Importance-Driven 3D Gaussian SLAM for Efficient Mapping and Communication-Aware Sharing

Zhaolin Yang[1] Huilin Yin[1*] Linchuan Zhang[2] Mingyu Liu[3] Alois C. Knoll[3]

[1] College of Electronic and Information Engineering, Tongji University, China
[2]Shanghai Research Institute for Intelligent Autonomous Systems, Tongji University, China
[3]Chair of Robotics, Artificial Intelligence and Real-Time Systems, Technical University of Munich, Germany

[*]Corresponding author: yinhuilin@tongji.edu.cn

Abstract

3D Gaussian Splatting (3DGS) enables efficient dense Simultaneous Localization and Mapping (SLAM) but typically produces highly redundant maps, leading to increased computational and communication costs in resource-constrained settings. In this work, we propose a unified importance-driven Gaussian selection framework that addresses both map redundancy and multi-agent communication efficiency. We introduce Proximity-Weighted Accumulation Score (PWAS), a multi-view Gaussian importance metric that quantifies the structural contribution and cross-view consistency of each Gaussian primitive. Based on this metric, we design an importance-aware pruning strategy that removes redundant Gaussians while preserving reconstruction accuracy. Furthermore, we extend the same importance metric to cooperative SLAM and propose a selective sharing mechanism under limited communication budgets. Instead of transmitting full maps, agents exchange only the most informative Gaussians. Experiments on Replica and TUM RGB-D demonstrate that our method significantly reduces map size and communication overhead while maintaining competitive tracking and rendering performance, and consistently outperforms random sharing under identical bandwidth constraints.

1. Introduction

Recent advances in 3D Gaussian Splatting (3DGS) have significantly improved real-time dense scene representation and novel view synthesis quality [7]. By modeling scenes as collections of anisotropic 3D Gaussians with learnable appearance and opacity, 3DGS achieves high-fidelity rendering with efficient differentiable rasterization. Building upon this representation, several recent works have integrated Gaussian-based mapping into Simultaneous Localization and Mapping (SLAM) systems, enabling dense reconstruction with strong photometric consistency and real-time performance [6, 11, 21].

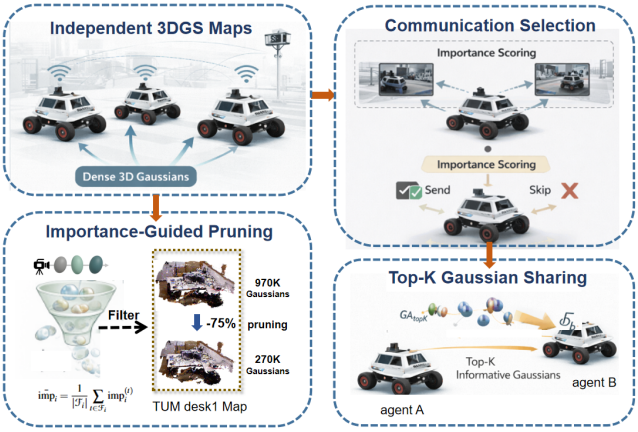

Figure 1. Overview of the proposed importance-guided cooperative Gaussian framework. Each agent performs temporal importance accumulation and local pruning to obtain a compact representation, while importance-guided communication selects and shares only top-K informative Gaussians across agents.

Despite these advantages, Gaussian-based SLAM systems typically suffer from severe redundancy. During incremental mapping, large numbers of Gaussians are continuously densified to improve local reconstruction quality. However, many of these Gaussians contribute marginally to rendering or exhibit unstable geometric behavior across frames. As a result, maps grow rapidly in size, leading to increased memory consumption, slower optimization, and degraded long-term stability. Efficiently identifying and removing redundant Gaussians therefore becomes critical for scalable 3DGS-based SLAM.

At the same time, cooperative perception and multi-agent SLAM have emerged as key components in vehicle-

to-everything (V2X) systems and intelligent transportation scenarios [1, 18, 20]. In such settings, multiple agents collaboratively build and refine maps under strict communication constraints. Transmitting full dense Gaussian maps between agents is often impractical due to bandwidth limitations. Instead, only a subset of informative map elements should be shared. However, existing cooperative mapping approaches typically rely on heuristic selection strategies or transmit entire submaps without principled redundancy modeling [3, 8].

In this work, we argue that redundancy reduction and communication-efficient sharing can be addressed within a unified framework. We introduce a multi-view Gaussian importance metric that evaluates each Gaussian according to its accumulated rendering contribution and cross-frame geometric stability. This metric provides a principled measure of both reliability and redundancy.

Based on this importance estimation, we first design an importance-aware pruning strategy that removes low-value Gaussians while preserving reconstruction fidelity and tracking performance. This leads to more compact and stable maps in single-agent SLAM. Furthermore, we extend the same metric to a cooperative setting and propose an importance-aware selective sharing mechanism. Under limited communication budgets, agents transmit only the top-ranked Gaussians according to their importance scores, instead of full maps or randomly selected subsets.

Extensive experiments on Replica [15] and TUM RGB-D [16] datasets demonstrate that our framework simultaneously achieves effective map compression and communication-efficient cooperative mapping. Under identical bandwidth constraints, importance-based sharing consistently outperforms random selection and enables partial-observation agents to approach the performance of full-information baselines.

Our contributions can be summarized as follows:

- We propose a unified multi-view Gaussian importance metric that captures rendering contribution and geometric stability.
- We design an importance-aware pruning strategy for compact and stable 3DGS-based SLAM.
- We extend the importance framework to cooperative mapping and introduce a selective sharing mechanism under communication constraints.
- We validate the effectiveness of our approach in both single-agent and multi-agent settings.

