

LiDAR-GS: Real-time LiDAR Re-Simulation using Gaussian Splatting

Qifeng Chen¹, Sicong Du¹, Tao Tang², Rengan Xie³, Peng Chen¹, Yuchi Huo³, Sheng Yang¹✉

¹Unmanned Vehicle Dept., CaiNiao Inc., Alibaba Group

²State Key Laboratory of CAD&CG Zhejiang University

³Sun Yat-sen University

{cqf7419, shengyang93fs}@gmail.com

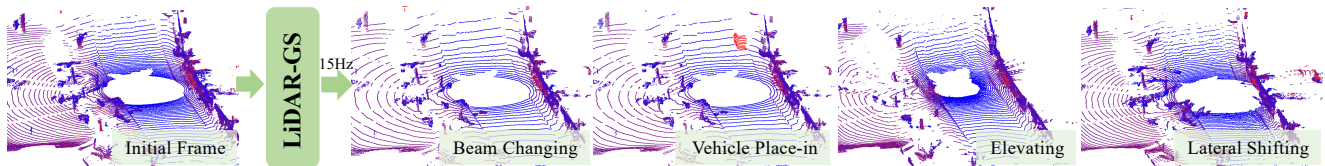


Figure 1: We present LiDAR-GS, the LiDAR Gaussian Splatting for re-simulating LiDAR frames rendered in real-time.

Abstract

We present LiDAR-GS, a Gaussian Splatting (GS) method for real-time, high-fidelity re-simulation of LiDAR scans in public urban road scenes. Recent GS methods proposed for cameras have achieved significant advancements in real-time rendering beyond Neural Radiance Fields (NeRF). However, applying GS representation to LiDAR, an active 3D sensor type, poses several challenges that must be addressed to preserve high accuracy and unique characteristics. Specifically, LiDAR-GS designs a differentiable laser beam splatting, using range-view representation for precise surface splatting by projecting lasers onto micro cross-sections, effectively eliminating artifacts associated with local affine approximations. Furthermore, LiDAR-GS leverages Neural Gaussian Representation, which further integrates view-dependent clues, to represent key LiDAR properties that are influenced by the incident direction and external factors. Combining these practices with some essential adaptations, e.g., dynamic instances decomposition, LiDAR-GS succeeds in simultaneously re-simulating depth, intensity, and ray-drop channels, achieving state-of-the-art results in both rendering frame rate and quality on publically available large scene datasets when compared with the methods using explicit mesh or implicit NeRF. Our source code is publicly available.

Keywords: lidar reconstruction and re-simulation, rendering, gaussian splatting.

1. Introduction

Closed-loop simulation [3, 32, 34] is the basis for evolving *sensor-to-control* end-to-end autonomous driving systems [5, 16]. High quality Novel View Synthesis (NVS) enables sequentially deducing the subsequent effects of a series of vehicle control instructions virtually and safely. In recent advances, camera re-simulation capabilities based on Gaussian Splatting [11] have gradually been integrated into these platforms, making real-time rendering in such reconstructed scenes feasible and applicable. However, academia and industry still lack efficient simulation methods that are both real-time and fidelity, for the LiDAR sensor prevalently equipped by the L4 autonomy.

There are four key distinctions between camera and LiDAR simulation: (1) Most pinhole cameras share the common perspective projection model (even with varying distortions), whose local affine approximations in Normalized Device Coordinates (NDC) accumulates artifacts [7], and thus not directly applicable for precise modeling of 360° rotating LiDAR sensors. (2) Spherical Harmonics (SH) [11], designed for view-dependent color rendering, are suboptimal for intensity modeling of LiDAR. Compared to passive cameras, the variance of the incident direction and distance on the scanned surface would heavily influence the measurement. (3) Also due to such an active sensing characteristic, some emitted laser beams would return no range measurement when the observed return signal has either too low amplitude or no clear peak, called ray-drop characteristics. (4) Unlike rasterized camera pixels with uniform horizontal/vertical sampling, LiDAR is extremely sparse on the vertical FOV, typically only with 16-128 rotating laser beams.

Our approach addresses these four differences to bring-

up LiDAR-GS accordingly: (1) We choose range-view representation [12] to organize laser beams and design a differentiable laser beam splatting method, projecting 3D Gaussians onto beam-aligned micro-planes instead of the unified NDC. This eliminates affine approximation errors while preserving geometric fidelity. (2) For high-fidelity modeling and rendering on these LiDAR properties sensitive to the incident direction and distance, we propose to use Neural Gaussian Representation instead of the Spherical Harmonics, reaching higher quality of re-simulation according to our experiments (Table 5 and Figure 8). (3) For simulating the ray-drop characteristics, we assign a learnable ray-drop probability to each Gaussian primitive, letting forward rendering to judge the validity of emitted laser beam. (4) For vertical sparsity, we introduce spatial scale regularization and compact bounding-box strategies to constrain Gaussian scales and pixel coverage, ensuring precise gradient propagation. A dynamic decomposition mechanism further adapts the model to real-world driving scenarios without compromising real-time performance.

