

Geometry-Aware Gaussian Splat Fields for Sparse-View Tomographic Reconstruction

Abstract

Recent advances in 3D Gaussian splatting (3DGS) have significantly improved speed and quality in tasks such as surface reconstruction and view synthesis. While 3DGS has been extended to sparse-view tomographic scenarios, existing methods face challenges in accurately reconstructing volumetric data due to limitations in handling noisy point-cloud initialization and capturing complex details. In this paper, we propose Geometry-Aware Gaussian Splat Fields, which introduce two key innovations to address these issues. First, we incorporate a spatial autocorrelation bias to establish feature relationships among neighboring points, mitigating the impact of imprecise initialization. Second, to enforce the 3D scene geometry constraint during deformation, we explicitly extract 3D geometry features and integrate them in learning the 3D deformation. Experiments on synthetic and real-world datasets demonstrate that our method outperforms state-of-the-art approaches in accuracy and efficiency. Code and models are available on the project page https://github.com/zyuxiang0822/geo_gaussian.

Keywords: sparse-view tomographic reconstruction, spatial autocorrelation bias, 3D geometry features.

1 Introduction

Computed tomography (CT) is an essential imaging technique for noninvasively examining the internal structure of objects. Most CT systems rely on X-rays as the imaging source due to their strong penetration capabilities through solid materials [19]. During a CT scan, an X-ray machine captures multi-angle 2D projections, which measure the attenuation of rays passing through the material. The core objective of CT, tomographic reconstruction, is to recover the 3D density field of the object from these projections. This task poses two significant challenges, each rooted in the practical constraints and application demands of CT technology. Firstly, the harmful nature of X-ray radiation limits the acquisition of sufficient and noise-free projections, making reconstruction a complex and inherently ill-posed problem. Secondly, time-sensitive applications, such as medical diagnosis and industrial inspection, demand algorithms that can deliver results promptly and efficiently.

Existing tomography methods face challenges in balancing reconstruction quality and processing speed. Traditional CT algorithms [2, 12, 48] can produce results within minutes; however, they often introduce severe artifacts due to their reliance on limited projections. Supervised learning-based approaches [8, 27, 28, 31] leverage semantic priors to achieve promising reconstruction quality but struggle when handling out-of-distribution objects, limiting their generalizability. More recently, neural radiance fields (NeRF) [38] have been adapted

for tomography, demonstrating strong performance in per-case reconstruction tasks [6, 42, 46, 58, 60]. Despite their success, these methods are computationally expensive, often requiring over 30 minutes per case for conventional CT scans. For higher-resolution imaging modalities like micro-CT, the processing time can exceed 10 hours due to the extensive sampling of points required for accurate volume rendering.

Recently, 3D Gaussian splatting (3DGS) [21] has outperformed NeRF in both quality and efficiency for tasks such as view synthesis [26, 34, 56] and surface reconstruction [15, 17, 57]. However, its application to volumetric reconstruction—particularly for X-ray tomography—remains limited and largely ineffective. While some concurrent works [5, 57] adapt 3DGS for X-ray view synthesis, they primarily use it to augment training data rather than to obtain accurate 3D density volumes. To address these shortcomings, R^2 -Gaussian [59] rectifies the integration bias in 3DGS and introduces a CUDA-based differentiable voxelizer for direct volumetric reconstruction, achieving both high efficiency and reconstruction quality. Nonetheless, our empirical assessment of methods for CT reconstruction demonstrates that its performance still falls short of expectations, especially as the object becomes sparser and the number of projections decreases. We identify the root cause to be the degradation of the initialized Gaussian centers, which rely heavily on point cloud data derived from FDK [12]. In scenarios with fewer projections, noise in the FDK-generated point cloud rises sharply, leading to inaccurate Gaussian parameter initialization and ultimately degrading the final reconstruction quality. Furthermore, to achieve geometrically consistent 3D volume, it is vital to preserve local geometric or structural information, as real-world object deformations are closely tied to their underlying 3D structures.

In this paper, we propose geometry-aware gaussian splat fields, an optimization strategy that effectively regularizes splat features for Sparse-View Tomographic Reconstruction. Our main contributions are as follows: (1) We propose a novel optimization strategy, which introduces spatial bias into the 3D Gaussian Splatting technique to stabilize the optimization process under sparse views. (2) We introduced 3D sparse convolution to extract local structural information effectively and efficiently for deformation learning. (3) We conduct comprehensive evaluations on the synthetic and real datasets, demonstrating that our method significantly outperforms state-of-the-art methods in reconstruction quality, highlighting its practical value.

2 Related works

2.1 Tomographic reconstruction

Computed tomography (CT) is widely used for non-intrusive inspection in medicine [16, 20], biology [10, 22, 35], and industry [9, 11]. Conventional fan-beam CT produces a 3D volume by reconstructing each slice from 1D projection arrays. Recently, the cone-beam scanner has become popular for its fast scanning and high resolution [45], leading to the demand for 3D tomography, i.e., recovering the volume directly from 2D projection images. Our work focuses on 3D sparse-view reconstruction where less than a hundred projections are captured to reduce radiation exposure. Traditional algorithms are mainly grouped into analytical and iterative methods. Analytical methods like filtered back projection (FBP) and its 3D variant FDK [12] produce results instantly by solving the Radon transform and its inverse [41]. However, they introduce serious streak artifacts in sparse-view scenarios. Iterative methods [2, 37, 44, 48] formulate tomography as a maximum-a-posteriori problem and iteratively minimize the energy function with regularizations. They successfully suppress artifacts but take longer time and lose structure details. Deep learning methods can be

categorized as supervised and self-supervised families. Supervised methods learn semantic priors from CT datasets. They then use the trained networks to inpaint projections [3, 14], denoise volumes [8, 25, 31, 33] or directly output results [1, 18, 27, 28, 55]. Supervised learning methods perform well in cases similar to training sets but suffer from poor generation ability when applied to unseen data. To overcome this limitation, some studies [6, 42, 46, 58, 60] handle tomography in a self-supervised learning fashion. Inspired by NeRF [38], they model the density field with coordinate-based networks and optimize them with photometric losses. Although NeRF-based methods excel in per-case reconstruction, they are time-consuming due to the extensive point sampling in volume rendering.