2. Related Work

Recent advances in dense SLAM have shifted from sparse point-based representations toward more expressive scene models. Neural radiance fields (NeRF) [12] demonstrate high-fidelity view synthesis, but their reliance on volumetric sampling and deep neural networks limits real-time applicability. To address these limitations, 3D Gaussian Splatting (3DGS) [7] introduces an efficient differentiable representation based on anisotropic Gaussian primitives, enabling real-time photorealistic rendering.

Building upon 3DGS, several recent works have integrated Gaussian representations into SLAM pipelines. Systems such as GS-SLAM [21] and SplaTAM [6] achieve real-time dense reconstruction with strong photometric consistency. However, these methods typically accumulate Gaussian primitives continuously during densification without explicit redundancy control, leading to uncontrolled map growth and increasing memory consumption. While approaches such as LSG-SLAM [19] introduce submap-level consolidation and dynamic variants like D4DGS-SLAM [17] address scene motion, principled management of Gaussian redundancy remains largely unexplored.

In parallel, the rendering community has proposed various strategies for compact 3DGS representations. Methods such as Scaffold-GS [9] and Taming 3DGS [10] control densification and optimize representation efficiency, while MiniSplatting [5], LightGaussian [4], HAC [2], and C3DGS [13] focus on pruning, clustering, and attribute compression. Although effective for offline view synthesis, these approaches do not directly address the online requirements of SLAM, where temporal consistency, pose coupling, and incremental updates must be preserved.

Beyond single-agent efficiency, distributed and cooperative SLAM have long been studied in multi-robot systems [1, 18, 20]. Traditional approaches focus on pose graph fusion and distributed optimization, typically exchanging keyframes or sparse constraints. In contrast, cooperative perception in V2X systems explores feature-level or intermediate representation sharing under bandwidth constraints [3, 8]. However, dense Gaussian maps introduce a new challenge: transmitting the full set of primitives is often infeasible due to large memory footprint.

Unlike heuristic submap transmission or random selection strategies, our work introduces a unified importance metric that simultaneously addresses Gaussian redundancy and communication-aware selection. By ranking primitives according to multi-view contribution and geometric stability, we enable both compact single-agent mapping and bandwidth-efficient multi-agent sharing within a consistent framework.

3. Method

3.1. 3DGS SLAM

As shown in Fig. 2, we build upon a standard 3DGS SLAM pipeline and introduce a proximity-weighted accumulation score (PWAS) to guide both local pruning and inter-agent communication. The scene is represented by a set of Gaus-

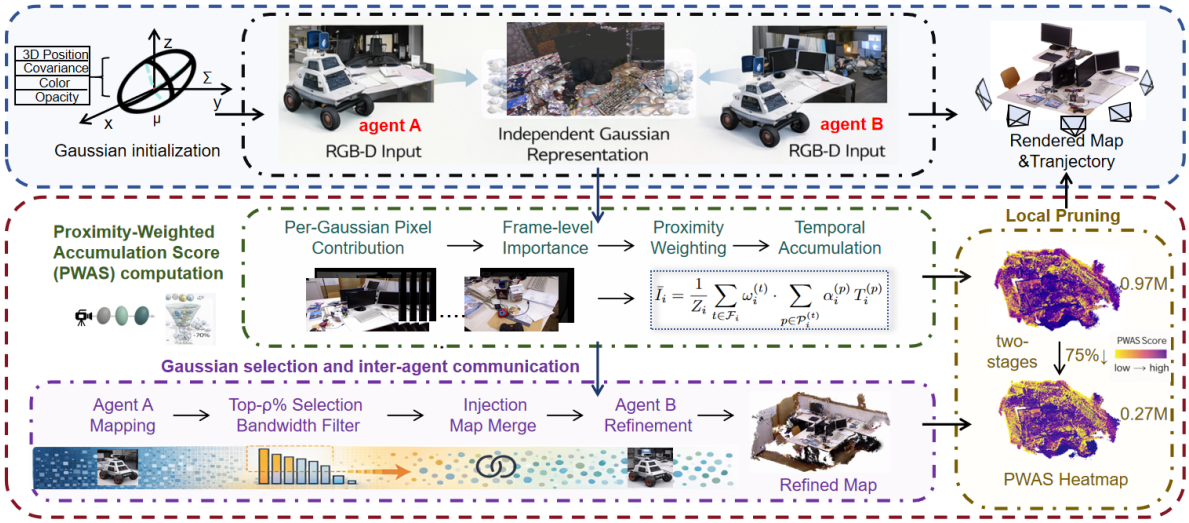

Figure 2. Framework of the proposed PWAS-guided 3DGS SLAM system. We introduce a proximity-weighted accumulation score (PWAS) that quantifies long-term structural contribution of each Gaussian primitive. PWAS is computed from multi-view pixel contributions with distance-aware weighting, and subsequently governs (1) stage-wise local pruning for model compression and (2) importance-aware inter-agent communication under bandwidth constraints. This unified metric bridges memory efficiency and communication efficiency within a single structure-aware optimization framework.

sian primitives $\mathcal{G} = \{(\mu_i, \Sigma_i, \alpha_i, \mathbf{c}_i)\}_{i=1}^{N}$, with $\mu_i \in \mathbb{R}^3$ the Gaussian center, Σ_i the covariance, α_i the opacity, and $\mathbf{c}_i$ the view-dependent appearance coefficients.