In conclusion, by combining methods and strategies mentioned above, we present the Gaussian Splatting approach for LiDAR re-simulation tasks called LiDAR-GS, which achieves state-of-the-art performance in both rendering quality and efficiency on publicly available datasets [27, 17] compared to previous methods using explicit mesh or NeRF. Our solution also effectively addresses dynamic instances through decomposed reconstruction, seamlessly reintegrating them into the static background.

2. Related Work

Among applicable simulators [15], CARLA [3] builds the traditional render engine upon manufactured virtual scenes with assets modeling, texture editing, and shading parameters tuning. Such a high expense hinders a large-scale generation of virtual scenes. Latest advancements such as MARS [32] and UniSim [34] take advantage of NeRF’s realistic rendering ability, enabling re-simulating sequences captured from real-world environments and guaranteeing a smaller sim-to-real gap [21, 18]. However, these NeRF-based simulators are comparably less efficient than GS-based rendering, requiring more effort to bring-up a real-time simulation system, and they neither support structuralized scene editing due to their implicit representation. Consequently, several approaches substitute the scene representation to GS for cameras (Section 2.1), whereas such a representation has not been thoroughly considered for LiDAR (Section 2.2).

2.1. Gaussian Splatting for Cameras

Evolution of Scene Representation. Using reconstruction methods to recover the driving scene for camera re-simulation has experienced three stages: Explicit recon-

struction with ray-casting [2], NeRF [23], and GS [11]. While rendering and Novel View Synthesis (NVS) quality is gradually advancing throughout these stages, both the efficiency and the asset editing ability drop in the NeRF stage. Thus, several sensor simulators [38] refactor their rendering pipeline by GS representation to support efficient rendering and flexible scene editing.

Gaussian Splatting. 3D Gaussian Splatting (3DGS) [11] proposes to rasterize a set of Gaussian ellipsoids to approximate the appearance of a 3D scene, achieving fast convergence of reconstruction. It also chooses to reversely splat intersected primitives onto rasterized sensor frames for acceleration, ensuring real-time rendering performance. 2DGS [6] replaces 3D Gaussians with 2D Gaussians for a more accurate ray-splat intersection, and employs a low-pass filter to avoid degenerated line projection. Scaffold-GS [20] proposes to initialize a voxel grid and attach learnable features onto each voxel point and all attributes of Gaussians are determined by interpolated features and lightweight neural networks, and Octree-GS [26] further introduces a level-of-detail strategy to better capture details. GS++ [7] analyzes the mathematical expectation of the 3DGS error function, and optimizes the projection to adapt to different camera models, such as fisheye. AbsGS [35] proposes the absolute value accumulation of gradients to further optimize the growth and splitting strategies.

2.2. LiDAR Re-simulation Methods

Explicit reconstruction with ray-casting. Explicit reconstruction methods mainly use two types of 3D data structures: The first category, 3D mesh, is reconstructed by either a stand-alone point-to-mesh process such as SPSR [10], or Truncated Signed Distance Function (TSDF) Volumes [2] with Marching Cubes [19]. The second category, Surface Elements (Surfels), regards points as planar discs with additional normal and radius attributes, and gradually refine their attributes through scan integration [30, 1]. LiDARsim [22] reconstructs scenes as a Surfel-based 3D mesh and applies a neural network to learn the ray-drop characteristics as a binary classification problem. PC-Gen [14] further presents First Peak Averaging (FPA) ray-casting and a surrogate model of ray-drop, achieving improvements in the domain gap between real and simulated LiDAR frames.

NeRF based methods. Recent LiDAR re-simulation methods [28, 31, 8, 29, 36] extend the NeRF formulation for novel LiDAR view synthesis. Among them, LiDAR-NeRF [28] and NFL [8] firstly propose the differentiable LiDAR re-simulation framework, which treats the oriented LiDAR laser beams as a set of rays and render 3D points and intensities in a similar manner as RGB. DyNFL [31] and LiDAR4D [36] extend to dynamic scenes. AlignMiF [29]

further proposes to incorporate and align multimodal inputs for better LiDAR re-simulation. Through our experiments, we found refactoring LiDAR re-simulation with Gaussian Splatting follows the trend of camera re-simulation discussed in section 2.1.

Contemporaneous methods. Contemporaneous publicly available work includes GS-LiDAR [9] and LiDAR-RT [37]. In contrast to these work, we introduce a micro cross-section projection assumption and rasterization for LiDAR rendering and instantiate it on 3DGS primitives. Furthermore, we parameterize the properties of GS primitives as functions of range and viewing direction, achieving state-of-the-art performance in both rendering quality and training efficiency.

