Towards Enhanced Sparse-View Tomographic Reconstruction Using 3D Gaussian Splatting

Aqsa Yousaf, Paul Agbaje, Afia Anjum, Arkajyoti Mitra, Habeeb Olufowobi
The University of Texas at Arlington
{aqsa.yousaf, pauloluwatowoju.agbaje, afia.anjum, arkajyoti.mitra, habeeb.olufowobi}@uta.edu

Abstract

Sparse-view tomographic reconstruction aims to recover 3D volumes from limited projection views, but often suffers from incomplete structures and volumetric artifacts. Gaussian splatting has recently emerged as an efficient representation for continuous volumetric modeling, reducing memory cost compared to voxel grids and training time compared to implicit methods. However, existing Gaussian splatting methods for CT reconstruction struggle with needle-like artifacts in sparse-view settings. To address this, we introduce two key contributions. First, we propose a structure-aware initialization strategy that uses gradient and density magnitude from preliminary reconstructions to intelligently place Gaussian primitives in high-contrast regions. Second, we adapt the well-established Beer-Lambert law from CT physics to stabilize Gaussian splatting optimization, transforming the exponential attenuation relationship into a linear domain that mitigates vanishing gradients, and stabilizes optimization. Together, these innovations lead to sharper and more stable reconstructions, achieving average improvements of 2.32% in PSNR and 2.41% in SSIM while using 6.47% fewer primitives across three standard CT datasets.

1. Introduction

Computed tomography (CT) is a widely used medical imaging technique that reconstructs 3D volumetric structures from multiple 2D X-ray projections, offering critical insights into internal anatomy. However, conventional CT requires dense angular sampling, which increases radiation dose and hinders its use in low-dose or real-time scenarios. Sparse-view CT reduces the number of projections, but this leads to ill-posed inverse problems with structural artifacts, noise, and volumetric inconsistencies [13, 20].

The research reported in this paper was supported by AIM-AHEAD Coordinating Center, award number OTA-21-017, and was, in part, funded by the National Institutes of Health Agreement No. 1OT2OD032581.

Classical analytical methods, such as FBP [6] and FDK [8], are computationally efficient but prone to streaking artifacts under sparse sampling [5]. Iterative approaches like ART [9], SART [1], and MBIR [17] offer improved robustness, but degrade under extreme sparsity. More recent deep learning methods [7, 12] use data-driven priors to enhance reconstruction, yet face challenges of high memory usage, limited resolution, and poor generalization across anatomies. Implicit neural representations (INR) [15, 18, 22] provide continuous modeling but require extensive computation and long optimization times.

Gaussian splatting (3DGS) [10, 21] is a newly emerging technique for continuous 3D representation, offering compact storage, real-time rendering, and high-fidelity volumetric modeling. Its explicit yet continuous parameterization makes it attractive for CT reconstruction, as it balances memory efficiency with real-time optimization. Several works have extended Gaussian splatting to sparse-view CT, including 3DGR-CT[14], DIF-Gaussian [16], and R^2-Gaussian [23]. While these methods demonstrate promising reconstructions, 3DGR-CT and DIF-Gaussian require hours of training, whereas R^2-Gaussian frequently exhibits needle-like artifacts under severe sparsity.

In this paper, we address the limitations of existing Gaussian-based CT reconstruction with two physics-inspired adaptations. First, we introduce a structure-aware initialization that uses gradient and density cues from preliminary reconstructions to place Gaussians in anatomically meaningful regions, reducing artifacts and improving convergence. Second, we propose a log-domain optimization grounded in the Beer–Lambert law, which stabilizes training and ensures attenuation consistency by avoiding vanishing gradients in dense regions. Together, these contributions directly tackle the problem of needle-like artifacts and lead to more accurate, stable, and efficient reconstructions under sparse-view conditions.

The main contributions of this paper are:

- A structure-aware initialization that uses gradient and density cues from preliminary reconstructions to place Gaussians in structurally meaningful regions, reducing

needle-like artifacts and enabling faster convergence.

- A physics-inspired log-domain optimization based on the Beer–Lambert law, which stabilizes training, avoids vanishing gradients in dense regions, and enforces attenuation consistency in sparse-view CT.
- Extensive experiments on chest, jaw, and head CT datasets demonstrate that our method achieves the best performance among reconstruction techniques with comparable runtimes, consistently improving both quality (PSNR/SSIM) and efficiency.

2. Related Work

Analytical methods such as Filtered Back Projection (FBP) [6] and the Feldkamp–Davis–Kress (FDK) algorithm [8] reconstruct volumes using direct inversion formulas. While computationally efficient, they are highly sensitive to sparse sampling and often produce streaking and noise artifacts [5]. Iterative approaches, including ART [9], SART [1], and MBIR [17], improve noise handling but remain computationally expensive and degrade under extremely sparse-view conditions due to their reliance on dense angular sampling and explicit discretization.

Neural networks have shown strong performance in sparse-view CT, leveraging priors learned from large datasets [7, 12]. However, voxel-based CNNs suffer from prohibitive memory costs, which limits achievable resolution, while implicit neural representations (INRs) such as neural attenuation fields (NAF) [22], deep intensity fields [15], and neural representations with priors [18] achieve continuous representations but require long optimization times and heavy computation. Although powerful, these methods struggle with generalization to unseen anatomies or acquisition protocols and are less practical in resource-constrained scenarios.

