 □□束的有效性。通□用□局部表面□何□□的二□高 斯□□替代□□的三□□球□,我□在保留□踪精度的同 □□少了空□模糊。□得注意的是,□管施加了更強的□ 何先□,我□的模型在位姿□□精度方面□未比三□高斯 □球□差。□些□果表明,二□□束不□□保了更好的□ 何一致性,□增強了多□化室□□景中的整□□踪□定性。

C. 渲染性能和系□速度

我□在Replica和TUM-RGBD□据集上□□渲染□量 和系□速度,其□果如表格 III 和表格 IV 所□□。□然先 前的工作□□性地表明,2D Gaussian Splatting [12] 在渲 染□量上通常不如□□3DGS [15],但我□的□果表明, 表面□□的2D高斯建模仍然可以提供具有□□力的性能。 具□而言,G²S-ICP SLAM□□7□GS-ICP SLAM相 □的渲染□量,□在渲染性能上略有下降。由于使用表面 □束的2D高斯,□□□微的退化是一□□期的□衡,因□ □□先考□□何□□而不是依□□□的外□混合。□管存 在□□□衡,我□的方法在Replica□据集上的渲染□量 和效率方面□□7 SplaTAM、MonoGS和GS-SLAM等其 他比□。此外,我□的方法在表格 III 中PSNR排名第二, LPIPS排名第三。

□5□一步展示了我□方法的□□。在room1和office0 □景中,G²S-ICPSLAM相比于GS-ICPSLAM生成了 更□□的□□同和更低的深度□差。定性□果□示出更 □□的□□□界、改□的物□□□以及在深度重建中更少 的缺失□域。尤其是,深度□差□展示了我□的方法更好 地保持了表面完整性,□得益于我□SLAM系□中嵌入的 □何感知建模。

□之,G²S-ICP SLAM 在□□完整性和□何精度之□ □行了深思熟□的折衷,提供快速且一致的3D重建,具有 TABLE III: 在Replica上的系□速度和渲染□量□□。最 佳□果被突出□示□ 第一 、秒 和 第三 。我□的方 法在系□速度和渲染□像□量方面表□出色。

Methods	Metrics	R0	R1	R2	Of0	Of1	Of2	Of3	Of4	Avg.
GS-SLAM [11]	$PSNR[dB] \uparrow$ $SSIM \uparrow$ $LPIPS \downarrow$ $FPS \uparrow$	31.56 0.968 0.094 8.34	32.86 0.973 0.075	32.59 0.971 0.093	38.70 0.986 0.050	41.17 0.993 0.033	32.36 0.978 0.094	32.03 0.970 0.110	32.92 0.968 0.112	34.27 0.975 0.082 8.34
SplaTAM [9]	$PSNR[dB] \uparrow$	32.86	33.89	35.25	38.26	39.17	31.97	29.70	31.81	34.11
	$SSIM \uparrow$	0.97	0.98	0.97	0.98	0.98	0.97	0.94	0.95	0.97
	$LPIPS \downarrow$	0.07	0.10	0.08	0.09	0.09	0.10	0.12	0.15	0.10
	$FPS \uparrow$	0.24	0.19	0.19	0.20	0.22	0.27	0.26	0.24	0.23
MonoGS [*] [10]	$PSNR[dB] \uparrow$	33.45	36.27	37.07	40.41	41.42	35.82	35.54	33.62	36.70
	$SSIM \uparrow$	0.943	0.959	0.965	0.974	0.977	0.964	0.959	0.939	0.960
	$LPIPS \downarrow$	0.070	0.069	0.064	0.052	0.045	0.055	0.054	0.100	0.064
	$FPS \uparrow$	0.59	0.72	0.64	0.76	0.93	0.63	0.65	0.66	0.70
GS-ICP SLAM*	$PSNR[dB] \uparrow$	34.47	36.92	37.37	41.76	42.49	36.01	36.17	38.21	37.92
	$SSIM \uparrow$	0.956	0.966	0.969	0.981	0.981	0.969	0.965	0.969	0.970
	$LPIPS \downarrow$	0.057	0.053	0.059	0.034	0.038	0.052	0.048	0.051	0.049
	$FPS \uparrow$	29.95	29.94	29.94	29.96	29.94	29.95	29.95	29.95	29.95
G ² S-ICP SLAM	$PSNR[dB] \uparrow$	33.40	35.73	36.12	40.62	41.72	35.32	35.14	37.02	36.88
	$SSIM \uparrow$	0.945	0.958	0.964	0.979	0.978	0.961	0.954	0.962	0.963
	$LPIPS \downarrow$	0.077	0.075	0.079	0.045	0.051	0.069	0.071	0.071	0.067
	$FPS \uparrow$	29.94	29.94	29.94	29.94	29.95	29.95	29.94	29.94	29.94