3. Preliminaries

3.1. Range-view For LiDAR Representation

In the range-view representation of LiDAR [12, 28], image rows and columns correspond to discretized vertical and horizontal FOV within the vertical FOV f_v . During range-view conversion, each point (x, y, z) in the sensor euclidean coordinate can be projected to a corresponding pixel location (h, w) on the $H \times W$ range-view image, as:

$$\begin{pmatrix} h \\ w \end{pmatrix} = \begin{pmatrix} (1 - (\phi + f_\downarrow)/f_v) \\ 0.5 \cdot (1 - \theta/\pi) \end{pmatrix} \cdot \begin{pmatrix} H \\ W \end{pmatrix}, \quad (1)$$

where the vertical angle $\phi = \arcsin(z, \sqrt{x^2 + y^2})$ and the horizontal angle $\theta = \arctan(y, x)$. The vertical FOV $f_v \triangleq f_\downarrow + f_\uparrow$ is divided into the lower and upper part from the $z = 0$ horizontal plane. The reverse operation of Equation (1) is:

$$\begin{pmatrix} \phi \\ \theta \end{pmatrix} = \begin{pmatrix} f_\uparrow - f_v \cdot h/H \\ \pi - 2\pi \cdot w/W \end{pmatrix}, \quad (2)$$

and the back-projection of Equation (1) can be operated as:

$$(x, y, z) = d \cdot \mathbf{d} \triangleq d \cdot (\cos \theta \cos \phi, \sin \theta \sin \phi, \cos \phi), \quad (3)$$

where d stands for the distance along the ray to the scanned target, and the ray-direction \mathbf{d} is determined by (ϕ, θ) .

Through such projection, we can use the position (h, w) in the range-view image to encode the corresponding information for (x, y, z) in a multi-channel format, including depth, intensity, and ray-drop. Both f_v and the exact vertical angle ϕ of each laser beam can be found in the LiDAR intrinsic parameters.

3.2. 3D Gaussian Splatting

3DGS [11] uses Gaussian ellipsoids to explicitly represent the scene and introduces a point-based differentiable volume splatting. Specifically, the parameterization of the Gaussian primitives follows:

$$\mathbb{G}(\mathbf{x}) = \exp(0.5 \cdot \partial \mathbf{x}^\top \Sigma^{-1} \partial \mathbf{x}), \quad (4)$$

where $\partial \mathbf{x}$ is the relative position from the center of the Gaussian ellipsoids, and Σ is its 3D covariance matrix. The covariance matrix $\Sigma = \mathbf{R} \mathbf{S} \mathbf{S}^\top \mathbf{R}^\top$ is factorized into a scaling matrix \mathbf{S} and a rotation matrix \mathbf{R} . To render an image, Gaussian ellipsoids will be projected into 2D image space, where the 2D projected covariance matrix $\Sigma' = \mathbf{J} \mathbf{W} \Sigma \mathbf{W}^\top \mathbf{J}^\top$ undergoes viewing transformation \mathbf{W} and projective transformation, with \mathbf{J} the Jacobian matrix of the local affine approximation during projection, designed to ensure that the ellipsoids maintain their regular shape after being projected from 3D to 2D. After the projection transformation, the color $\mathbb{C}(\mathbf{p})$ of each pixel \mathbf{p} is computed by integrating overlapping Gaussian ellipses with α -weighted:

$$\mathbb{C}(\mathbf{p}) = \sum_{i \in \mathbf{N}_d} c_i \cdot \alpha_{\xi_i} \prod_{j=1}^{i-1} (1 - \alpha_{\xi_j}), \quad (5)$$

where c_i is the color from Spherical Harmonics [11] of each Gaussian primitive $\xi_i \in \mathbf{N}_d$ that is intersected by ray \mathbf{d} emitted from the sensor viewpoint to the image pixel \mathbf{p} .

4. LiDAR Gaussian Splatting

4.1. LiDAR-GS Overview

Backward reconstruction. The reconstructed scene is represented by a set of Gaussians $\bigcup \xi$, and each Gaussian primitive $\xi = \{\mathbf{x}, \mathbf{R}, \mathbf{S}, \rho, r, \alpha\}$ stores position \mathbf{x} , rotation quaternion \mathbf{R} , scale coefficient \mathbf{S} , intensity ρ , ray-drop probability r , and opacity α .

To enhance the fidelity of LiDAR depth and intensity, while incorporating the characteristics of this active sensor, we employ the Neural Gaussian Representation (NGR) motivated by NeRF [23] and Scaffold-GS [20]. Specifically, we indirectly utilize a scene feature vector \mathbf{v}_ξ for each Gaussian (32-dim, initialized to zero), and use four parallel lightweight Multi-Layer Perceptrons (MLPs) to obtain the desired attributes for each Gaussian during rendering, as

\mathbb{F}_Σ , *Offset and Covariance MLP*: We use \mathbb{F}_Σ to learn the covariance components $\Sigma = \mathbf{R} \mathbf{S} \mathbf{S}^\top \mathbf{R}^\top$ and offset $\delta \mathbf{x}$ of init position \mathbf{x} , with inputs also related to view direction \mathbf{d}' :

$$\mathbb{F}_\Sigma(\mathbf{v}_\xi, \mathbf{d}', d) \rightarrow \delta \mathbf{x}, \mathbf{R}, \mathbf{S}, \quad (6)$$

where we apply the offset $\delta \mathbf{x}$ as $\mathbf{x} = \mathbf{x} + \delta \mathbf{x} \cdot \mathbf{S}$ [20] during reconstruction. This facilitates the movement and expansion of the Gaussian in the direction of the gradient, promoting their growth and split progress.