Building on 3DGS for novel view synthesis [10], several methods adapt Gaussian primitives to CT reconstruction. 3DGR-CT [14] introduced volumetric Gaussians initialized with FBP and iteratively refined by comparing rendered and measured projections. Although effective, this approach demanded a large number of primitives and high computational cost. DIF-Gaussian [16] improved spatial consistency and generalization by embedding Gaussian features in an autoencoder-based latent space, but incurred long training times. R^2-Gaussian [23] enhanced efficiency through a radiative Gaussian kernel and CUDA-based differentiable voxelizer, yet exhibited structural artifacts and volumetric inconsistencies under sparse projections. Collectively, these approaches highlight the promise of Gaussian representations for tomography but also reveal open challenges: (1) poor robustness to sparse-view inputs, (2) sensitivity to initialization, and (3) unstable optimization in high-attenuation regions.

The Beer–Lambert law and log-domain transformations have long been used in CT physics to linearize the attenuation model and mitigate gradient vanishing [13, 20]. While standard in traditional reconstruction pipelines, Gaussian splatting frameworks have yet to systematically incorporate such physics-inspired principles. In this work, we adapt these established principles to stabilize optimization in sparse-view Gaussian-based reconstruction, bridging physics-inspired modeling with modern continuous volumetric representations.

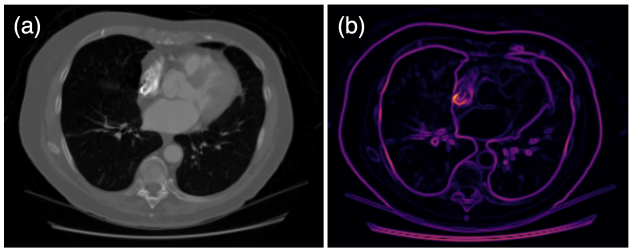


Figure 1. Visualization of structural cues in the chest dataset: (a) an axial slice from the CT volume, and (b) the corresponding axial slice of its gradient magnitude map, which highlights regions of significant density variation used for structure-aware initialization.

3. Preliminaries

Inspired by the Radiative Gaussian representation in R^2-Gaussian [23], we represent the density field of CT as a mixture of Gaussian kernels. The density field at a spatial location $\mathbf{x} \in \mathbb{R}^3$ is defined as:

$$\sigma(\mathbf{x}) = \sum_{i=1}^{M} G_i^3(\mathbf{x} \mid \rho_i, \mathbf{p}_i, \mathbf{\Sigma}_i) \tag{1}$$

where M is the number of Gaussians, ρ_i is the central density of the ith Gaussian, $\mathbf{p}_i \in \mathbb{R}^3$ represents its center position, and $\mathbf{\Sigma}_i \in \mathbb{R}^{3\times 3}$ is the covariance matrix encoding its spatial extent and anisotropy. Each Gaussian is defined as:

$$G_i^3(\mathbf{x} \mid \rho_i, \mathbf{p}_i, \mathbf{\Sigma}_i) = \rho_i \exp\left[-\tfrac{1}{2}(\mathbf{x}-\mathbf{p}_i)^\top \mathbf{\Sigma}_i^{-1}(\mathbf{x}-\mathbf{p}_i)\right] \tag{2}$$

This formulation provides a flexible and continuous parameterization of the density field, making it well-suited for tomographic reconstruction.

4. Method

Our end-to-end pipeline for sparse-view CT reconstruction (Figure 2) comprises two key stages: a structure-aware initializer that guides Gaussian placement and a log-based optimizer that stabilizes training and enforces attenuation consistency. These modules address volumetric inconsistencies and needle-like artifacts common in sparse-view settings. First, we estimate an initial CT volume from the input projections using SART [1]. We then compute its gra-