TABLE IV: 在TUM-RGBD上的系□速度和渲染□量□ □。所提出的方法□示了□人的系□速度和具有□□力的 渲染□量。

Methods	$\mathrm{PSNR}[\mathrm{dB}]\uparrow$	SSIM \uparrow	$\mathrm{LPIPS}\downarrow$	$\mathrm{FPS} \uparrow$
Photo-SLAM* [7]	21.14	0.738	0.211	-
$MonoGS^*$ [10]	17.91	0.716	0.311	1.77
GS-ICP SLAM [*] [8] (limited to 30 FPS)	20.42	0.764	0.226	29.97
G 2 S-ICP SLAM (limited to 30 FPS)	20.06	0.758	0.233	29.97

* denotes the reproduced results by running official code.

□□力的渲染性能——□固了2D高斯散点在□□SLAM 系□中的可行性。

D. 消融□究

□了□□我□□架中每□□件的□□,我□在Replica □据集上□行了消融□究,如表 V所示。□基□□始,我 □首先□□□3D□球□替□□我□提出的2D高斯□□表 示。□□一修改就□著提高了□何精度,□渲染深度L1□ 差□4.180厘米降至2.077厘米,同□改善了位姿□□精度 (ATE)□0.17厘米到0.15厘米。然而,□一□化伴□着光 度□量(PSNR)的下降,突出了渲染□□感□□何一致性 之□的□衡。

接下□,我□□合□何感知□化□失,□□失同□□督 光度、一致性深度和表面法□,以及尺度正□化。□□一步 □深度□差降低到0.441厘米,同□保持有□□力的ATE □0.16厘米,□部分恢□光度退化。完整模型在□何精度



Fig. 5: Replica □据集的定性□果。(a), (b) 可□化了房□1和□公室0□景中□□□点的RGB渲染,深度渲染和深度□ 差□。第1和第3行由GS-ICP SLAM渲染,第2和第4行由G²S-ICP SLAM渲染。G²S-ICP SLAM□□了更□□的□ 何形□和更低的深度□差。



Fig. 6: 模□消融□究。每一行□示渲染□像的□果、渲染 深度的L1□差以及渲染法□□。

TABLE V: Replica□据集上各模□的消融□究

Method	PSNR[dB]	Rendered Depth L1 [cm]	ATE [cm]
Baseline	37.92	4.180	0.17
+ 2D Gaussian Disk	34.72	2.077	0.15
+ GA Optimization	36.82	0.441	0.16
Full model	36.88	0.437	0.15

上整□表□最佳,表明二□高斯□□和□何感知□督在提升SLAM性能方面的互□□□。

□些□果□□,每□提□的模□都□最□系□有□□性 的□□,□且表面□□表示和□何感知□失之□的□同作 用在定位和重建保□度上都□到了最新的性能水平。

我□提出了 G² S-ICP SLAM, 一□□何感知的□□ SLAM□架, □□架□高斯点云的□□表□力□面□□的 二□高斯表示的空□一致性相□合。通□□束每□高斯 点位于局部切平面上, 我□的方法解□了常□于基于□□ 3DGS的SLAM系□中的深度不一致□□。此外, 我□引入 了一□□何感知的□失公式,□公式通□光度、一致深度 和表面法□一致性□□督□□,同□施加各向□性尺度正 □化以在□踪□程中□□表面□□。

通□在Replica和TUM-RGBD□据集上的大量□□,我 □的方法在□像机定位和3D重建□量方面展示了□著的 改□,在基于GS的SLAM方法中□到了最先□的性能。□ 得注意的是,我□的系□在不影□渲染□量的情□下,保 持了□□速度的同□增強了深度保□度和□何一致性。□ 些□果突出了□表面感知先□和□何□督集成到SLAM 系□中的有效性,□密集的、逼□的和□□上精□的□景 映射□□了新的方向。