\mathbb{F}_ρ , *Intensity MLP*: The definition of \mathbb{F}_ρ is as follows:

$$\mathbb{F}_\rho(\mathbf{v}_\xi, \mathbf{d}', d) \rightarrow \rho, \quad (7)$$

where we found MLP achieves better deduction than Spherical Harmonics when an input direction \mathbf{d}' and the distance-of-flight d is given (See Table 5).

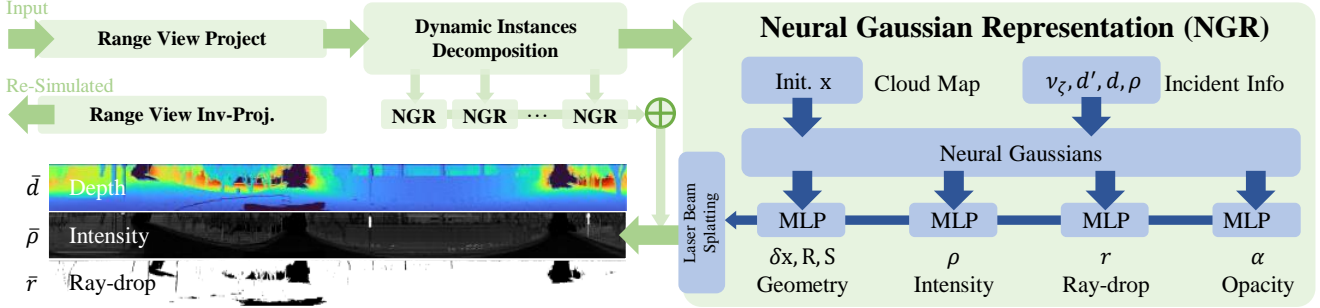


Figure 2: **The overview of LiDAR-GS pipeline.** We choose to represent LiDAR frames as range-view images (Section 3.1), and use Gaussian Splatting to perform both reconstruction and rendering (Section 3.2). Specifically, we choose to use Neural Gaussian Representation (NGR) with four MLPs to learn final attributes for each Gaussian (Section 4.1), where projecting and rasterizing these LiDAR Gaussians are performed through our proposed Laser Beam Splating strategy (Section 4.2). Dynamic instances are separately reconstructed in their stand-alone coordinates for final stitching (Section 4.3).

\mathbb{F}_r , *Ray-drop MLP*: We additionally use \mathbb{F}_r to learn about ray-drop probability, which means that laser beam do not return a value. The phenomenon of ray-drop usually exists in blind spots at close range and unmeasurable far range. Only in rare cases, the attenuation of laser caused by special materials can also exhibit ray-drop. Therefore, the input of the network contains information on view direction \mathbf{d}' , feature vector \mathbf{v}_ξ for each Gaussian and distance d :

$$\mathbb{F}_r(\mathbf{v}_\xi, \mathbf{d}', d) \rightarrow r. \quad (8)$$

\mathbb{F}_α , *Opacity MLP*: Similar with the previous method [11], we have also set up an opacity network \mathbb{F}_α , as:

$$\mathbb{F}_\alpha(\mathbf{v}_\xi, \mathbf{d}', d) \rightarrow \alpha, \quad (9)$$

where α interprets occlusions and invisibility among individual Gaussians, which is an indispensable attribute.

Optimizing \mathbf{v}_ξ and MLPs. Given input real-world LiDAR scans, we use the following objective loss function to optimize the feature vector \mathbf{v}_ξ of each Gaussian and four MLPs:

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_d + \mathcal{L}_\rho + \mathcal{L}_r + \mathcal{L}_S + \mathcal{L}_{cd}, \\ \mathcal{L}_\rho &= (1 - \lambda_\rho) \cdot \mathcal{L}_1 + \lambda_\rho \cdot \mathcal{L}_{D-SSIM}, \\ \mathcal{L}_S &= \frac{1}{N} \sum_{\xi_i} \text{Prod}(\mathbf{S}_{\xi_i}), \end{aligned} \quad (10)$$

where the depth loss \mathcal{L}_d follows \mathcal{L}_1 loss, and we set $\lambda_\rho = 0.2$ for the intensity loss \mathcal{L}_ρ . The CD loss \mathcal{L}_{cd} follows the chamfer distance [4] between the rendered point cloud and GT. The ray-drop probability loss \mathcal{L}_r follows \mathcal{L}_2 loss, and we use a regularization loss \mathcal{L}_S to inhibit the enlargement of scale.