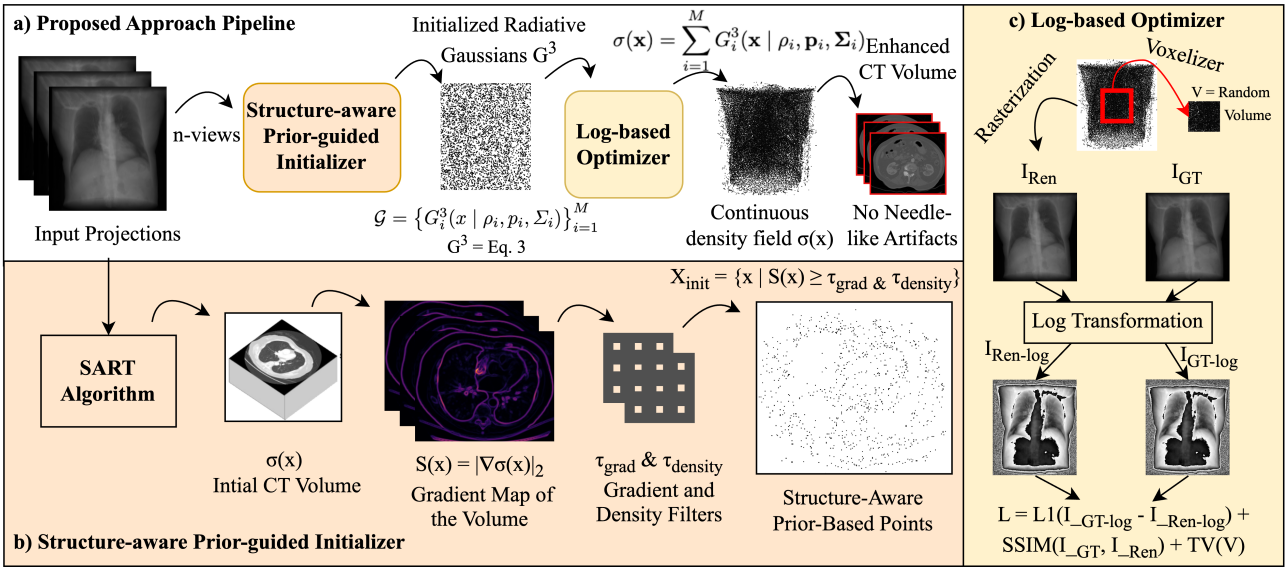


Figure 2. Overview of the proposed X-ray CT reconstruction pipeline. (a) Pipeline architecture, highlighting key components. (b) Structure-aware initialization, which extracts structural information from input projections to enhance reconstruction in sparse-view settings. (c) Log-based optimization, which improves reconstruction accuracy. These innovations enhance volumetric consistency, improve structural fidelity, and effectively eliminate needle-like artifacts, addressing the challenges of sparse-view CT reconstruction.

dient magnitude to identify high-contrast regions and apply gradient and density filters to extract structurally salient points. These points guide the placement of Gaussian centers, yielding a compact and informed initialization of the radiative Gaussian primitives. In the optimization phase, we represent the volume as a continuous mixture of Gaussians, each parameterized by position, density, scale, and orientation. To improve gradient stability in highly attenuated regions, we apply a log-based transformation to the rendered intensities. This transformation enhances numerical conditioning and enforces attenuation consistency. Overall, our pipeline combines structure-aware initialization with log-based regularization to preserve anatomical detail and suppress directional artifacts under severe view sparsity.

4.1. Structure-Aware Initialization

To initialize the reconstruction, we define a preliminary volumetric density field $\sigma(\mathbf{x})$ using SART [2], which takes approximately one minute to complete. We use SART over the FDK algorithm due to its demonstrated superior performance in sparse-view reconstruction scenarios.

Next, we compute the gradient magnitude of this density field to detect high-contrast regions. If $\sigma(\mathbf{x})$ represents the density at a specific location $\mathbf{x} = (x, y, z)$, we can represent the gradient computation as:

$$\nabla\sigma(\mathbf{x}) = \left(\frac{\partial\sigma}{\partial x}, \frac{\partial\sigma}{\partial y}, \frac{\partial\sigma}{\partial z}\right) \tag{3}$$

The gradient magnitude function quantifies the rate of change in density and is computed using equation 4, where $S(\mathbf{x})$ is a structural saliency measure, highlighting regions of significant density variation as shown in Figure 1.

$$S(\mathbf{x}) = \|\nabla\sigma(\mathbf{x})\|_2 = \sqrt{\left(\frac{\partial\sigma}{\partial x}\right)^2 + \left(\frac{\partial\sigma}{\partial y}\right)^2 + \left(\frac{\partial\sigma}{\partial z}\right)^2} \tag{4}$$

To identify the most structurally important regions, we apply a threshold τ_{grad}, defining the set of selected locations ($\mathcal{X}_{\text{grad}}$) as:

$$\mathcal{X}_{grad} = \{\mathbf{x} \mid S(\mathbf{x}) \geq \tau_{grad}\} \tag{5}$$

While gradient-based detection identifies structurally significant regions, a refined representation is needed for accurate density modeling. After identifying the most structurally significant regions in the volumes, we refine the selection by filtering the points based on their density. Specifically, we retain points within a specified density range ($\mathcal{X}_{\text{den}}$), given by:

$$\mathcal{X}_{den} = \{\mathbf{x} \in \mathcal{X}_{grad} \mid \rho(\mathbf{x}) \geq \tau_{density}\}. \tag{6}$$

To obtain the final set of Gaussian centers, we randomly sample M points from $\mathcal{X}_{den}$, where $M = 50,000$:

$$\mathcal{X}_{final} = \text{RandomSample}(\mathcal{X}_{den}, M) \tag{7}$$

where, $\text{RandomSample}(\mathcal{X}_{den}, M)$ denotes a uniform random selection of M points from the set $\mathcal{X}_{den}$. Each selected point $\mathbf{x}_i \in \mathcal{X}_{final}$ is then used to define a Gaussian primitive, forming the density $\sigma(\mathbf{x}) = \sum_{i=1}^{M} G_i^3(\mathbf{x} \mid \rho_i, \mathbf{p}_i, \mathbf{\Sigma}_i)$ as shown in equation (1).

Table 1. Performance on the **Head dataset [11]** under sparse-view settings. Each entry reports mean ± std over 3 runs. Metrics include reconstruction quality (SSIM, PSNR) and efficiency (rendering time, Gaussian count) for 13 and 25 views.