Given the current camera pose $\mathbf{T}_t \in SE(3)$ and intrinsics, projected Gaussians are then sorted by geometric depth z_i and alpha-blended to compute pixel colors (C_p) and rendered depths (D_p):

$$C_p = \sum_{i=1}^{N} \gamma_i \alpha_i \prod_{j=1}^{i-1}(1-\alpha_j), \quad D_p = \sum_{i=1}^{N} z_i \alpha_i \prod_{j=1}^{i-1}(1-\alpha_j), \quad (1)$$

where γ_i is the color of the i-th Gaussian and N denotes depth-ordered Gaussians, z_i is the depth of μ_i in camera coordinates.

In mapping, we adopt the structure-preserving supervision [23], where a pseudo-target image I_{fuse} is constructed by combining the rendered image with geometric and structural guidance derived from depth and edge consistency. Formally, $I_{\text{fuse}} = \Phi(I_{\text{render}}, D_{\text{gt}}, \nabla D_{\text{gt}})$, where $\Phi(\cdot)$ denotes a fusion operator that enhances structural boundaries and suppresses illumination-induced distortions.

Concretely, we define an illumination residual $L_{\text{illum}} = \|I_{\text{render}} - I_{\text{fuse}}\|_1$, a depth residual $L_{\text{depth}} = \|D_{\text{render}} - D_{\text{gt}}\|_1$, and a photometric term $L_{\text{color}} = 0.8\,\|I_{\text{render}} - I_{\text{gt}}\|_1 + 0.2\,(1 - \text{SSIM}(I_{\text{render}}, I_{\text{gt}}))$, leading to the final mapping loss

$$L_{\text{map}} = 0.5\,L_{\text{color}} + L_{\text{depth}} + \omega_{\text{illum}}\,L_{\text{illum}}, \quad (2)$$

where the illumination weight is adaptively adjusted to avoid over-dominance:

$$\omega_{\text{illum}} = \min\left(\frac{L_{\text{illum}}}{L_{\text{color}} + \varepsilon}, \tau\right), \quad (3)$$

where ε avoids division by zero and τ limits the weight to prevent it from dominating the mapping loss.

For tracking, we jointly uses color and depth residuals but restricts supervision to reliable high-opacity pixels. Specifically, we compute adaptive residual weights

$$w_{\text{im}} = \frac{\gamma_{\text{im}}}{L_{\text{color}} + \gamma_{\text{im}}}, \quad w_{\text{depth}} = \frac{\gamma_{\text{depth}}}{L_{\text{depth}} + \gamma_{\text{depth}}}, \quad (4)$$

where γ_{im} and γ_{depth} are residual sensitivity coefficients that modulate the sensitivity of the respective weights to residual magnitudes. and optimize camera pose while keeping Gaussian parameters fixed using

$$\mathcal{L}_{\text{Tracking}} = \mathbf{1}_{\{O_p > 0.99\}} \Big(w_{\text{im}} \|C_p\|_1 + w_{\text{depth}} \|D_p\|_1 + \lambda_R (\log \frac{w_{\text{depth}}}{w_{\text{im}}} - \log \rho)^2 \Big), \quad (5)$$

where $\mathbf{1}_{\{\cdot\}}$ is the indicator function, O_p denotes rendered opacity, ρ is a pre-calibrated ideal color-to-depth weight ratio, and the log-ratio regularizer prevents the adaptive weights from drifting to extreme values, improving tracking stability across scenes with different texture, depth range, and noise characteristics.

3.2. Proximity-Weighted Accumulation Score

In 3DGS-based SLAM, the number of Gaussians grows continuously due to densification, which inevitably introduces redundancy. We therefore seek a principled metric to quantify the long-term utility of each Gaussian across time and viewpoints. Our key observation is that a structurally important Gaussian should consistently contribute to rendered observations and receive stable support from geometrically relevant camera views.

Per-frame contribution. For a Gaussian g_i observed at frame t, we define its rendering contribution as the accumulated opacity-weighted visibility over all covered pixels:

$$I_i^{(t)} = \sum_{p \in \mathcal{P}_i^{(t)}} \alpha_i^{(t)}(p)\, T_i^{(t)}(p), \tag{6}$$

where $\alpha_i^{(t)}(p)$ and $T_i^{(t)}(p)$ denote the projected opacity and transmittance along the viewing ray at pixel p, respectively. The term αT naturally arises from alpha compositing and reflects the effective radiance contribution of g_i to the rendered image.

Proximity weighting. To avoid over-emphasizing incidental distant observations, we introduce a proximity-based attenuation factor:

$$\omega_i^{(t)} = \exp\left(-\frac{\|\mu_i - c_t\|_2}{\sigma_d}\right), \tag{7}$$

where μ_i is the Gaussian center and c_t is the camera center at frame t. The parameter σ_d controls the decay rate, encouraging contributions from nearby and geometrically stable viewpoints. By accumulating rendering contributions over multiple frames, PWAS reduces the influence of single-view outliers and ensures that importance reflects long-term structural consistency rather than instantaneous visibility.

Temporal accumulation. The final importance score of Gaussian g_i is defined as the normalized multi-view accumulation:

$$\bar{I}_i = \frac{\sum_{t \in \mathcal{F}_i} \omega_i^{(t)} I_i^{(t)}}{\sum_{t \in \mathcal{F}_i} \omega_i^{(t)} + \varepsilon}, \tag{8}$$

where $\mathcal{F}_i$ denotes the set of frames in which g_i is visible. This formulation ensures that Gaussians consistently contributing to reconstruction across time obtain higher scores, while transient or weakly observed primitives are naturally suppressed.