2.2 3DGS

3D Gaussian splatting (3DGS) [21] outperforms NeRF in speed by leveraging parallelized rasterization for image rendering, representing objects as a set of trainable Gaussian-shaped primitives. It has demonstrated notable success in numerous RGB tasks, including surface reconstruction [15, 17, 57], dynamic scene modeling [29, 52, 53], human avatars [24, 29, 32], and 3D generation [7, 49, 54]. Recently, there have also been attempts to extend 3DGS to X-ray imaging. For instance, X-Gaussian [5] adapts 3DGS to synthesize novel-view X-ray projections, while Gao et al. [13] incorporate noise-inducing physical effects into X-Gaussian. However, these methods primarily use 3DGS as a data augmentation tool for conventional tomography and do not directly tackle volumetric reconstruction. To address this gap, R^2 -Gaussian [59] extends 3DGS to sparse-view tomographic reconstruction by resolving critical limitations such as integration bias and difficulties in extracting accurate 3D density volumes. It introduces radiative Gaussian kernels tailored for X-ray imaging, rectifies the rasterization process, and employs a differentiable voxelizer for efficient density-volume retrieval. While R^2 -Gaussian improves performance in sparse-view settings compared to prior methods, its heavy reliance on projection-based representations still compresses 3D information into 2D. Consequently, it struggles to capture intricate geometric details, and reconstruction quality degrades as the number of views becomes highly limited.

In this work, we propose integrating neural networks into the splatting framework to regularize splat behavior and fully leverage 3D information. Specifically, we regress splat features from their 3D coordinates, introducing a spatial autocorrelation bias that substantially enhances reconstruction quality even in very sparse-view scenarios, as demonstrated in our experiments.

3 Method

3.1 Preliminaries of Radiative Gaussian Splatting

In the following, we provide a brief overview of the Radiative Gaussian splatting rendering technique, which is a fundamental building block of our model.

Radiative Gaussian splatting represents the target object as a group of learnable 3D Gaussian kernels $\mathcal{G}^3 = \{G_i^3\}_{i=1,\dots,M}$. Each kernel G_i^3 defines a local Gaussian-shaped density field, expressed as:

$$G_i^3(\mathbf{x} \mid \rho_i, \mathbf{p}_i, \Sigma_i) = \rho_i \cdot \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{p}_i)^\top \Sigma_i^{-1}(\mathbf{x} - \mathbf{p}_i)\right), \quad (1)$$

where ρ_i is the central density, $\mathbf{p}_i \in \mathbb{R}^3$ is the position, and $\Sigma_i \in \mathbb{R}^{3 \times 3}$ is the covariance matrix, parameterized as $\Sigma_i = \mathbf{R}_i \mathbf{S}_i \mathbf{S}_i^\top \mathbf{R}_i^\top$ with a rotation matrix \mathbf{R}_i and a scale matrix \mathbf{S}_i .

The overall density field at position $\mathbf{x} \in \mathbb{R}^3$ is computed by summing the contributions of all Gaussian kernels:

$$\sigma(\mathbf{x}) = \sum_{i=1}^M G_i^3(\mathbf{x} | \rho_i, \mathbf{p}_i, \Sigma_i). \quad (2)$$

To render a Gaussian from a given viewpoint, the 3D Gaussians are transformed into a ray space where the viewing rays are aligned with the coordinate axis. The transformation can be expressed as:

$$\mathbf{x}' = \mathbf{W}(\mathbf{x} - \mathbf{p}_i), \quad \Sigma'_i = \mathbf{W}\Sigma_i\mathbf{W}^\top, \quad (3)$$

where \mathbf{W} is the transformation matrix that aligns the viewing direction with the ray axis.

Each transformed 3D Gaussian is then projected onto the 2D image plane by dropping the third dimension. The resulting 2D Gaussian is represented as:

$$G_i^2(\mathbf{x} | \rho'_i, \mathbf{p}'_i, \Sigma'_i) = \rho'_i \cdot \exp\left(-\frac{1}{2}(\mathbf{x} - \mathbf{p}'_i)^\top \Sigma'^{-1}_i (\mathbf{x} - \mathbf{p}'_i)\right), \quad (4)$$

where the projected position \mathbf{p}'_i and covariance matrix Σ'_i are derived by excluding the third row and column of the 3D parameters. The projected density ρ'_i is corrected for integration bias as:

$$\rho'_i = \rho_i \cdot \sqrt{\frac{|2\pi\Sigma'_i|}{|\Sigma_i|}}. \quad (5)$$

The final rendered image is computed by summing the contributions of all 2D Gaussians:

$$\hat{I}(\mathbf{x}) = \sum_{i=1}^M G_i^2(\mathbf{x}). \quad (6)$$

The model is optimized by minimizing the total loss:

$$\mathcal{L}_{\text{total}} = \mathcal{L}_1(\mathbf{I}_r, \mathbf{I}_m) + \lambda_{\text{SSIM}} \mathcal{L}_{\text{SSIM}}(\mathbf{I}_r, \mathbf{I}_m) + \lambda_{\text{TV}} \mathcal{L}_{\text{TV}}(\mathbf{V}_{\text{TV}}), \quad (7)$$

where \mathcal{L}_1 is the pixel-wise loss, $\mathcal{L}_{\text{SSIM}}$ evaluates structural similarity, and \mathcal{L}_{TV} regularizes the reconstructed density volume.