Method	13 Views				25 Views			
	SSIM ↑	PSNR ↑	Time ↓	Gaussians ↓	SSIM ↑	PSNR ↑	Time ↓	Gaussians ↓
FDK [8]	0.116 ± 0.000	19.341 ± 0.000	–	–	0.204 ± 0.000	22.999 ± 0.000	–	–
SART [1]	0.757 ± 0.000	29.523 ± 0.000	00:43 ± 00:01	–	0.788 ± 0.000	31.731 ± 0.000	01:23 ± 00:01	–
CGLS [19]	0.522 ± 0.000	27.498 ± 0.000	00:07 ± 00:00	–	0.567 ± 0.000	29.559 ± 0.000	00:08 ± 00:01	–
ASD_POCS [20]	0.876 ± 0.000	30.842 ± 0.000	01:07 ± 00:01	–	0.937 ± 0.000	34.731 ± 0.000	01:45 ± 00:01	–
R^2-Gaussian [23]	0.895 ± 0.001	31.517 ± 0.006	02:45 ± 00:02	57,749 ± 94	0.948 ± 0.001	36.062 ± 0.010	02:47 ± 00:01	58,118 ± 51
Ours	0.914 ± 0.000	32.286 ± 0.002	02:21 ± 00:00	54,052 ± 184	0.961 ± 0.000	37.065 ± 0.002	02:22 ± 00:01	53,014 ± 19

4.2. Log-Based Optimization

X-ray image formation follows the Beer–Lambert law, which models exponential attenuation of photons along a projection path:

$$I = I_0 \exp\Big(- \int \mu(x)\, ds \Big) \tag{8}$$

where I is the measured intensity at the detector, I_0 is the incident source intensity, and $\mu(x)$ is the linear attenuation coefficient at spatial location x. For notational simplicity, we denote the line integral of attenuation along the path s by $p = \int \mu(x)\, ds$. Thus, Eq. (8) can be written as

$$I = I_0 \exp(-p) \tag{9}$$

Limitations of intensity-domain optimization. Most existing Gaussian-based approaches minimize a discrepancy between rendered and ground-truth intensities, e.g.,

$$\mathcal{L}_{\text{intensity}} = |I_r - I_g| \tag{10}$$

where I_r and I_g denote rendered and ground-truth intensities, respectively. However, differentiating Eq. (9) with respect to $\mu(x)$ shows why this choice is problematic:

$$\frac{\partial I}{\partial \mu(x)} = -I \cdot \frac{\partial p}{\partial \mu(x)} \tag{11}$$

Because the gradient is scaled by I, it vanishes in strongly attenuating regions ($I \to 0$). This leads to poor supervision where it is needed most (dense tissue or bone) and biases optimization toward low-attenuation regions, ultimately producing inconsistent density estimates.

Log-domain reformulation. To mitigate this, we exploit the standard log-transformation widely used in conventional CT physics. Taking the logarithm of Eq. (9) yields:

$$\log I = \log I_0 - p \quad \Rightarrow \quad p = \log I_0 - \log I \tag{12}$$

This linearizes the exponential attenuation model into an additive form consistent with the Radon transform, thereby improving the conditioning of the forward operator. In this domain, the loss can be formulated as:

$$\mathcal{L}_{\log} = \big| \log I_r - \log I_g \big| \tag{13}$$

with gradient

$$\frac{\partial \log I}{\partial \mu(x)} = -\frac{\partial p}{\partial \mu(x)} \tag{14}$$

Crucially, the gradient is now independent of I, eliminating the vanishing-gradient problem and ensuring stable updates even in high-attenuation regions.

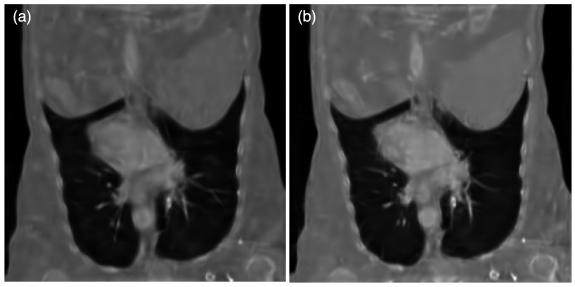


Figure 3. Qualitative comparison on the Chest dataset: (a) reconstruction using R^2-Gaussian [23], which exhibits needle-like artifacts, and (b) reconstruction from our method, where such artifacts are effectively removed, leading to a cleaner and more accurate representation.

Integration with Gaussian splatting. By incorporating this physically grounded log-domain formulation into our differentiable Gaussian rendering framework, we achieve numerically stable optimization that preserves structural fidelity while suppressing artifacts. This synergy between physics-based modeling and continuous Gaussian parameterization enables more accurate and robust reconstructions under sparse-view conditions. The optimization process minimizes the objective function $\mathcal{L}_{total}$, defined as:

$$\mathcal{L}_{\text{total}} = \lambda_{\log} \mathcal{L}_{\log}(I_r, I_g) + \lambda_{\text{ssim}} \mathcal{L}_{\text{ssim}}(I_r, I_g) + \lambda_{\text{tv}} \mathcal{L}_{\text{tv}}(V_{\text{tv}}) \tag{15}$$

where I_r and I_g denote the reconstructed and ground truth images, respectively. The log-based loss $\mathcal{L}_{\log}$ improves contrast preservation, while the structural similarity loss $\mathcal{L}_{ssim}$ enhances fine-grained detail reconstruction. The total variation regularization $\mathcal{L}_{tv}$ encourages smoothness by penalizing high-frequency noise. $\mathcal{V}_{tv}$ is a randomly selected 32×32 size volume queried through a voxelizer [23] to regularize the volumetric density field. The weighting factors

Table 2. Performance on the **Chest dataset [3]** under sparse-view settings (13 and 25 views).