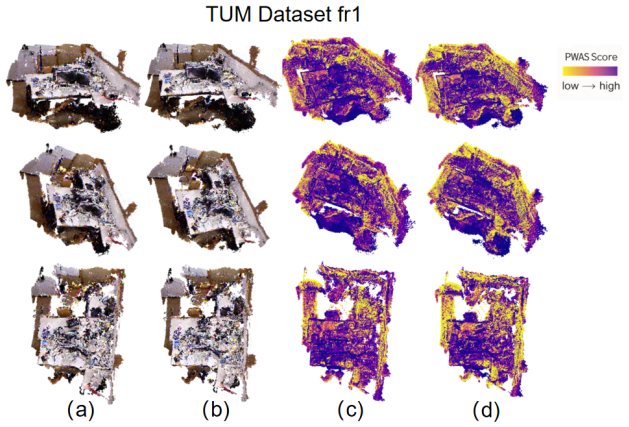

Figure 3. Visualization of PWAS-guided pruning on the TUM fr1 sequence. (a) Gaussian center point cloud before pruning. (b) Gaussian center point cloud after stage-wise PWAS pruning. (c) PWAS importance distribution before pruning. (d) Importance distribution after pruning. High-importance regions (yellow) concentrate around structurally stable and frequently observed areas, while low-importance regions (purple) correspond to redundant or weakly supported primitives.

Compared with heuristic pruning strategies based solely on opacity or screen-space extent, PWAS explicitly measures long-term structural relevance, making it suitable as a unified criterion for both local map compression and inter-agent Gaussian selection.

3.3. Importance-Guided Local Map Compression

To prevent uncontrolled growth of Gaussian primitives during long-term mapping, we adopt a hybrid pruning strategy. Conventional 3DGS-SLAM systems typically rely on periodic visibility-based pruning to remove inactive Gaussians. Following this practice, we perform lightweight cyclic pruning every 100 mapping iterations, removing primitives with persistently low opacity:

$$\mathcal{G}_{\text{vis}} = \{g_i \in \mathcal{G} \mid \alpha_i > \alpha_{\min}\}, \tag{9}$$

where α_i denotes Gaussian opacity and $\alpha_{\min}$ is a small visibility threshold. While effective for eliminating clearly inactive primitives, such local criteria are inherently myopic and do not capture long-term structural relevance across viewpoints.

To address this limitation, we introduce stage-wise importance-driven compression based on the accumulated PWAS score $\bar{I}_i$. Instead of continuously pruning small portions of Gaussians, we perform two global compression stages positioned at $\frac{1}{3}$ and $\frac{2}{3}$ of the total mapping iterations. At each stage, Gaussians are ranked by their importance scores, and only the top 30% are retained:

$$\mathcal{G}' = \text{Top}_{30\%}\left(\{\bar{I}_i \mid g_i \in \mathcal{G}\}\right). \tag{10}$$

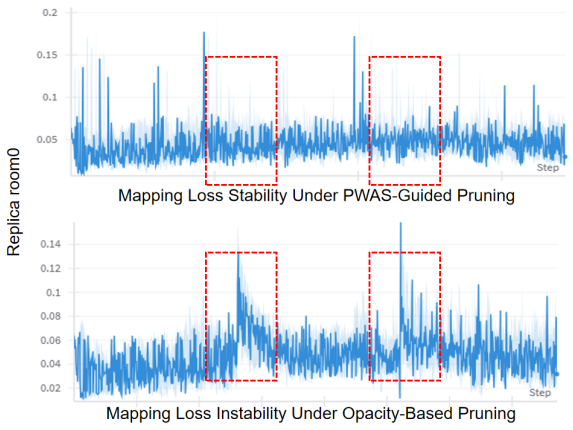

Figure 4. Mapping loss evolution under PWAS-guided pruning (top) and opacity-based pruning (bottom) on Replica Room0. The highlighted regions correspond to pruning intervals. PWAS pruning maintains a smoother loss profile, whereas opacity-based pruning results in larger fluctuations. This suggests that importance-aware selection better aligns with structural contribution in the scene representation.

Because $\bar{I}_i$ aggregates multi-view rendering contribution over time, this selection favors structurally stable and frequently observed primitives, while aggressively removing redundant Gaussians that lack sustained geometric support. Unlike opacity-only heuristics, the PWAS score reflects view-consistent transmittance interaction and cross-frame evidence, making it better aligned with structural significance. As further illustrated in Fig. 3, the PWAS heatmap reveals that high-importance Gaussians concentrate around structurally stable and frequently observed regions, while low-importance primitives correspond to redundant interior areas. After pruning, geometric structure is largely preserved.

Importantly, importance accumulation is updated incrementally across frames, preventing premature removal of newly inserted primitives that have not yet accumulated sufficient evidence. As shown in Fig. 4, even under aggressive compression which removes 70% of Gaussians at each stage, the mapping loss remains stable, indicating that discarded primitives contribute minimally to structural reconstruction. This demonstrates that PWAS enables principled large-scale model reduction while preserving mapping consistency and visual fidelity.

3.4. Importance-Aware Communication

While local pruning ensures compact single-agent mapping, multi-agent systems introduce an additional constraint: limited communication bandwidth. Transmitting the full Gaussian map between agents is prohibitively expensive, especially as the scene scale grows. We therefore extend the proposed importance metric to guide communication-efficient Gaussian exchange.