3.2 Geometry-aware Gaussian Splat Fields

Limitations of 3D GS. Modeling 3D scenes with irregularly spaced point primitives offers significant flexibility and enables rapid optimization, particularly when abundant training views are available. However, in the context of CT reconstruction, which inherently relies on sparse-view data, these independent point primitives are highly susceptible to overfitting and fail to adequately capture the underlying structure of the object. Moreover, 3DGS has limited capability in extracting detailed geometric information, which is critical for modeling complex objects. To address these limitations, we propose integrating a spatial autocorrelation bias within the splats and deriving splat features and volume information through implicit neural models. This approach not only enhances the ability to encode geometric details but also provides a means to constrain and regularize the optimization process, mitigating the ill-posed nature of sparse-view reconstruction tasks.

Deep structural prior. First, we follow the idea of a deep image prior [47, 50] and aim to utilize CNNs to model locally structured patterns of splat features. In the original work [50], the CNN takes as input low-dimension Gaussian noise $\epsilon \sim \mathcal{N}(0, I)$ and gradually upsamples it into the desired image resolution; the weights of the network are then optimized to fit the observed noisy image. In our case, we aim to generate a 3D field of splat features; as 3D CNNs are computationally prohibitive, we use 2D CNNs that generate axis-aligned tri-plane representations [7, 40]. Overall, the step is a splat-based variation of the approach utilized in [47] for a fully volumetric NeRF-based sparse rendering.

Specifically, given a randomly initialized noise ϵ , the convolutional network g_θ regresses the three $H \times W$ -resolution planes \mathbf{F} :

$$\mathbf{F} = g_\theta(\epsilon) \in \mathbb{R}^{3 \times H \times W \times l}, \quad (8)$$

where l denotes the feature dimension and θ indicates learnable network weights. The overall CNN structure resembles the one originally proposed in [47].

Geometry feature learning. To capture the local geometric structure of the CT voxels, we regard the 3D Gaussian as the 3D point cloud, *i.e.*, we only use the 3D coordinates of the 3D Gaussian. By integrating the geometric branch and identity branch, we can efficiently obtain point-level features at high resolution while embedding the local geometric information of the point cloud. The geometric branch leverages the sparse convolution [30] on the sparse voxels to extract local geometric features at different receptive fields. Given the point cloud $\mathbf{P} \in \mathbb{R}^{N \times 3}$, we first transform the high-resolution point clouds into low-resolution voxels by dividing the space through a fixed grid size s :

$$\mathbf{V} = \text{floor}(\mathbf{P}/s), \quad (9)$$

where the size of \mathbf{V} is $M \times 3$ and M is the number of voxels. Then, we construct a sparse 3D U-Net by stacking a set of sparse convolutions with a skip connection. Taking \mathbf{V} as input, we perform sparse 3D U-Net to aggregate local features (dubbed as $\mathbf{F}_v \in \mathbb{R}^{M \times C}$) of the point clouds.

Geometry-aware splat fields. Next, the splat center \mathbf{p}_k is projected onto each of the three feature planes to obtain feature vectors via bilinear interpolation. These features are then concatenated along the feature dimension and denoted as $\mathbf{f}_k \in \mathbb{R}^l$. The feature and the initial point are propagated through a deformation MLP f_θ which refines the position of the input point:

$$\hat{\mathbf{p}}_k = f_\theta(\mathbf{p}_k, \mathbf{f}_k, V), \quad (10)$$

Finally, the updated point location, along with the inferred feature vector, is provided as input to a set of compact (5-6 layers, 64-128 neurons) neural fields $\{f_{\theta_p}, f_{\theta_s}, f_{\theta_\sigma}\}$ to obtain properties of Gaussian splats. The obtained splats are then rendered w.r.t. the input views to optimize the learnable modules by minimizing the photometric loss.

Splat norm regularization. For static reconstruction, we add additional norm regularization $\|\hat{\mathbf{p}}_k\|_2$ to the loss function to bias the resulting splats to not deviate significantly from the origin, similar to the floor loss considered in [36].

Geometry-aware splat fields (Fig. 1) builds on the core property of neural networks to discover local patterns and fit low frequencies of a signal first. To that end, we implement SplatFields as a neural generator that infers the attributes of Gaussian splats. The neural generator combines key properties of convolutional neural

networks, which model local structured patterns, with multi-layer perceptrons that serve as global approximators. Then We extract positional features using an MLP, as well as local geometric features using a 3D U-Net, fused by another MLP to form the geometry-aware features.