Forward rendering. Given a required free viewpoint and orientation for rendering, we choose to use the range-view rasterization and generate view-directions for interacting with splatted Gaussians. Specifically, each pixel \mathbf{p}

in the rasterized range-view determines a view-direction \mathbf{d} to fetch all related Gaussians \mathbf{N}_d , and differential rendering for intensity ρ presents an integration approach as $\bar{\rho} = \sum_{i \in \mathbf{N}_d} \omega_i \rho_i$. By the same logic, the range-view image of depth and ray-drop probability can be calculated accordingly through the distance between the sensor and primitives. To determine the occurrence of ray-drop, we establish a threshold for the ray-drop probability, defined as $\bar{r} < \lambda_r$ ($\lambda_r = 0.5$). These criteria are employed to assess whether the laser beam is lost in a given ray direction. If so, clear all attributes for such a pixel.

4.2. Laser Beam Splating

In most pinhole camera GS methods [11], 3D Gaussians are splatted to the normalized image plane, employing the Jacobian matrix to approximate the non-linear projection process. Such a local affine approximation consequently degrades the quality of photo-realistic rendering [7], and is also not applicable for LiDAR frames. Even in the range-view representation with equirectangular projection, the degradation still masters, because the LiDAR sampling is not equidistant in the vertical direction. Therefore, the differences between pinhole shuttering for passive cameras and range-finding for active LiDAR sensors cannot be bridged by an single approximate projection.

In order to migrate Gaussian Splatting from passive camera to active LiDAR and minimize the distortion caused by the Gaussian projection and rasterization process, we introduce a new laser beam splating method, called *Micro Cross-Section Projection*, with implementation details illustrated in Figure 3 and provided in algorithm 1.

Gaussian projection. We use the distance of flight $d = |\mathbf{x} - \mathbf{o}|$ (where \mathbf{o} represents the sensor viewpoint) instead of z-buffering in the NDC space [11] for depth buffering, and establish the target projection plane Π perpendicular to the direction \mathbf{d}' for each Gaussian, precisely located

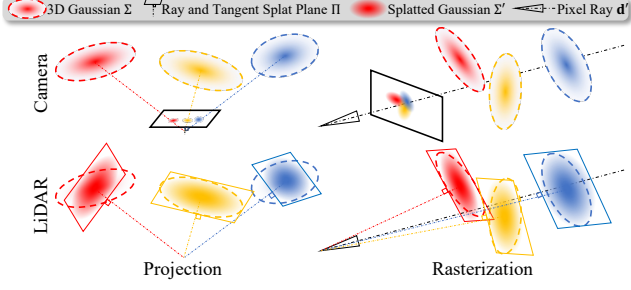


Figure 3: **Illustration of the proposed Micro Cross-Section Projection.** Instead of splatting Gaussians to the normalized image plane like 3DGS [11], we calculate the cross-section micro-plane through the ray established between Gaussian mean and sensor viewpoint for splatting. During forward rendering, we collect acrossed Gaussians along the ray direction to integrate attributes for each pixel.

at its mean position \mathbf{x} . Hence, the projection of the Gaussian onto this plane Π corresponds to the cross-section of the Gaussian itself, thus eliminating any distortion in the projection process.

Since there is no longer a normalized image plane or sphere to determine the range of affected pixels for each Gaussian during projection, we employed a compact Axis-Aligned Bounding-Box (AABB) to ascertain the pixel coverage of each Gaussian. Specifically, we establish the major axes \mathbf{a} and minor axes \mathbf{b} of the Gaussian projection on the projection plane and transform it into a rectangular bounding box (b_h, b_w) , where $b_h = \mathbf{b} / \tan \phi$, $b_w = \mathbf{a} / \tan \theta$. Due to the sparsity of the laser beam in the vertical FOV, the vertical angle ϕ is usually bigger than the horizontal angle θ , leading to significant variations in gradient. Therefore, compared to directly using the maximum radius as the selection box [11], our strategy reduces the unnecessary pixel coverage, which is beneficial for optimizing the vertical direction. The ablation study in Table 5 confirms the effectiveness of this strategy.

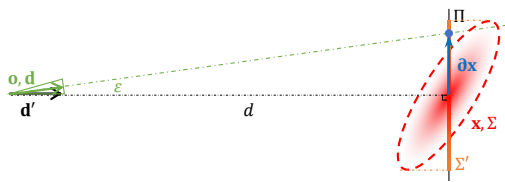


Figure 4: Sideview of the back-projection for each Gaussian during rendering.