Communication budget formulation. Assume agent A maintains a local Gaussian set $\mathcal{G}_A$ and intends to share a subset $\mathcal{G}_{A\to B}$ with agent B. Given a communication budget allowing transmission of only ρ fraction of primitives, the selection problem becomes:

$$\mathcal{G}_{A\to B} = \arg \max_{\substack{\mathcal{S}\subset\mathcal{G}_A \\ |\mathcal{S}|=\rho|\mathcal{G}_A|}} \sum_{g_i\in\mathcal{S}} \bar{I}_i, \tag{11}$$

where $\bar{I}_i$ is the importance score defined in Eq. (8). The optimal solution corresponds to selecting the top-$\rho\%$ Gaussians ranked by importance.

Offline injection protocol. In practice, we simulate cooperative mapping by assuming two agents traverse the same underlying camera trajectory while observing disjoint subsets of frames. Specifically, the input sequence is partitioned into alternating frames, where agent A receives only odd-indexed frames and agent B receives only even-indexed frames. Each agent performs mapping and pruning independently using only its assigned observations.

Let $\mathcal{G}_A$ and $\mathcal{G}_B$ denote the Gaussian sets constructed from the odd and even frame subsets, respectively. This setup ensures that each agent has partial but complementary scene observations, enabling controlled evaluation of importance-aware Gaussian sharing under consistent motion.

At the communication stage, agent A selects a bandwidth-constrained subset $\mathcal{G}_{A\to B}$ according to the optimization formulation in Eq. (11). The selected Gaussians are injected into agent B prior to its mapping optimization. Concretely, we concatenate the transmitted Gaussian parameters into $\mathcal{G}_B$, including means, covariance representations, opacity, and appearance coefficients, then reinitialize the optimizer to ensure that all primitives jointly participate in subsequent gradient updates.

This design allows the transmitted Gaussians to serve as structured initialization rather than frozen priors, enabling agent B to refine both locally observed and remotely injected primitives under its own viewpoint distribution. Because selection is governed by the accumulated importance score $\bar{I}_i$, transmitted Gaussians correspond to primitives with sustained multi-view contribution and structural stability.

Notably, the same importance metric governs both local map compression and inter-agent communication. This unifies memory efficiency and bandwidth efficiency under a single quantitative criterion, positioning PWAS as a general resource allocation principle for scalable multi-agent 3DGS-based SLAM.

Table 1. Camera tracking result on Replica(ATE RMSE ↓[cm])

	Method	Avg.	R0	R1	R2	Of0	Of1	Of2	Of3	Of4
NeRF based	NICE SLAM	1.06	0.97	1.31	1.07	0.88	1.00	1.06	1.10	1.13
	Point SLAM	0.52	0.61	0.41	0.37	0.38	0.48	0.54	0.69	0.72
	Vox Fusion	3.09	1.37	4.70	1.47	8.48	2.04	2.58	1.11	2.94
3DGS based	MonoGS	0.58	0.44	0.32	0.31	0.44	0.52	**0.23**	0.17	2.25
	SplaTAM	0.36	0.31	0.40	0.29	0.47	0.27	0.29	0.32	0.55
	Ours	**0.27**	**0.24**	**0.31**	**0.24**	**0.37**	**0.22**	0.28	**0.15**	**0.36**

Table 2. Camera tracking result on TUM(ATE RMSE ↓[cm]). *Indicates that only three sequences are counted in the average

	Method	Avg.	fr1/ desk	fr1/ desk2	fr1/ room	fr2/ xyz	fr3/ off
NeRF based	NICE SLAM	15.87	4.26	4.99	34.49	31.73	3.87
	Point SLAM	8.92	4.34	**4.54**	30.92	1.31	3.48
	Vox Fusion	11.31	3.52	6.00	19.53	1.49	26.01
3DGS based	*MonoGS	*1.47	1.50	-	-	1.44	**1.49**
	SplaTAM	5.48	3.35	6.54	**11.13**	**1.24**	5.16
	Ours	**4.46**	**0.53**	4.83	11.3	1.32	4.34

4. Experiments

4.1. Experimental Setup

Datasets. We evaluate our method on standard RGB-D benchmarks widely adopted in dense SLAM and 3D Gaussian Splatting literature. Specifically, we use sequences from the Replica dataset [15] and TUM RGB-D dataset [16], which provide synchronized color and depth streams with ground-truth camera trajectories. These datasets contain indoor scenes with varying texture richness, illumination conditions, and geometric complexity, enabling evaluation under diverse scenarios.

Evaluation Metrics. To assess both geometric accuracy and rendering quality, we report the following metrics: (1) Absolute Trajectory Error (ATE RMSE) for pose estimation accuracy, (2) Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Learned Perceptual Image Patch Similarity (LPIPS) for rendered image fidelity, (3) the total number of Gaussians as a measure of map compactness.

Implementation Details. Our system is implemented in PyTorch and built upon a 3DGS-based RGB-D SLAM pipeline. The proximity bandwidth parameter σ_d in PWAS is fixed 20 across all experiments. For local pruning, we retain 30% Gaussians ranked by the accumulated importance score. For multi-agent communication, we select the top-20% Gaussians for sharing. All experiments are conducted on a single NVIDIA RTX 4090 GPU.

4.2. Overall SLAM Performance

We first evaluate the tracking performance of the overall SLAM pipeline. Table 1 and Table 2 report camera tracking accuracy in terms of ATE RMSE (cm) on the Replica and TUM RGB-D benchmarks.

On the Replica dataset, our method achieves the lowest average ATE among all 3DGS-based approaches and outperforms SplaTAM across most sequences. Compared to MonoGS [11] and SplaTAM [6] , the proposed method reduces the average tracking error while maintaining stable performance across different scene layouts. Importantly, this improvement is achieved despite aggressive stage-wise Gaussian compression, indicating that importance-guided pruning does not degrade geometric consistency.