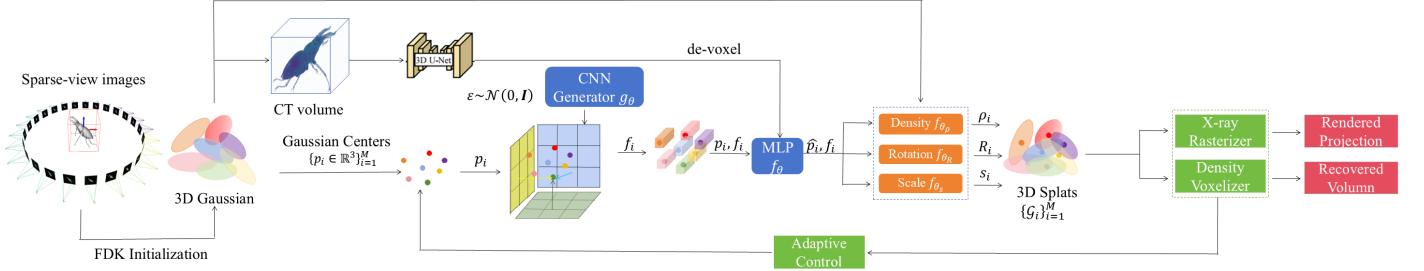


Figure 1. Overview of the method. SplatFields takes as input a point cloud (*e.g.*, initialized from FDK [12]), for which it models the geometric (position \mathbf{p}_k , scale \mathbf{s}_k , rotation \mathbf{O}_k). These attributes represent the point set as 3D splats that are then rendered with the 3DGS rasterizer. First, the point location set $\{\mathbf{p}_k \in \mathbb{R}^3\}_{k=1}^K$ is encoded into features $\{\mathbf{f}_k\}_{k=1}^K$ by sampling the tri-plane representation generated by a CNN generator g_θ to provide a deep structural prior [50]. We extract positional features using an MLP, as well as local geometric features using a 3D U-Net, fused by another MLP to form the geometry-aware features. These values are then propagated through a deformation MLP f_θ to refine the point locations $\hat{\mathbf{p}}_k$. The new point set, along with the features, is then propagated through a series of compact neural fields to predict the properties of rendering primitives $\{\mathbf{G}_k\}_{k=1}^K$ that are rendered with respect to arbitrary viewpoints.

4 Experiments

4.1 Experimental settings

Dataset. We conduct experiments on both synthetic and real-world datasets. For the synthetic dataset, we collect 20 real CT volumes, ranging from organisms to micro objects (Fig.2). We then use the tomography toolbox TIGRE [4] to synthesize X-ray projections and add Compton scatter and electric noise. For real-world experiments, we use three cases from the FIPS dataset [48], each with 721 real projections [3]. Since ground truth volumes are unavailable, we use FDK [12] to create pseudo-ground truth using all views and then subsample views for sparse-view experiments. We set 50 and 25 views for both synthetic and real-world data as two sparse-view scenarios.

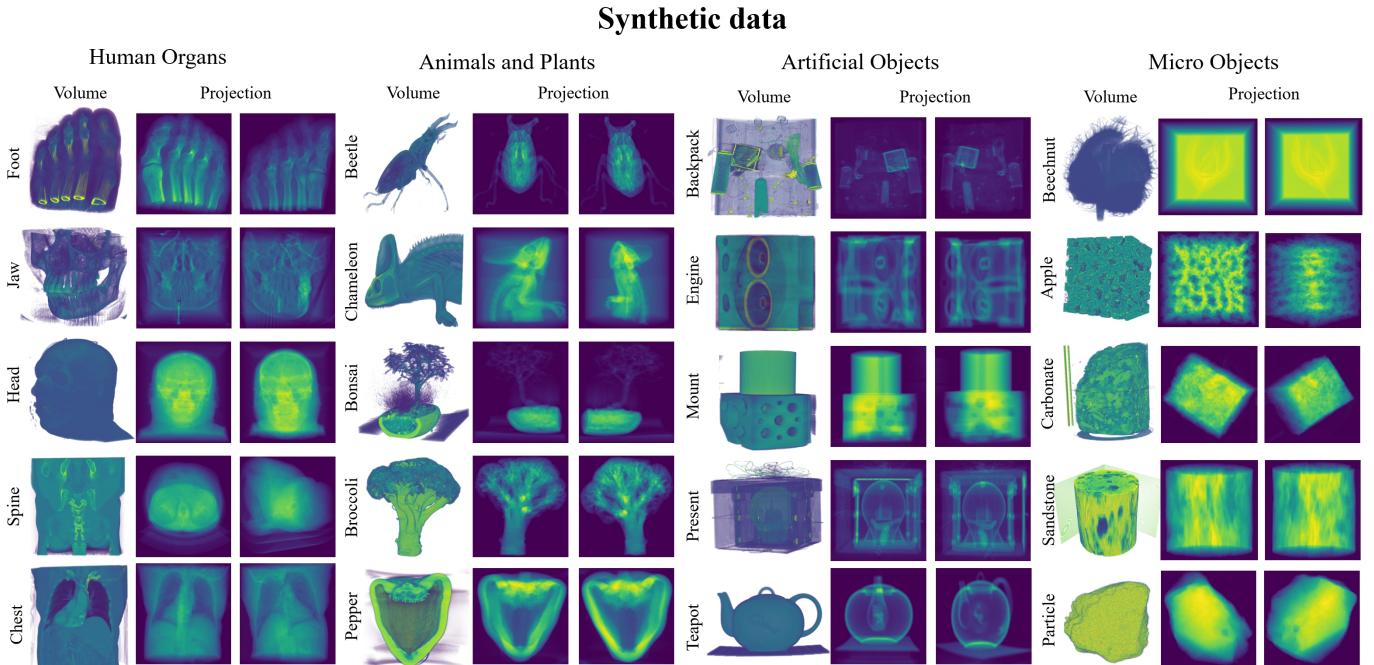


Figure 2. Synthetic datasets used for experiments. We show half volume and projection examples for each case.

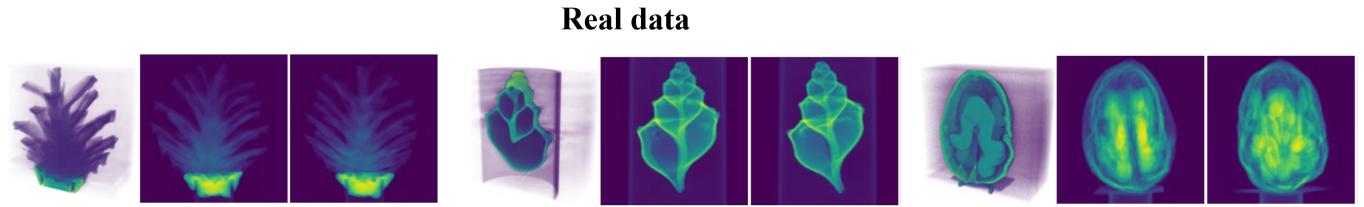


Figure 3. Real datasets used for experiments. We show half volume and projection examples for each case.