Gaussian rasterization. As shown in Figure 4, during rendering of such a pixel with ray-direction \mathbf{d} , we back-project the relative position of the ray to the Gaussian onto the corresponding cross-section micro-plane Π . This step is crucial for calculating the distribution of the set of inter-

Algorithm 1 Laser Beam Splatting

```

1: {1. Gaussian Reconstruction}
2:  $\xi : \{\mathbf{x}, \mathbf{R}, \mathbf{S}, \rho, r, \alpha\}$ 
3: {2. Gaussian Projection}
4: for  $\xi$  do
5:    $d \leftarrow \|\mathbf{x} - \mathbf{o}\|$ ,  $\mathbf{d}' \leftarrow (\mathbf{x} - \mathbf{o})/d$   $\triangleright$  Depth and Ray
6:    $n_1 \leftarrow n_1 \perp \mathbf{d}'$ ,  $n_2 \leftarrow n_2 \perp \mathbf{d}'$ ,  $n_1 \triangleright$  Basis of plane

7:    $\Pi, \mathbf{J} \leftarrow n_1, n_2$   $\triangleright$  Tangent plane and Transformation
8:    $\Sigma' \leftarrow \mathbf{J}\mathbf{W}\Sigma\mathbf{W}^\top \mathbf{J}^\top$   $\triangleright$  Projection of covariance
9:    $b_h, b_w \leftarrow \Sigma'$   $\triangleright$  Range of affected pixels
10:  end for
11: {3. Gaussian Rasterization}
12: for  $\mathbf{p} \leftarrow \mathbf{p} : (h, w)$  do
13:    $\mathbf{N}_d \leftarrow \xi : \{b_h, b_w\}$   $\triangleright$  Depth buffering
14:    $\mathbf{p} \leftarrow [\bar{\rho}, \bar{d}, \bar{r}] \leftarrow [0, 0, 0]$   $\triangleright$  Render Target
15:   for  $\xi_i \in \mathbf{N}_d$  do
16:      $\epsilon \leftarrow \arccos(\mathbf{d}' \cdot \mathbf{d})$ 
17:      $\partial \mathbf{x} \leftarrow d, \epsilon$   $\triangleright$  Back-projection
18:      $\alpha_i \leftarrow \alpha_{\xi_i} \cdot \mathbb{G}_i(\partial \mathbf{x})$   $\triangleright$  Splatting and  $\alpha$ -weighting
19:      $[\bar{\rho}, \bar{d}, \bar{r}] \leftarrow \alpha_i \sum_{j=1}^{i-1} (1 - \alpha_j) [\rho_{\xi_j}, d_{\xi_j}, r_{\xi_j}]$ 
20:   end for
21:   if  $\bar{r} < \lambda_r$  then
22:      $\bar{d}, \bar{\rho} \leftarrow 0$   $\triangleright$  Ray-drop
23:   end if
24: end for

```

sected Gaussians $\xi_i \in \mathbf{N}_d$. Specifically for a pixel \mathbf{p} , if its corresponding ray $\{\mathbf{o}, \mathbf{d}\}$ passes through a Gaussian at an actual distance of flight d , we pick the relative position $\partial \mathbf{x}$ as:

$$\partial \mathbf{x} = d \cdot [\mathbf{d} / (\mathbf{d} \cdot \mathbf{d}') - \mathbf{d}']. \quad (11)$$

On the corresponding cross-section micro-plane Π , the real depth at the intersection of the ray and the intersected Gaussian can be calculated by $d / (\mathbf{d} \cdot \mathbf{d}')$. Following Equation (4), the Gaussian splatting of ξ_i w.r.t. pixel \mathbf{p} is used to calculate the rendering weight α_i for Equation (5). Consequently, all attribute of each pixel are computed by integrating these intersected Gaussians with α -weighting.

4.3. Dynamic Instances Decomposition

In dynamic scenes, LiDAR-GS decomposes dynamic instances by utilizing detection results represented as bounding boxes, effectively distinguishing them from the static background. As illustrated in Figure 5 and inspired by NSG [25], each dynamic instance is reconstructed within a standalone coordinate system, and we use the Kabsch algorithm [13] to estimate the relative pose between sequentially observed frames (If not provided). Using these tracked relative poses, we integrate observations from subsequent frames to enhance the completeness of instance-wise reconstruction. During rasterization, reconstructed Gaussians of

each instance are placed w.r.t. the relative pose of the static background for a joint rendering.

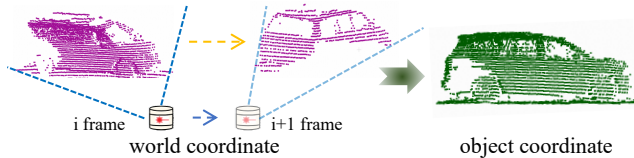


Figure 5: **Dynamic object decomposition and Gaussian modeling.** Given detection and segmentation results of each moving instance between consecutive frames, we separately stitch and reconstruct them in a standalone coordinate system for improving the completeness.

5. Experiments

5.1. Experimental Setting

Real-world Large Scene Datasets. We verified the LiDAR-GS using 12 sequences from Waymo Open Dataset [27] and KITTI-360 Dataset [17], which are public real-world datasets for autonomous driving scenarios. Consistent with previous LiDAR re-simulation methods [28, 31], we utilized 50 continuous frames per sequence, designating every 10th frame for validation and the remaining frames for training. Please refer to Table 1 for the specific sequences chosen.