On the TUM dataset, our approach achieves competitive performance with respect to both NeRF-based [14, 22, 24] and 3DGS-based baselines [6, 11]. Although tracking accuracy is not the primary focus of this work, the results demonstrate that PWAS-guided compression preserves pose estimation reliability even under significant model reduction.

Overall, these results confirm that the proposed importance-driven resource allocation framework maintains camera tracking accuracy, establishing a solid foundation for subsequent compression and communication analysis.

4.3. Importance-Guided Map Compression

We evaluate the lightweight capability of PWAS-guided pruning on both Replica and TUM datasets. Table 3 summarizes the number of Gaussian primitives before and after pruning.

On Replica, PWAS reduces the primitive count by an average of 77.4%, with reductions consistently ranging from 64% to over 84% across different scenes. On the real-world TUM dataset, an average reduction of 63.5% is achieved. Despite aggressive pruning, tracking accuracy remains stable, indicating that structurally important Gaussians are preserved.

Fig. 5 further visualizes the compression on Replica. The radar comparison shows a uniform shrinkage of the Gaussian footprint under PWAS, demonstrating that redundant primitives are selectively removed rather than arbitrarily discarded. Combined with the stable loss behavior observed during pruning, these results confirm that PWAS enables structure-aware model compression instead of heuristic opacity filtering.

4.4. Importance-Guided Communication

We evaluate the proposed PWAS-based transmission strategy under a fixed 20% communication budget and compare it with random 20% selection as well as independent agent mapping. Quantitative results are reported in Table 4.

Quantitative comparison. Under the same bandwidth constraint, PWAS Top-20% selection consistently achieves lower ATE and LPIPS while improving PSNR and SSIM

Table 3. Gaussian primitive reduction under PWAS-guided pruning on Replica and TUM datasets.

Scene	**Before** (Gaussians)	**After** (Gaussians)	**Reduction** (%)
	Replica		
room0	5,050,946	1,226,354	75.7
room1	6,712,262	1,029,088	84.7
room2	5,775,957	1,597,422	72.3
office0	6,230,915	1,563,085	74.9
office1	5,660,015	1,462,063	74.2
office2	4,547,238	1,086,412	76.1
office3	5,256,856	1,876,223	64.3
office4	5,347,215	148,830	97.2
Average (Replica)			77.4
	TUM		
fr1 desk	2,719,797	982,323	63.9
fr1 desk2	1,383,576	330,700	76.1
fr1 room1	4,014,810	998,750	75.1
fr2 xyz	6,343,054	3,249,425	48.8
fr3 off	830,278	385,623	53.6
Average (TUM)			63.5

Table 4. Multi-agent communication results under 20% transmission budget.

	Metric	Replica			TUM		
		room0	room1	room2	fr1 room	fr1 desk	fr3 off
agentA	ATE(cm) ↓	0.23	0.29	0.23	0.35	1.00	4.27
	PSNR ↑	32.6	33.8	34.2	17.49	19.80	21.27
	SSIM ↑	0.975	0.972	0.974	0.797	0.797	0.871
	LPIPS ↓	0.083	0.084	0.085	0.341	0.290	0.209
agentB	ATE(cm) ↓	0.25	1.62	0.33	0.38	1.18	4.64
	PSNR ↑	32.6	33.7	34.0	19.09	20.50	21.36
	SSIM ↑	0.976	0.972	0.972	0.795	0.795	0.876
	LPIPS ↓	0.082	0.086	0.085	0.305	0.299	0.211
Random 20%	ATE(cm) ↓	0.28	1.52	0.33	0.37	1.10	4.55
	PSNR ↑	32.5	33.7	33.8	19.23	20.66	21.55
	SSIM ↑	0.973	0.972	0.971	0.810	0.810	0.877
	LPIPS ↓	0.085	0.085	0.085	0.291	0.277	0.207
PWAS 20%	ATE(cm) ↓	**0.24**	**1.47**	**0.30**	**0.37**	**0.98**	**4.51**
	PSNR ↑	**32.7**	**33.9**	**35.0**	**19.91**	**21.01**	**21.65**
	SSIM ↑	**0.978**	**0.973**	**0.978**	**0.810**	**0.810**	**0.890**
	LPIPS ↓	**0.080**	**0.083**	**0.083**	**0.275**	**0.272**	**0.194**

across both Replica and TUM datasets. In particular, trajectory error is reduced in all evaluated sequences compared to random transmission, demonstrating that importance-aware Gaussian selection better preserves geometry relevant to pose estimation. The performance gap is more pronounced in complex scenes such as *fr1 desk* and *fr3 off*, where random selection leads to noticeable degradation.

Visual reconstruction quality. Qualitative comparisons are shown in Fig. 6. Under identical 20% transmission, PWAS preserves structural continuity and fine details in high-texture regions, whereas random selection introduces visible artifacts and incomplete surfaces. The highlighted areas in TUM and Replica scenes clearly illustrate that importance-guided transmission maintains geometric consistency despite aggressive compression.

Trajectory stability. Figure. 7 further visualizes the ATE distribution along trajectories. PWAS Top-20% transmission yields lower accumulated drift and smoother error distribution compared to random selection. This indicates that transmitting structurally informative Gaussians directly benefits downstream pose refinement and reduces long-term drift.