Implementation details. Our geo-Gaussian is implemented in PyTorch [39] and CUDA [43], and trained with the Adam optimizer [23] for 30k iterations. Learning rates for position, density, scale, and rotation are initially set as 0.0002, 0.01, 0.005, and 0.001, respectively, and exponentially to 0.1 of their initial values. Loss weights are $\lambda_{ssim} = 0.25$, $\lambda_{norm} = 0.01$ and $\lambda_{tv} = 0.05$. We initialize $M = 50k$ Gaussians with a density threshold $\tau = 0.05$ and scaling term $k = 0.15$. The TV volume size is $D = 32$. Adaptive control runs from 500 to 15k iterations with a gradient threshold of 0.00005. All methods run on six NVIDIA L40 GPUs. We evaluate reconstruction quality using PSNR and SSIM [51], with PSNR calculated in 3D volume and SSIM averaged over 2D slices in axial, coronal, and sagittal directions.

4.2 Results and evaluation

For fairness, we do not compare methods that require external training data but focus on those that solely use 2D projections of arbitrary objects. We compare geo-Gaussian with three traditional methods (FDK [12], SART [2], ASD-POCS [48]) and three SOTA NeRF-based methods (IntraTomo [58], NAF [60], SAX-NeRF [6]) and two gaussian-based methods (XGaussian [5], R^2 -Gaussian [59]). Tab.1 reports the quantitative results on sparse-view tomography.geo-Gaussian achieves the best performance across all synthetic and

Methods	50-view		25-view	
	PSNR↑	SSIM↑	PSNR↑	SSIM↑
Synthetic dataset				
FDK [12]	26.5	0.4222	22.99	0.317
SART [2]	34.37	0.875	31.14	0.825
ASD-POCS [48]	34.34	0.914	30.48	0.847
IntraTomo [58]	35.25	0.923	34.68	0.914
NAF [60]	36.65	0.932	33.91	0.893
NeRF [38]	36.86	0.938	34.13	0.901
SAX-NeRF [6]	37.63	0.949	34.33	0.906
XGaussian [5]	37.23	0.941	35.09	0.922
R^2 -Gaussian [59]	37.98	0.952	35.19	0.923
Ours	38.71	0.967	36.21	0.941
Real-world dataset				
FDK [12]	27.38	0.449	23.3	0.335
SART [2]	33.61	0.827	31.52	0.790
ASD-POCS [48]	34.58	0.861	31.32	0.810
IntraTomo [58]	36.99	0.854	35.85	0.835
NAF [60]	36.44	0.818	32.92	0.772
NeRF [38]	36.63	0.822	33.25	0.780
SAX-NeRF [6]	34.89	0.840	33.49	0.793
XGaussian [5]	34.52	0.821	34.83	0.833
R^2 -Gaussian [59]	37.52	0.866	35.10	0.840
Ours	38.24	0.864	35.91	0.851

Table 1. Comparison of different methods on synthetic and real-world datasets.

most real-world experiments. Specifically, our method delivers a 0.93 dB higher PSNR than R^2 -Gaussian, on the synthetic dataset, and a 0.95 dB improvement over IntraTomo on the real-world dataset. Fig.4 and Fig.5 show the visual comparisons of different methods. FDK and SART introduce streak artifacts, while ASD-POCS and IntraTomo blur structural details. NAF and SAX-NeRF are better than other baseline methods but have salt-and-pepper noise. In comparison, our method successfully recovers sharp details.