Table 1: Sequences from Waymo and KITTI-360 for evaluation.

DATASET	SCENE ID
Static Waymo	11379226583756500423_6230_810_6250_810
	10676267326664322837_311_180_331_180
	17761959194352517553_5448_420_5468_420
Dynamic Waymo	1172406780360799916_1660_000_1680_000
	1083056852838271990_4080_000_4100_000
	13271285919570645382_5320_000_5340_000
	10072140764565668044_4060_000_4080_000
KITTI-360	Seq 1538–1588, Seq 1728–1778
	Seq 1908–1958, Seq 3353–3403

Evaluation Metrics. Following previous methods [28, 8, 29, 31], our main evaluation metrics for depth including Chamfer Distance (CD), F-Score (using 5cm as threshold), Root Mean Squared Error (RMSE) and mean absolute error (MAE). Our evaluation metrics for intensity include Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), Root Mean Squared Error (RMSE), and mean absolute error (MAE).

Baseline and Implementation. For static dataset, we used the official settings to conduct experiments with LiDAR-NeRF [28] and AlignMiF [29]. For NFL [8] which

have not yet released their source-code, we stripped the dynamic graph from DyNFL [31] and followed its settings to reproduce the results. Additionally, we maintained the experimental settings as described in LiDAR-NeRF [28] to reproduce the results on PCGen [14] and LiDARsim [22]. For dynamics dataset, we reproduced the results using the same data sequences and official settings as DyNFL [31] and LiDAR4D [36], while also recording the official results of LiDARsim [22] and UniSim [34] from DyNFL [31]. Meanwhile, we also compared our method with contemporaneous works, LiDAR-RT [37] and GS-LiDAR [9], by reproducing their results according to the official source code.

LiDAR-GS was trained for 5,000 iterations on RTX3090. We initialized 500,000 GS anchors, randomly sampled from all LiDAR frames and adopted the same growth and split strategy as Scaffold-GS [20] and the same accumulative gradient strategy as AbsGS [35]. For densification, we adjusted the gradient threshold to 0.006, and set the final densification iterations to 3,000. The network structure employs double-layer 32-dims tiny MLP, implemented using tinycudnn [24].

5.2. Main Results

Table 2 and Table 3 present our quantitative results on the public static dataset and dynamic dataset, respectively. As shown, our proposed method LiDAR-GS achieves best quantitative performance among most metrics. Meanwhile, Table 4 demonstrates that our approach significantly reduces training and inference time compared to existing explicit reconstruction, NeRF methods and contemporaneous GS methods. This improvement marks a milestone that LiDAR re-simulation can now achieve the same or even higher frame rate than real-world sensors, typically at 10Hz, making large-scale re-simulation of real-world sequences feasible.

Results on Static Dataset. The quantitative comparison on the Waymo dataset with the static scene is displayed in Table 2, illustrating LiDAR-GS has achieved state-of-the-art in most indicators. As shown in Figure 6, LiDAR-GS is superior to previous methods in modeling scene surfaces and intensity, such as roads, vehicles, and poles. We attribute this superior performance to our well-designed GS reconstruction process and efficient attribute prediction.

Results on Dynamic Dataset. The quantitative comparison on the Waymo dataset with dynamic objects is displayed in Table 3. We can find our proposed method achieves remarkable performance on most evaluation metrics, underscoring the realism of our re-simulated LiDAR scans. Although LiDAR-GS scores lower than LiDAR4D on some intensity metrics (attributed to its two-stage convolutional network optimization), our rendered results, as shown in Figure 7, appear more realistic. And our reconstruction efficiency is significantly higher than that of GS-

Table 2: Average quantitative results for static scenes from the Waymo Open Dataset and KITTI-360.

Method	Waymo								KITTI-360							
	Point				Intensity				Point				Intensity			
	CD↓	F-score↑	RMSE↓	MAE↓	MAE↓	RMSE↓	PSNR↑	SSIM↑	CD↓	F-score↑	RMSE↓	MAE↓	MAE↓	RMSE↓	PSNR↑	SSIM↑
LiDARsim [22]	18.87	0.58	11.40	4.75	0.074	0.12	18.56	0.32	1.04	0.66	6.04	3.00	0.15	0.20	13.64	0.09
PCGen [14]	0.39	0.77	9.87	3.07	0.072	0.12	18.29	0.33	0.26	0.84	5.08	2.43	0.16	0.22	13.11	0.06
LiDARNeRF [28]	0.19	0.87	7.52	1.79	0.039	0.07	23.22	0.59	0.09	0.92	3.70	1.35	0.10	0.15	16.25	0.25
AlignMiF [29]	0.16	0.88	7.39	1.75	0.038	0.07	23.27	0.60	0.09	0.92	3.69	1.34	0.10	0.15	16.19	0.25
NFL [8]	0.27	0.84	8.40	1.92	0.035	0.08	21.49	0.50	0.21	0.84	4.69	1.59	0.10	0.14	16.61	0.18
LiDAR-GS	0.14	0.91	6.47	1.46	0.032	0.06	24.52	0.64	0.09	0.92	3.39	0.91	0.08	0.13	17.38	0.49