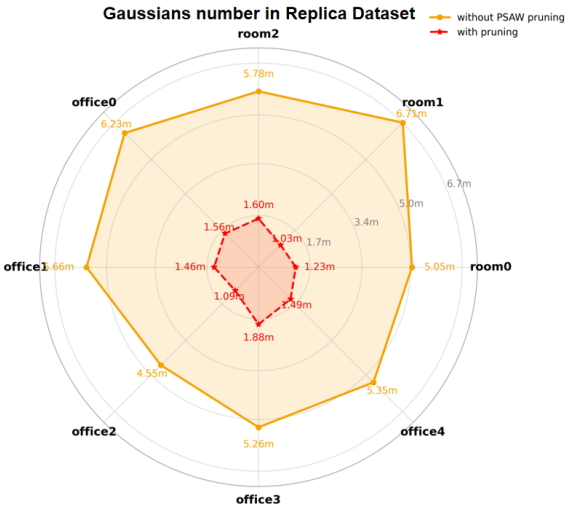

Figure 5. Comparison of Gaussian primitive counts on the Replica dataset. The orange curve represents the baseline without pruning, while the red curve corresponds to PWAS-guided pruning. Across all scenes, PWAS consistently reduces the number of Gaussians by a large margin while maintaining structural coverage.

Figure 6. Visual reconstruction comparison under identical 20% transmission budget. PWAS Top-20% selection maintains structural integrity and detail preservation in challenging regions (green boxes), while random 20% selection results in noticeable structural loss and texture distortion.

Overall, these results confirm that PWAS serves not only as a local compression metric but also as an effective bandwidth allocation criterion, enabling improved reconstruction fidelity and trajectory accuracy under strict communication constraints.

4.5. Ablation Study

The proposed PWAS score serves two roles in our framework: (i) importance-guided local pruning and (ii)

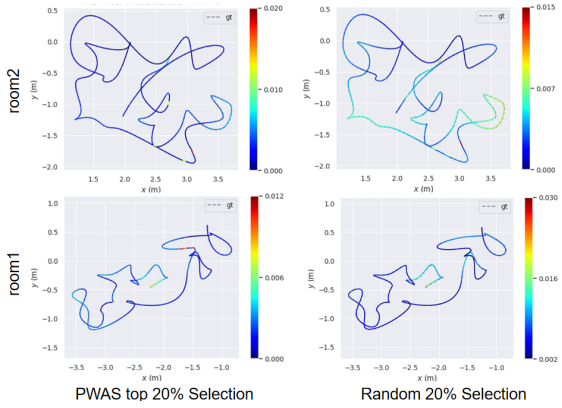

Figure 7. Trajectory error visualization under 20% communication budget. Color encodes ATE magnitude. PWAS Top-20% selection (left) produces lower accumulated drift and improved trajectory consistency compared to Random 20% selection (right).

bandwidth-aware inter-agent transmission. While the communication experiments in Table 4 already demonstrate that PWAS-based selection outperforms random transmission under identical bandwidth constraints, those results are evaluated within our full system pipeline. To isolate the contribution of the pruning criterion itself, we conduct a controlled ablation study.

Pruning transfer to SplaTAM. We integrate PWAS-based pruning into the original SplaTAM pipeline without modifying any other components, including tracking loss, densification strategy, keyframe policy, and periodic cleanup. The only change is replacing the default visibility-based pruning rule with our stepwise PWAS pruning schedule.

As shown in Table 5, introducing PWAS pruning significantly reduces the number of Gaussian primitives while maintaining comparable tracking accuracy. The ATE remains stable, indicating that aggressive compression does not degrade pose estimation quality.

This experiment confirms that the effectiveness of PWAS does not rely on our specific system design. Instead, it functions as a general importance metric that can be seamlessly incorporated into existing 3DGS-based SLAM frameworks.

Combined with the communication experiments, which already compare PWAS against random selection under identical transmission ratios, the ablation results demonstrate that PWAS consistently improves both memory efficiency and information efficiency through a unified importance formulation.

5. Conclusion

This work addresses the problem of principled Gaussian redundancy management in 3D Gaussian Splatting SLAM under both memory and communication constraints. While recent 3DGS-based SLAM systems have improved reconstruction fidelity and scalability, efficient selection of informative Gaussian primitives remains insufficiently studied.

We introduced PWAS, an importance-aware formulation that unifies local model compression and multi-agent communication through multi-view accumulated rendering contribution and proximity-aware weighting. Instead of relying on opacity heuristics or submap-level transmission, PWAS provides a continuous and structurally grounded criterion for Gaussian selection.

Extensive experiments on Replica and TUM demonstrate that PWAS significantly reduces the number of Gaussian primitives while maintaining tracking accuracy and rendering quality. Under strict bandwidth constraints, importance-guided transmission consistently outperforms random selection in both visual fidelity and trajectory stability. Furthermore, transferring PWAS pruning to the original SplaTAM pipeline confirms its generality and system-independence.

Overall, PWAS establishes a unified importance metric for resource-efficient 3DGS SLAM, bridging model compression and communication through a single quantitative principle.

Table 5. Ablation on pruning strategy (Replica dataset).

Method	ATE ↓ (cm)	Gaussians ↓
SplaTAM (original pruning)	0.36	6.71M
SplaTAM + PWAS pruning	**0.35**	**1.58M**

Limitations and Future Work. PWAS relies on accumulated rendering statistics, which assume sufficient multi-view coverage to stabilize importance estimation. In highly dynamic or extremely sparse observation settings, importance estimation may become less reliable. Future work will explore adaptive importance updating mechanisms, dynamic-scene-aware importance modeling, and fully decentralized multi-agent scheduling strategies.

6. Acknowledgements

This work was supported by the National Natural Science Foundation of China under Grant No. 62433014 and No.62133011.