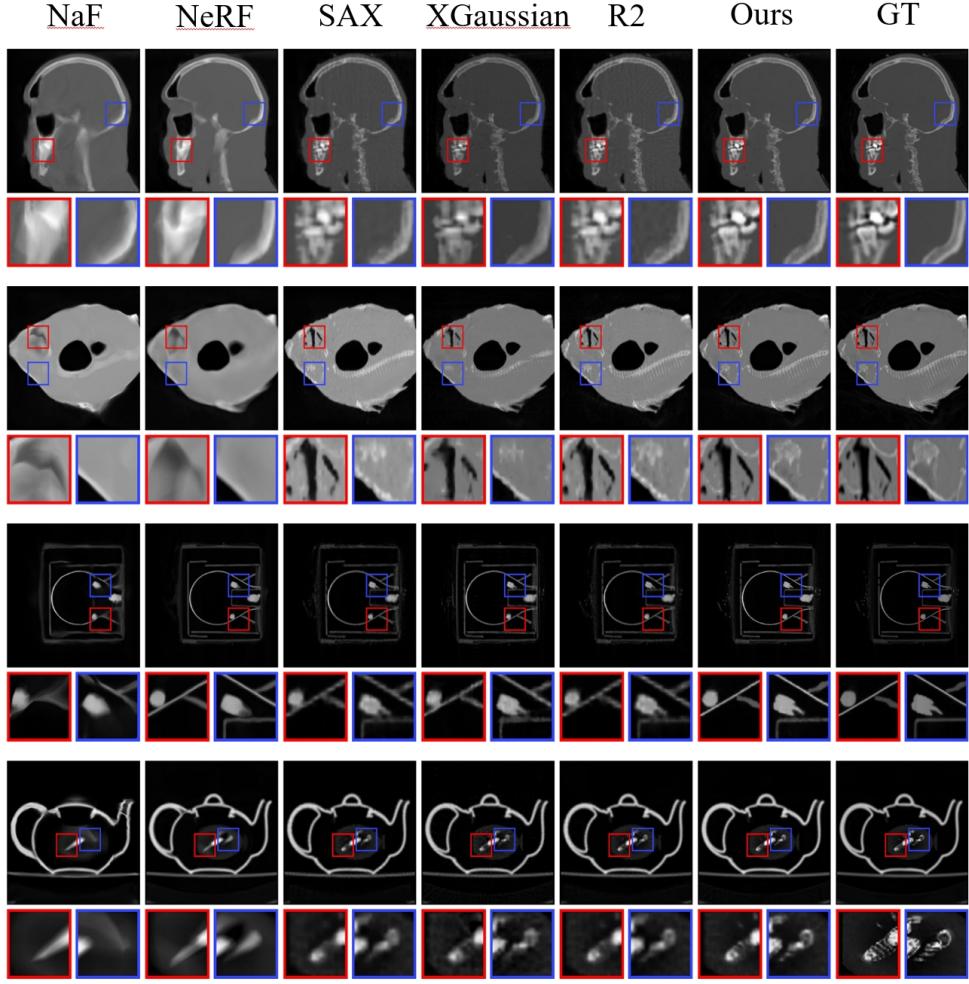


Figure 4. Visual results of novel view synthesis on the scenes. Please zoom in for a better view.

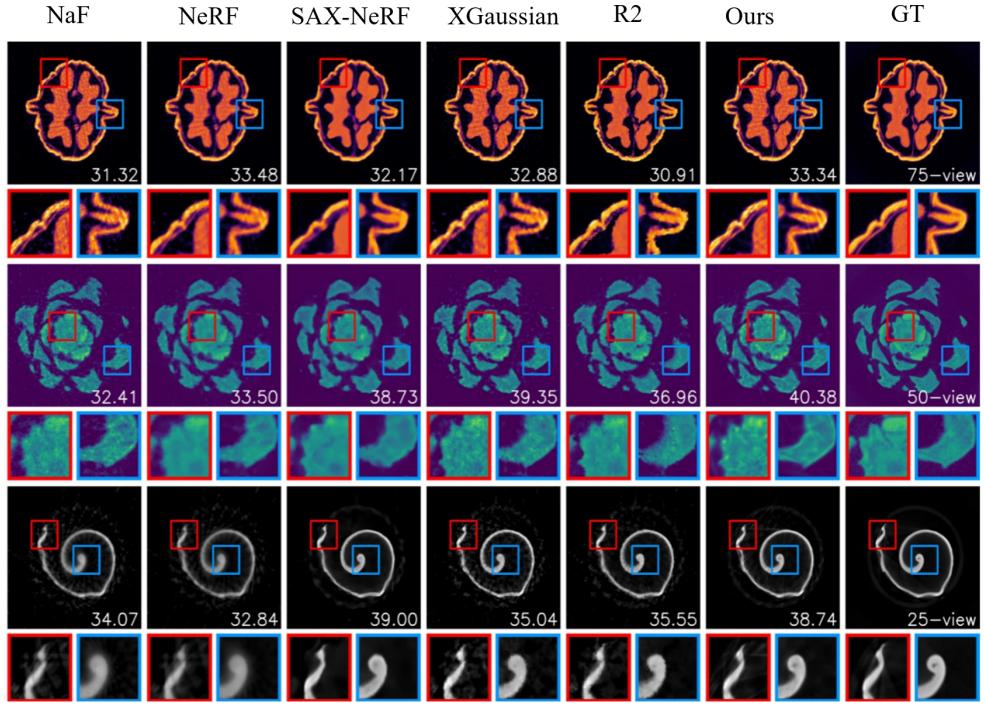


Figure 5. Colorized slice examples of different methods with PSNR (dB) shown at the bottom right of each image. The first three rows are from the synthetic dataset and the last row is from the real-world dataset. Our method recovers more details and suppresses artifacts.

Table 2. Ablation study of Our model. (Synthetic dataset)

Methods	26 Views		24 Views		22 Views		20 Views		18 Views	
	SSIM↑	PSNR↑								
basic (MLP-only)	0.89	24.82	0.89	23.93	0.88	23.32	0.86	21.58	0.82	19.07
+ \mathcal{L}_2 -norm reg.	0.89	24.98	0.89	24.21	0.88	23.64	0.86	21.79	0.82	19.46
+tri-CNN	0.91	25.23	0.90	24.66	0.88	23.19	0.86	21.46	0.81	18.72
full model	0.91	25.80	0.90	24.94	0.89	23.98	0.87	22.26	0.82	19.16

4.3 Ablation study

We validate the impact of the proposed triplane CNN generator on the performance of the Geo-aware splatFields model in Tab.2. Here, the basic pipeline implies using only the set of MLPs to directly predict the splat rendering features and point displacements from the initial splat locations, without conditioning on the deep features produced by the triplane CNN. Results indicate that utilizing the deep features regressed by a CNN improves the quality, with the splat L2-norm regularization term further benefiting the reconstruction. Note that the regularization has a marginal improvement on the results of our pipeline that does not utilize the CNN feature generator, demonstrating the synergy of both modeling strategies.

5 Conclusion and future work

In this work, we proposed an effective optimization strategy that introduces spatial biases and local structural information into the 3D Gaussian splats during optimization process by modeling them through a continuous neural field. We demonstrated that our optimization strategy considerably enhances CT reconstruction quality in the sparse setups, without the need for any external, data-driven priors. We plan to further investigate a multi-task learning framework to optimize imaging quality under sparse-view conditions for different CT scan tasks.