Method	13 Views				25 Views			
	SSIM ↑	PSNR ↑	Time ↓	Gaussians ↓	SSIM ↑	PSNR ↑	Time ↓	Gaussians ↓
FDK [8]	0.089 ± 0.000	15.643 ± 0.002	–	–	0.156 ± 0.000	19.158 ± 0.000	–	–
SART [1]	0.672 ± 0.000	26.300 ± 0.000	00:51 ± 00:06	–	0.718 ± 0.000	28.987 ± 0.000	01:33 ± 00:05	–
CGLS [19]	0.405 ± 0.000	24.270 ± 0.000	00:13 ± 00:07	–	0.435 ± 0.000	26.307 ± 0.000	00:11 ± 00:02	–
ASD_POCS [20]	0.780 ± 0.000	27.493 ± 0.000	01:25 ± 00:18	–	0.856 ± 0.000	30.904 ± 0.000	02:01 ± 00:08	–
R^2-Gaussian [23]	0.782 ± 0.002	27.425 ± 0.014	03:04 ± 00:01	61,124 ± 59	0.876 ± 0.001	31.679 ± 0.010	03:16 ± 00:02	64,342 ± 101
Ours	0.805 ± 0.001	28.036 ± 0.014	02:27 ± 00:02	55,964 ± 436	0.893 ± 0.000	32.378 ± 0.007	02:32 ± 00:01	54,920 ± 180

Table 3. Performance on the **Jaw dataset** [11] under sparse-view settings (13 and 25 views).

Method	13 Views				25 Views			
	SSIM ↑	PSNR ↑	Time ↓	Gaussians ↓	SSIM ↑	PSNR ↑	Time ↓	Gaussians ↓
FDK [8]	0.126 ± 0.000	18.955 ± 0.000	–	–	0.237 ± 0.000	22.291 ± 0.000	–	–
SART [1]	0.604 ± 0.000	26.813 ± 0.000	00:43 ± 00:01	–	0.682 ± 0.000	29.385 ± 0.000	01:23 ± 00:01	–
CGLS [19]	0.571 ± 0.000	26.491 ± 0.000	00:07 ± 00:00	–	0.626 ± 0.000	28.682 ± 0.000	00:08 ± 00:01	–
ASD_POCS [20]	0.689 ± 0.000	28.043 ± 0.000	01:06 ± 00:01	–	0.789 ± 0.000	31.619 ± 0.000	01:44 ± 00:01	–
R^2-Gaussian [23]	0.704 ± 0.000	29.216 ± 0.003	02:18 ± 00:01	54,098 ± 22	0.818 ± 0.000	33.386 ± 0.005	02:22 ± 00:01	55,775 ± 8
Ours	0.733 ± 0.000	29.902 ± 0.001	02:18 ± 00:01	54,616 ± 35	0.834 ± 0.001	34.030 ± 0.002	02:18 ± 00:01	54,919 ± 57

Note: Each entry reports mean ± std over 3 runs. Metrics include reconstruction quality (SSIM, PSNR) and efficiency (runtime and Gaussian count).

λ_{log}, λ_{ssim}, and λ_{tv} control the relative contributions of each term in the optimization process.

5. Experiments

5.1. Dataset

We evaluate our approach on three CT datasets: a chest CT from LIDC-IDRI [3], and jaw and head CTs from SciVis [11]. To ensure consistency across datasets, we apply a preprocessing pipeline that includes density normalization and volume resizing. Specifically, we normalize densities to $[0, 1]$ and rescale volumes to $256 \times 256 \times 256$ voxels. Projection data is generated using the TIGRE tomography toolbox [4], with a 512×512 detector over a $360°$ rotation. To simulate realistic conditions, we introduce Gaussian noise $\mathcal{N}(0, 10^2)$ for detector and Poisson noises with parameter $\lambda = 10^5$ for photon scattering.

5.2. Evaluation Metrics

To comprehensively assess the performance of our method, we report both quality and efficiency metrics. For reconstruction quality, we evaluate the Peak Signal-to-Noise Ratio (PSNR) computed over the entire 3D volume, and the Structural Similarity Index Measure (SSIM), averaged across 2D slices in the axial, coronal, and sagittal planes, to capture both global fidelity and perceptual consistency. For efficiency, we record the total rendering time required for reconstruction and the number of Gaussians used in the final representation, which jointly reflect the computational and memory cost of the method.

5.3. Implementation Details

Our PyTorch implementation runs on an NVIDIA RTX 6000 Ada GPU with CUDA acceleration. We train the model for 10e3 iterations, optimizing Gaussian parameters with initial learning rates: position ($2e-4$), density ($1e-2$), scale ($5e-3$), and rotation ($1e-3$), each decaying to 10% of its initial values. The loss function combines log attenuation loss ($\lambda_{log} = 0.1$ for 25-views and 0.05 for 13-views), SSIM ($\lambda_{SSIM} = 0.25$), and total variation ($\lambda_{TV} = 0.05$). To improve initialization, we apply a density threshold of 0.05 and dynamically compute the gradient threshold based on the 70^{th}(top 30% of gradients) of the gradient map. The TV volume is used to regularize smoothness, with adaptive control refining Gaussian updates from iteration 500 to 10,000.

6. Results

We evaluate our method against five established reconstruction techniques, FDK [8], SART [1], CGLS [19], ASD_POCS [20], and R^2-Gaussian [23], using the Chest, Jaw, and Head CT datasets. These baselines were selected because they operate in the range of minutes per reconstruction, allowing for a fair comparison in terms of both quality and efficiency. In contrast, NeRF-based and other deep neural network approaches typically require hours for training making them impractical for direct comparison in this setting.