References

[1] Qi Chen, Sihai Tang, Qing Yang, and Song Fu. Cooper: Cooperative perception for connected autonomous vehicles based on 3D point clouds. In *IEEE International Conference on Distributed Computing Systems (ICDCS)*, 2019.

[2] Yihang Chen, Qianyi Wu, Weiyao Lin, Mehrtash Harandi, and Jianfei Cai. HAC: Hash-grid assisted context for 3D

gaussian splatting compression. In *European Conference on Computer Vision (ECCV)*, 2024.

[3] Alexander Cunningham, Kai M. Wurm, Wolfram Burgard, and Frank Dellaert. Fully distributed scalable smoothing and mapping with robust multi-robot data association. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2012.

[4] Zhiwen Fan, Kevin Wang, Kairun Wen, Zehao Zhu, Dejia Xu, and Zhangyang Wang. LightGaussian: Unbounded 3D gaussian compression with 15x reduction and 200+ fps. In *Advances in Neural Information Processing Systems*, 2024.

[5] Guangchi Fang and Bing Wang. Mini-Splatting: Representing scenes with a constrained number of gaussians. In *European Conference on Computer Vision (ECCV)*, 2024.

[6] Nikhil Keetha, Jay Karhade, Krishna Murthy Jatavallabhula, Gengshan Yang, Sebastian Scherer, Deva Ramanan, and Jonathon Luiten. SplaTAM: Splat, track and map 3D gaussians for dense RGB-D SLAM. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

[7] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3D gaussian splatting for real-time radiance field rendering. *ACM Transactions on Graphics*, 2023.

[8] Pierre-Yves Lajoie, Benjamin Ramtoula, Fang Wu, and Giovanni Beltrame. Towards collaborative simultaneous localization and mapping: A survey of the current research landscape. *Field Robotics*, 2022.

[9] Tao Lu, Mulin Yu, Linning Xu, Yuanbo Xiangli, Limin Wang, Dahua Lin, and Bo Dai. Scaffold-GS: Structured 3D gaussians for view-adaptive rendering. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

[10] Shashanka Subhra Mallick, Rahul Goel, Bernhard Kerbl, Markus Steinberger, Francisco Vicente Carrasco, and Fernando De La Torre. Taming 3DGS: High-quality radiance fields with limited resources. In *SIGGRAPH Asia Conference Papers*, 2024.

[11] Hidenobu Matsuki, Riku Murai, Paul H. J. Kelly, and Andrew J. Davison. Gaussian splatting SLAM (MonoGS). In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

[12] Ben Mildenhall, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. NeRF: Representing scenes as neural radiance fields for view synthesis. In *European Conference on Computer Vision (ECCV)*, 2020.

[13] Jiating Qian, Yiming Yan, Fengjiao Gao, Baoyu Ge, Maosheng Wei, Boyi Shangguan, and Guangjun He. C3DGS: Compressing 3D gaussian model for surface reconstruction. *IEEE JSTARS*, 2025.

[14] Erik Sandström, Yue Li, Luc Van Gool, and Martin R. Oswald. Point-SLAM: Dense neural point cloud-based SLAM. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023.

[15] Julian Straub et al. The Replica dataset: A digital replica of indoor spaces. *arXiv preprint arXiv:1906.05797*, 2019.

[16] Jürgen Sturm, Nikolas Engelhard, Felix Endres, Wolfram Burgard, and Daniel Cremers. A benchmark for the evaluation of RGB-D SLAM systems. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2012.

[17] Zhicong Sun, Jacqueline Lo, and Jinxing Hu. Embracing dynamics: Dynamics-aware 4D gaussian splatting SLAM. *arXiv preprint arXiv:2504.04844*, 2025.

[18] Tsun-Hsuan Wang, Sivabalan Manivasagam, Ming Liang, Bin Yang, Wenyuan Zeng, James Tu, and Raquel Urtasun. V2VNet: Vehicle-to-vehicle communication for joint perception and prediction. In *European Conference on Computer Vision (ECCV)*, 2020.

[19] Zhe Xin, Chenyang Wu, Penghui Huang, Yanyong Zhang, Yinian Mao, and Guoquan Huang. Large-scale gaussian splatting SLAM. *arXiv preprint arXiv:2505.09915*, 2025.

[20] Runsheng Xu, Hao Xiang, Xin Xia, Xu Han, Jinlong Li, and Jiaqi Ma. OPV2V: An open benchmark dataset and fusion pipeline for perception with vehicle-to-vehicle communication. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2022.

[21] Congliang Yan, Dongdong Qu, Dong Xu, Botian Zhao, Zhenzhong Wang, Dong Wang, and Xuelong Li. GS-SLAM: Dense visual SLAM with 3D gaussian splatting. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024.

[22] Xingrui Yang, Hai Li, Hongjia Zhai, Yuhang Ming, Yuqian Liu, and Guofeng Zhang. Vox-Fusion: Dense tracking and mapping with voxel-based neural implicit representation. In *IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, 2022.

[23] Huilin Yin, Zhaolin Yang, Linchuan Zhang, Gerhard Rigoll, and Johannes Betz. RoGER-SLAM: A robust gaussian splatting SLAM system for noisy and low-light environment resilience. *IEEE Transactions on Instrumentation and Measurement*, 2026.

[24] Zihan Zhu, Songyou Peng, Viktor Larsson, Weiwei Xu, Hujun Bao, Zhaopeng Cui, Martin R. Oswald, and Marc Pollefeys. NICE-SLAM: Neural implicit scalable encoding for SLAM. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.