References

- [1] Jonas Adler and Ozan Öktem. Learned primal-dual reconstruction. *IEEE transactions on medical imaging*, 37(6):1322–1332, 2018.
- [2] 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.
- [3] Rushil Anirudh, Hyojin Kim, Jayaraman J Thiagarajan, K Aditya Mohan, Kyle Chamoley, and Timo Bremer. Lose the views: Limited angle ct reconstruction via implicit sinogram completion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 6343–6352, 2018.

- [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] Yuanhao Cai, Yixun Liang, Jiahao Wang, Angtian Wang, Yulun Zhang, Xiaokang Yang, Zongwei Zhou, and Alan Yuille. Radiative gaussian splatting for efficient x-ray novel view synthesis. In *European Conference on Computer Vision*, pages 283–299. Springer, 2025.
- [6] Yuanhao Cai, Jiahao Wang, Alan Yuille, Zongwei Zhou, and Angtian Wang. Structure-aware sparse-view x-ray 3d reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11174–11183, 2024.
- [7] Zilong Chen, Feng Wang, Yikai Wang, and Huaping Liu. Text-to-3d using gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21401–21412, 2024.
- [8] Hyungjin Chung, Dohoon Ryu, Michael T McCann, Marc L Klasky, and Jong Chul Ye. Solving 3d inverse problems using pre-trained 2d diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 22542–22551, 2023.
- [9] Leonardo De Chiffre, Simone Carmignato, J-P Kruth, Robert Schmitt, and Albert Weckenmann. Industrial applications of computed tomography. *CIRP annals*, 63(2):655–677, 2014.
- [10] Philip CJ Donoghue, Stefan Bengtson, Xi-ping Dong, Neil J Gostling, Therese Hultgren, John A Cunningham, Chongyu Yin, Zhao Yue, Fan Peng, and Marco Stampanoni. Synchrotron x-ray tomographic microscopy of fossil embryos. *Nature*, 442(7103):680–683, 2006.
- [11] Anton Du Plessis, Ina Yadroitseva, and Igor Yadroitsev. Effects of defects on mechanical properties in metal additive manufacturing: A review focusing on x-ray tomography insights. *Materials & Design*, 187:108385, 2020.
- [12] Lee A Feldkamp, Lloyd C Davis, and James W Kress. Practical cone-beam algorithm. *Josa a*, 1(6):612–619, 1984.
- [13] Zhongpai Gao, Benjamin Planche, Meng Zheng, Xiao Chen, Terrence Chen, and Ziyan Wu. Ddgs-ct: Direction-disentangled gaussian splatting for realistic volume rendering. *arXiv preprint arXiv:2406.02518*, 2024.
- [14] Muhammad Usman Ghani and W Clem Karl. Deep learning-based sinogram completion for low-dose ct. In *2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP)*, pages 1–5. IEEE, 2018.
- [15] Antoine Guédon and Vincent 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*, pages 5354–5363, 2024.
- [16] Godfrey N Hounsfield. Computed medical imaging. *Science*, 210(4465):22–28, 1980.

- [17] Binbin Huang, Zehao Yu, Anpei Chen, Andreas Geiger, and Shenghua Gao. 2d gaussian splatting for geometrically accurate radiance fields. In *ACM SIGGRAPH 2024 conference papers*, pages 1–11, 2024.
- [18] Kyong Hwan Jin, Michael T McCann, Emmanuel Froustey, and Michael Unser. Deep convolutional neural network for inverse problems in imaging. *IEEE transactions on image processing*, 26(9):4509–4522, 2017.
- [19] Avinash C Kak and Malcolm Slaney. *Principles of computerized tomographic imaging*. SIAM, 2001.
- [20] Shigehiko Katsuragawa and Kunio Doi. Computer-aided diagnosis in chest radiography. *Computerized Medical Imaging and Graphics*, 31(4-5):212–223, 2007.
- [21] 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.
- [22] Timo Kiljunen, Touko Kaasalainen, Anni Suomalainen, and Mika Kortesniemi. Dental cone beam ct: A review. *Physica Medica*, 31(8):844–860, 2015.
- [23] D Kinga, Jimmy Ba Adam, et al. A method for stochastic optimization. In *International conference on learning representations (ICLR)*, volume 5, page 6. San Diego, California;, 2015.
- [24] Muhammed Kocabas, Jen-Hao Rick Chang, James Gabriel, Oncel Tuzel, and Anurag Ranjan. Hugs: Human gaussian splats. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 505–515, 2024.
- [25] Suhyeon Lee, Hyungjin Chung, Minyoung Park, Jonghyuk Park, Wi-Sun Ryu, and Jong Chul Ye. Improving 3d imaging with pre-trained perpendicular 2d diffusion models. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10710–10720, 2023.
- [26] Zhihao Liang, Qi Zhang, Ying Feng, Ying Shan, and Kui Jia. Gs-ir: 3d gaussian splatting for inverse rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21644–21653, 2024.
- [27] 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.
- [28] Yiqun Lin, Jiewen Yang, Hualiang Wang, Xinpeng Ding, Wei Zhao, and Xiaomeng Li. C[^] 2rv: Cross-regional and cross-view learning for sparse-view cbct reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 11205–11214, 2024.
- [29] Youtian Lin, Zuozhuo Dai, Siyu Zhu, and Yao Yao. Gaussian-flow: 4d reconstruction with dynamic 3d gaussian particle. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21136–21145, 2024.