Tables 1, 2, and 3 summarize the quantitative performance of all methods on the Head, Chest, and Jaw datasets, respectively, under sparse-view reconstruction. Each table reports reconstruction quality (SSIM, PSNR) and efficiency

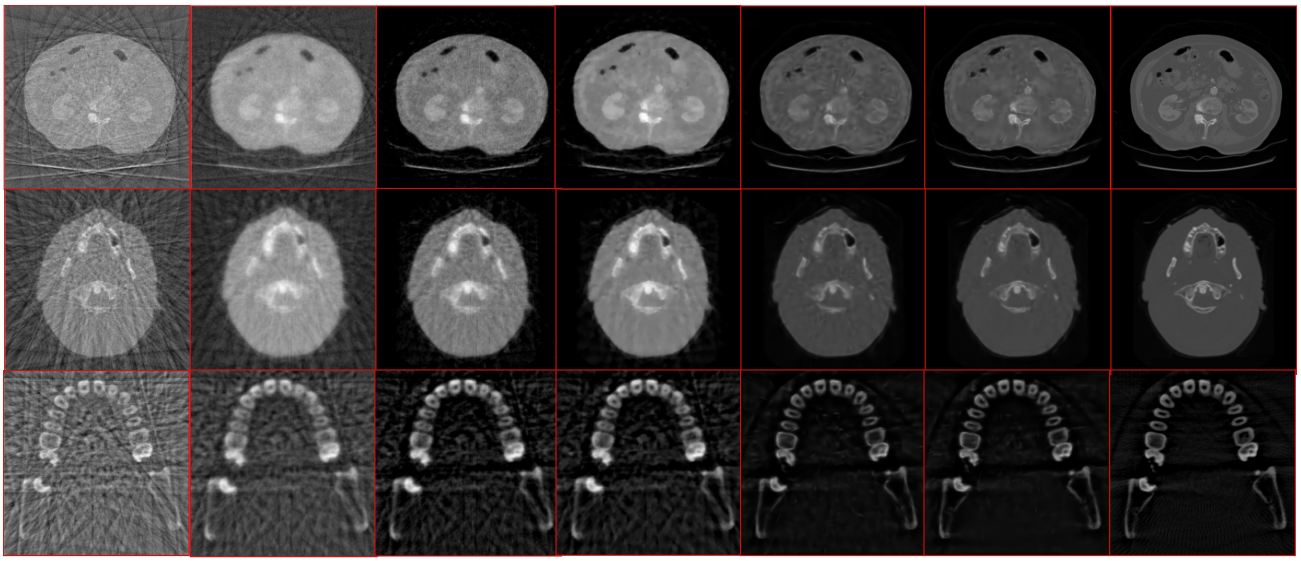


Figure 4. **Qualitative comparison of reconstructed CT slices (25-views) across different methods.** Columns represent reconstruction techniques: FDK, CGLS, SART, ASD-POCS, R^2-Gaussian, and Ours, with Ground Truth shown last. Rows correspond to different datasets: Chest (top), Head (middle), and Jaw (bottom).

(rendering time, Gaussian count) averaged over three runs. Across all datasets and view settings, our method consistently achieves the highest reconstruction fidelity while maintaining competitive efficiency.

Across Head, Chest, and Jaw (Tables 1–3), our method improves reconstruction quality and efficiency. At 13 views, average gains are 2.3% PSNR and 3.1% SSIM, with runtime reduced by 11.6% and primitives by 4.6% (approx). At 25 views, we obtain 2.3% PSNR and 1.8% SSIM improvements, while cutting runtime by 13.4% and primitives by 8.3% (approx). Peak gains reach +2.8% PSNR (Head, 25 views) and +4.1% SSIM (Jaw, 13 views), with efficiency benefits up to –22.5% runtime and –14.6% primitives (Chest, 25 views). These percentage improvements highlight the effectiveness of the structure-aware initialization and physics-inspired log formulation under severe sparsity.

Qualitative results support the quantitative findings. Figure 3 presents a direct comparison on the Chest dataset, where reconstructions with R^2-Gaussian exhibit needle-like artifacts in high-contrast regions. Our method eliminates these artifacts, resulting in smoother and more anatomically coherent reconstructions. Figure 4 further illustrates representative axial slices reconstructed from 25 views across the three datasets. Analytic baselines such as FDK, CGLS, and SART produce blurry or noisy outputs, while R^2-Gaussian improves sharpness at the cost of spurious artifacts. In contrast, our method produces sharper and more faithful reconstructions that closely resemble the ground truth.

These results demonstrate that the proposed structure-aware initialization combined with density-based refinement leads to improved reconstruction quality, effective suppression of artifacts, and reduced model complexity across diverse datasets and sparse-view conditions.

7. Ablation Study

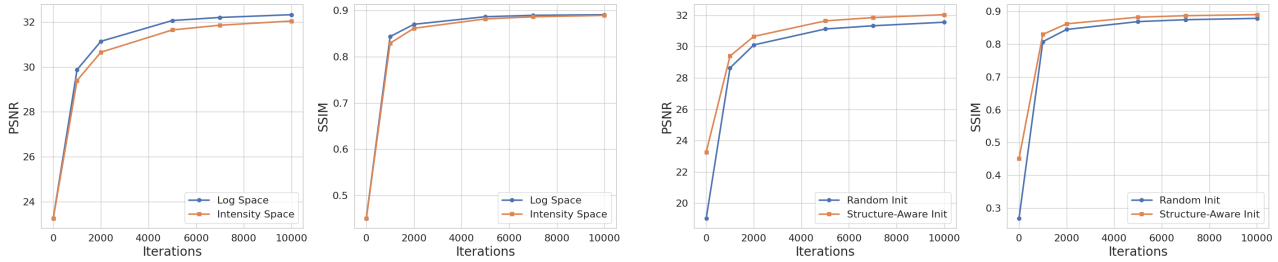
To evaluate the effectiveness of the individual components and hyperparameters in our proposed framework, we conduct a series of ablation experiments. These experiments analyze the contributions of the log-based supervision, initialization strategy, and associated thresholds. Each experiment isolates a specific variable while keeping all other components fixed, allowing us to assess its direct impact on reconstruction quality.