References

- R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos, "Orbslam: a versatile and accurate monocular slam system," *IEEE* transactions on robotics, vol. 31, no. 5, pp. 1147–1163, 2015.
- [2] R. Mur-Artal and J. D. Tardós, "Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras," *IEEE* transactions on robotics, vol. 33, no. 5, pp. 1255–1262, 2017.
- [3] C. Campos, R. Elvira, J. J. G. Rodríguez, J. M. Montiel, and J. D. Tardós, "Orb-slam3: An accurate open-source library for visual, visual-inertial, and multimap slam," *IEEE Transactions on Robotics*, vol. 37, no. 6, pp. 1874–1890, 2021.
- [4] T. Qin, P. Li, and S. Shen, "Vins-mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Transactions* on *Robotics*, vol. 34, no. 4, pp. 1004–1020, 2018.
- [5] C. Forster, M. Pizzoli, and D. Scaramuzza, "Svo: Fast semidirect monocular visual odometry," in 2014 IEEE international conference on robotics and automation (ICRA). IEEE, 2014, pp. 15–22.
- [6] H. M. Cho, H. Jo, and E. Kim, "Sp-slam: Surfel-point simultaneous localization and mapping," *IEEE/ASME Transactions* on *Mechatronics*, vol. 27, no. 5, pp. 2568–2579, 2021.
- [7] H. Huang, L. Li, H. Cheng, and S.-K. Yeung, "Photo-slam: Real-time simultaneous localization and photorealistic mapping for monocular, stereo, and rgb-d cameras," arXiv preprint arXiv:2311.16728, 2023.
- [8] S. Ha, J. Yeon, and H. Yu, "Rgbd gs-icp slam," in European Conference on Computer Vision. Springer, 2024, pp. 180– 197.
- [9] N. Keetha, J. Karhade, K. M. Jatavallabhula, G. Yang, S. Scherer, D. Ramanan, and J. Luiten, "Splatam: Splat, track & map 3d gaussians for dense rgb-d slam," arXiv preprint arXiv:2312.02126, 2023.

- [10] H. Matsuki, R. Murai, P. H. Kelly, and A. J. Davison, "Gaussian splatting slam," arXiv preprint arXiv:2312.06741, 2023.
- [11] C. Yan, D. Qu, D. Wang, D. Xu, Z. Wang, B. Zhao, and X. Li, "Gs-slam: Dense visual slam with 3d gaussian splatting," arXiv preprint arXiv:2311.11700, 2023.
- [12] B. Huang, Z. Yu, A. Chen, A. Geiger, and S. Gao, "2d gaussian splatting for geometrically accurate radiance fields," in ACM SIGGRAPH 2024 conference papers, 2024, pp. 1–11.
- [13] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp." in *Robotics: science and systems*, vol. 2, no. 4. Seattle, WA, 2009, p. 435.
- [14] P. J. Besl and N. D. McKay, "Method for registration of 3d shapes," in Sensor fusion IV: control paradigms and data structures, vol. 1611. Spie, 1992, pp. 586–606.
- [15] B. Kerbl, G. Kopanas, T. Leimkühler, and G. Drettakis, "3d gaussian splatting for real-time radiance field rendering," ACM Transactions on Graphics, vol. 42, no. 4, 2023.
- [16] A. Guédon and V. Lepetit, "Sugar: Surface-aligned gaussian splatting for efficient 3d mesh reconstruction and high-quality mesh rendering," in *Proceedings of the IEEE/CVF Conference* on Computer Vision and Pattern Recognition, 2024, pp. 5354– 5363.
- [17] D. Chen, H. Li, W. Ye, Y. Wang, W. Xie, S. Zhai, N. Wang, H. Liu, H. Bao, and G. Zhang, "Pgsr: Planar-based gaussian splatting for efficient and high-fidelity surface reconstruction," *IEEE Transactions on Visualization and Computer Graphics*, 2024.
- [18] J. Straub, T. Whelan, L. Ma, Y. Chen, E. Wijmans, S. Green, J. J. Engel, R. Mur-Artal, C. Ren, S. Verma *et al.*, "The replica dataset: A digital replica of indoor spaces," *arXiv preprint arXiv:1906.05797*, 2019.
- [19] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of rgb-d slam systems," in 2012 IEEE/RSJ international conference on intelligent robots and systems. IEEE, 2012, pp. 573–580.
- [20] A. Zeng, S. Song, M. Nießner, M. Fisher, J. Xiao, and T. Funkhouser, "3dmatch: Learning local geometric descriptors from rgb-d reconstructions," in CVPR, 2017.
- [21] H. Wang, J. Wang, and L. Agapito, "Co-slam: Joint coordinate and sparse parametric encodings for neural real-time slam," in *Proceedings of the IEEE/CVF Conference on Computer* Vision and Pattern Recognition, 2023, pp. 13293–13302.
- [22] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE transactions on image processing*, vol. 13, no. 4, pp. 600–612, 2004.
- [23] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric," in *CVPR*, 2018.
- [24] Z. Zhu, S. Peng, V. Larsson, W. Xu, H. Bao, Z. Cui, M. R. Oswald, and M. Pollefeys, "Nice-slam: Neural implicit scalable encoding for slam," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 12786–12796.
- [25] E. Sandström, Y. Li, L. Van Gool, and M. R. Oswald, "Pointslam: Dense neural point cloud-based slam," in *Proceedings* of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 18433–18444.