- [30] Baoyuan Liu, Min Wang, Hassan Foroosh, Marshall Tappen, and Marianna Pensky. Sparse convolutional neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 806–814, 2015.
- [31] Jiaming Liu, Rushil Anirudh, Jayaraman J Thiagarajan, Stewart He, K Aditya Mohan, Ulugbek S Kamilov, and Hyojin Kim. Dolce: A model-based probabilistic diffusion framework for limited-angle ct reconstruction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10498–10508, 2023.
- [32] Xian Liu, Xiaohang Zhan, Jiaxiang Tang, Ying Shan, Gang Zeng, Dahua Lin, Xihui Liu, and Ziwei Liu. Humangaussian: Text-driven 3d human generation with gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6646–6657, 2024.
- [33] Zhengchun Liu, Tekin Bicer, Rajkumar Kettimuthu, Doga Gursoy, Francesco De Carlo, and Ian Foster. Tomogan: low-dose synchrotron x-ray tomography with generative adversarial networks: discussion. *JOSA A*, 37(3):422–434, 2020.
- [34] Tao Lu, Mulin Yu, Lining Xu, Yuanbo Xiangli, Limin Wang, Dahua Lin, and Bo Dai. Scaffold-gs: Structured 3d gaussians for view-adaptive rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20654–20664, 2024.
- [35] Vladan Lučić, Friedrich Förster, and Wolfgang Baumeister. Structural studies by electron tomography: from cells to molecules. *Annu. Rev. Biochem.*, 74(1):833–865, 2005.
- [36] Jonathon Luiten, Georgios Kopanas, Bastian Leibe, and Deva Ramanan. Dynamic 3d gaussians: Tracking by persistent dynamic view synthesis. In *2024 International Conference on 3D Vision (3DV)*, pages 800–809. IEEE, 2024.
- [37] Stephen H Manglos, George M Gagne, Andrzej Krol, F Deaver Thomas, and Rammohan Narayanaswamy. Transmission maximum-likelihood reconstruction with ordered subsets for cone beam ct. *Physics in Medicine & Biology*, 40(7):1225, 1995.
- [38] 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. *Communications of the ACM*, 65(1):99–106, 2021.
- [39] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.
- [40] Songyou Peng, Michael Niemeyer, Lars Mescheder, Marc Pollefeys, and Andreas Geiger. Convolutional occupancy networks. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part III 16*, pages 523–540. Springer, 2020.
- [41] Johann Radon. On the determination of functions from their integral values along certain manifolds. *IEEE transactions on medical imaging*, 5(4):170–176, 1986.

- [42] Darius Rückert, Yuanhao Wang, Rui Li, Ramzi Idoughi, and Wolfgang Heidrich. Neat: Neural adaptive tomography. *ACM Transactions on Graphics (TOG)*, 41(4):1–13, 2022.
- [43] Jason Sanders. *CUDA by Example: An Introduction to General-Purpose GPU Programming*. Addison-Wesley Professional, 2010.
- [44] Ken Sauer and Charles Bouman. A local update strategy for iterative reconstruction from projections. *IEEE Transactions on Signal Processing*, 41(2):534–548, 1993.
- [45] William C Scarfe, Allan G Farman, Predag Sukovic, et al. Clinical applications of cone-beam computed tomography in dental practice. *Journal-Canadian Dental Association*, 72(1):75, 2006.
- [46] 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.
- [47] Ruoxi Shi, Xinyue Wei, Cheng Wang, and Hao Su. Zerorf: Fast sparse view 360deg reconstruction with zero pretraining. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 21114–21124, 2024.
- [48] 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.
- [49] Jiaxiang Tang, Jiawei Ren, Hang Zhou, Ziwei Liu, and Gang Zeng. Dreamgaussian: Generative gaussian splatting for efficient 3d content creation. *arXiv preprint arXiv:2309.16653*, 2023.
- [50] Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Deep image prior. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 9446–9454, 2018.
- [51] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.
- [52] Guanjun Wu, Taoran Yi, Jiemin Fang, Lingxi Xie, Xiaopeng Zhang, Wei Wei, Wenyu Liu, Qi Tian, and Xinggang Wang. 4d gaussian splatting for real-time dynamic scene rendering. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20310–20320, 2024.
- [53] Ziyi Yang, Xinyu Gao, Wen Zhou, Shaohui Jiao, Yuqing Zhang, and Xiaogang Jin. Deformable 3d gaussians for high-fidelity monocular dynamic scene reconstruction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20331–20341, 2024.
- [54] Taoran Yi, Jiemin Fang, Junjie Wang, Guanjun Wu, Lingxi Xie, Xiaopeng Zhang, Wenyu Liu, Qi Tian, and Xinggang Wang. Gaussiandreamer: Fast generation from text to 3d gaussians by bridging 2d and 3d diffusion models. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 6796–6807, 2024.

- [55] Xingde Ying, Heng Guo, Kai Ma, Jian Wu, Zhengxin Weng, and Yefeng Zheng. X2ct-gan: reconstructing ct from biplanar x-rays with generative adversarial networks. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10619–10628, 2019.
- [56] Zehao Yu, Anpei Chen, Binbin Huang, Torsten Sattler, and Andreas Geiger. Mip-splatting: Alias-free 3d gaussian splatting. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 19447–19456, 2024.
- [57] Zehao Yu, Torsten Sattler, and Andreas Geiger. Gaussian opacity fields: Efficient high-quality compact surface reconstruction in unbounded scenes. *arXiv preprint arXiv:2404.10772*, 2024.
- [58] Guangming Zang, Ramzi Idoughi, Rui Li, Peter Wonka, and Wolfgang Heidrich. Intratomo: self-supervised learning-based tomography via sinogram synthesis and prediction. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1960–1970, 2021.
- [59] Ruyi Zha, Tao Jun Lin, Yuanhao Cai, Jiwen Cao, Yanhao Zhang, and Hongdong Li. R²-gaussian: Rectifying radiative gaussian splatting for tomographic reconstruction. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2024.
- [60] 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.