7.1. Intensity-Space vs. Log-Space

We compared our log-space loss formulation with a conventional L_1 loss applied in intensity space. As shown in Figure 5(a), log-based optimization consistently results in higher PSNR and SSIM throughout the training process. The improvements are particularly evident in the early and intermediate stages, indicating better convergence behavior due to stabilized gradient flow.

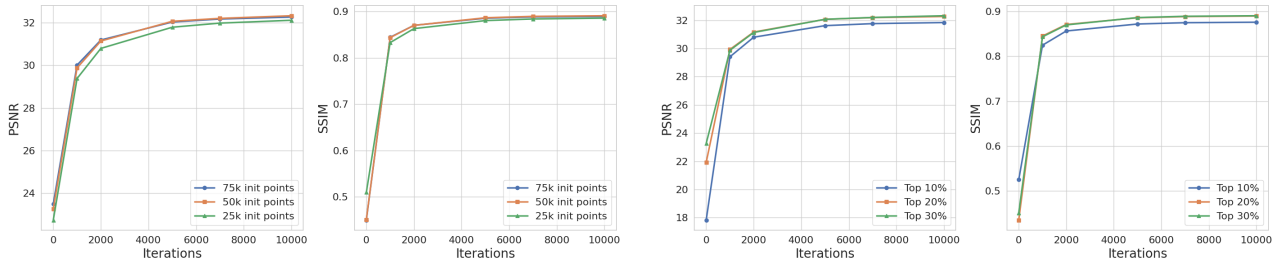
7.2. Random vs. Structure-Aware Initialization

We compare our structure-aware initialization against a random initialization strategy for placing Gaussian primitives. As shown in Figure 5(b), the structure-aware method



(a) (b)

Figure 5. (a) **Intensity-Space vs. Log-Space Optimization**, Log-space optimization outperforms intensity-space, offering faster convergence and higher final PSNR/SSIM. (b) **Random vs. Structure-Aware Initialization**, Structure-aware initialization achieves better reconstruction quality than random Gaussian initialization by focusing on structurally relevant regions.



(a) (b)

Figure 6. **(a) Effect of Initialization Point Count**, Using 50k initialization points yields the best balance between performance and efficiency, with higher or lower counts offering diminishing returns. **(b) Effect of Gradient Thresholding**, Selecting the top 30% of gradient magnitudes leads to the most effective reconstructions, indicating the importance of including moderately strong structural cues.

achieves consistently better PSNR and SSIM throughout training. It also converges faster due to its focus on structurally relevant regions. This confirms that incorporating structure information during initialization leads to more effective and accurate reconstruction compared to random sampling.

7.3. Effect of Initialization Point Count

We evaluate the impact of the number of initial Gaussian primitives by testing 25k, 50k, and 75k points. As illustrated in Figure 6(a), using 50k points consistently yields the best reconstruction quality across training steps, achieving the highest PSNR and SSIM.

Using fewer points (25k) results in lower early performance and slower convergence, likely due to insufficient spatial coverage and structural resolution. Increasing to 75k offers only marginal improvement over 50k and, in some cases, leads to minor degradation, possibly due to overparameterization or increased optimization complexity.

These results suggest that 50k Gaussians strike an optimal balance between model capacity and training efficiency, providing sufficient structural representation without incurring unnecessary computational overhead.

7.4. Effect of Gradient Thresholding

We examine the impact of different gradient thresholds used to select structurally salient regions for Gaussian initialization. Specifically, we test thresholds corresponding to the top 10%, 20%, and 30% of gradient magnitudes. The resulting PSNR and SSIM scores are presented in Figure 6(b).

Using a tighter threshold (top 10%) restricts initialization to only the most prominent edges, which leads to suboptimal performance. These settings miss out on broader structural cues necessary for smooth reconstruction. In contrast, selecting the top 30% provides a more balanced structural awareness, yielding the highest performance across both metrics.

8. Conclusion

This paper addresses the challenge of needle-like artifacts in sparse-view CT reconstruction within the Gaussian splatting framework. While Gaussian splatting is efficient for real-time reconstruction, its performance degrades under high sparsity due to poor initialization and unstable optimization. To overcome this, we propose a structure-aware initialization that places Gaussians in structurally significant regions, and a physics-inspired log-domain optimization based on the Beer–Lambert law that stabilizes train-

ing and enforces attenuation consistency. Together, these advances lead to sharper, artifact-free reconstructions with faster convergence, consistently outperforming existing approaches across multiple CT datasets.

References

[1] Anders H Andersen and Avinash C Kak. Simultaneous algebraic reconstruction technique (sart): a superior implementation of the art algorithm. *Ultrasonic imaging*, 6(1):81–94, 1984. 1, 2, 4, 5

[2] Ria Aprilliyani, Rian Gilang Prabowo, and Basari Basari. On the performance of sart and art algorithms for microwave imaging. In *AIP Conference Proceedings*. AIP Publishing, 2018. 3

[3] Samuel G Armato III, Geoffrey McLennan, Luc Bidaut, Michael F McNitt-Gray, Charles R Meyer, Anthony P Reeves, Binsheng Zhao, Denise R Aberle, Claudia I Henschke, Eric A Hoffman, et al. The lung image database consortium (lidc) and image database resource initiative (idri): a completed reference database of lung nodules on ct scans. *Medical physics*, 38(2):915–931, 2011. 5

[4] Ander Biguri, Manjit Dosanjh, Steven Hancock, and Manuchehr Soleimani. Tigre: a matlab-gpu toolbox for cbct image reconstruction. *Biomedical Physics & Engineering Express*, 2(5):055010, 2016. 5

[5] Philippe P Bruyant, Jacques Sau, and Jean-Jacques Mallet. Streak artifact reduction in filtered backprojection using a level line–based interpolation method. *Journal of Nuclear Medicine*, 41(11):1913–1919, 2000. 1, 2

[6] Nabil Chetih and Zoubeida Messali. Tomographic image reconstruction using filtered back projection (fbp) and algebraic reconstruction technique (art). In *2015 3rd International Conference on Control, Engineering & Information Technology (CEIT)*, pages 1–6. IEEE, 2015. 1, 2

[7] Zhiming Cui, Yu Fang, Lanzhuju Mei, Bojun Zhang, Bo Yu, Jiameng Liu, Caiwen Jiang, Yuhang Sun, Lei Ma, Jiawei Huang, et al. A fully automatic ai system for tooth and alveolar bone segmentation from cone-beam ct images. *Nature communications*, 13(1):2096, 2022. 1, 2

[8] Lee A Feldkamp, Lloyd C Davis, and James W Kress. Practical cone-beam algorithm. *Journal of the Optical Society of America A*, 1(6):612–619, 1984. 1, 2, 4, 5

[9] Richard Gordon, Robert Bender, and Gabor T Herman. Algebraic reconstruction techniques (art) for three-dimensional electron microscopy and x-ray photography. *Journal of theoretical Biology*, 29(3):471–481, 1970. 1, 2

[10] Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 3d gaussian splatting for real-time radiance field rendering. *ACM Trans. Graph.*, 42(4):139–1, 2023. 1, 2

[11] Pavel Klacansky. Open scivis datasets, 2024. Accessed: 2025-02-27. 4, 5

[12] Daeun Kyung, Kyungmin Jo, Jaegul Choo, Joonseok Lee, and Edward Choi. Perspective projection-based 3d ct reconstruction from biplanar x-rays. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1–5. IEEE, 2023. 1, 2

[13] Yinsheng Li, Ke Li, Chengzhu Zhang, Juan Montoya, and Guang-Hong Chen. Learning to reconstruct computed tomography images directly from sinogram data under a variety of data acquisition conditions. *IEEE transactions on medical imaging*, 38(10):2469–2481, 2019. 1, 2

[14] Yingtai Li, Xueming Fu, Shang Zhao, Ruiyang Jin, and S Kevin Zhou. Sparse-view ct reconstruction with 3d gaussian volumetric representation. *arXiv preprint arXiv:2312.15676*, 2023. 1, 2

[15] Yiqun Lin, Zhongjin Luo, Wei Zhao, and Xiaomeng Li. Learning deep intensity field for extremely sparse-view cbct reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 13–23. Springer, 2023. 1, 2

[16] Yiqun Lin, Hualiang Wang, Jixiang Chen, and Xiaomeng Li. Learning 3d gaussians for extremely sparse-view cone-beam ct reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 425–435. Springer, 2024. 1, 2

[17] Lu Liu. Model-based iterative reconstruction: a promising algorithm for today's computed tomography imaging. *Journal of Medical imaging and Radiation sciences*, 45(2):131–136, 2014. 1, 2

[18] Liyue Shen, John Pauly, and Lei Xing. Nerp: implicit neural representation learning with prior embedding for sparsely sampled image reconstruction. *IEEE Transactions on Neural Networks and Learning Systems*, 35(1):770–782, 2022. 1, 2

[19] Jonathan Richard Shewchuk et al. An introduction to the conjugate gradient method without the agonizing pain. 1994. 4, 5

[20] Emil Y Sidky and Xiaochuan Pan. Image reconstruction in circular cone-beam computed tomography by constrained, total-variation minimization. *Physics in Medicine & Biology*, 53(17):4777, 2008. 1, 2, 4, 5

[21] Aqsa Yousaf, Arkajyoti Mitra, Paul Agbaje, Afia Anjum, and Habeeb Olufowobi. Daps-agf: Depth-aware perceptual similarity with adaptive gradient filtering for enhanced outdoor scene reconstruction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2648–2656, 2025. 1

[22] Ruyi Zha, Yanhao Zhang, and Hongdong Li. Naf: neural attenuation fields for sparse-view cbct reconstruction. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 442–452. Springer, 2022. 1, 2

[23] Ruyi Zha, Tao Jun Lin, Yuanhao Cai, Jiwen Cao, Yanhao Zhang, and Hongdong Li. R^2-gaussian: Rectifying radiative gaussian splatting for tomographic reconstruction. *arXiv preprint arXiv:2405.20693*, 2024. 1, 2, 4, 5