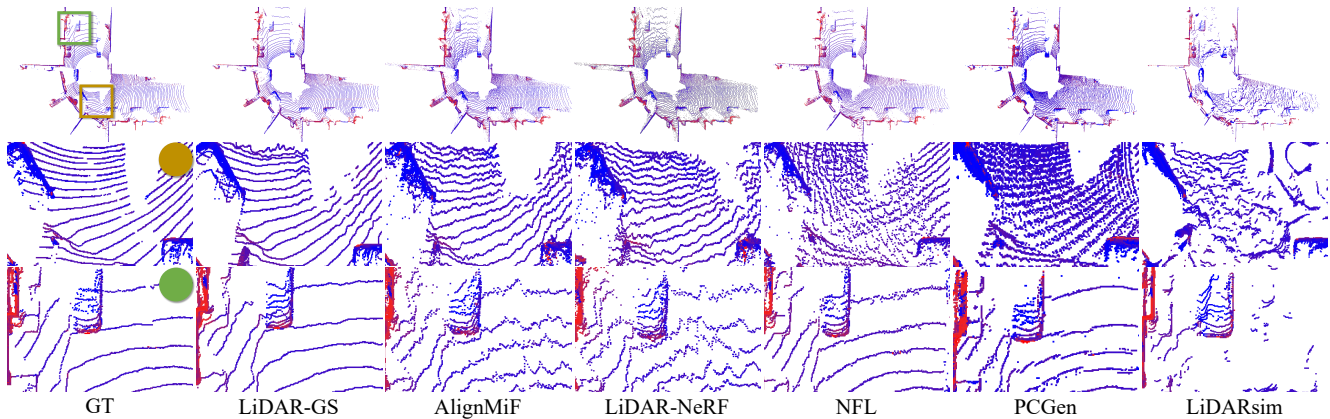


Figure 6: For static scenes, qualitative comparison between LiDAR-GS and other methods on Waymo sequences. Points are colored through their intensity from blue (0) to red (40).

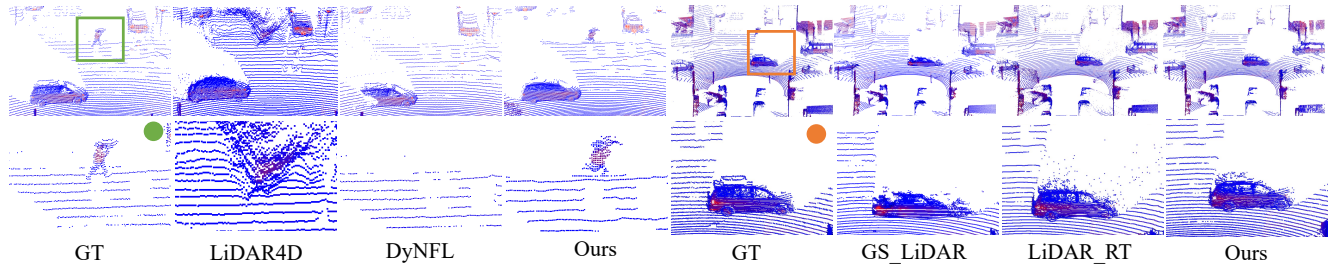


Figure 7: For dynamic scenes, qualitative comparison between LiDAR-GS and other methods on Waymo sequence. Points are colored through their intensity from blue (0) to red (40).

LiDAR, even though the performance of geometric accuracy is comparable.

Results on Novel View Synthesis. We adjusted the free-viewpoint by lateral shifting (Y) and elevating (Z), or change the number of laser beams (from 64 to 32) for LiDAR re-simulation on novel views as shown in Figure 10, demonstrating our effectiveness of NVS in driving scenarios. And supplementary video shows the more results of the re-simulation application, including **vehicle place-in, novel view, beam changing and simulated driving**.

5.3. Ablation Results

Ablations on Neural Gaussian Representation. We conducted an ablation experiment to remove the NGR. Table 5 supports our assertions in Section 4.1. The introduction of NGR significantly enhances the capability of scene representation. It can effectively consider the impact of incident direction and distance. As shown in Figure 8, if Gaussian properties are not related to above factors, it cannot identify positions of ray-drop, leading to incorrect optimization of intensity and depth, such as road blind spots.

Table 3: Average quantitative results for dynamic scenes.

Method	Point				Intensity			
	CD↓	F-scr↑	RMSE↓	MAE↓	MAE↓	RMSE↓	PSNR↑	SSIM↑
LiDARsim [22]	0.311	0.817	10.798	1.701	-	-	-	-
UniSim [34]	0.143	0.902	2.631	0.256	-	-	-	-
DyNFL [31]	0.109	0.929	2.348	0.308	0.024	0.050	26.012	0.858
LiDAR4D [36]	0.106	0.937	2.286	0.306	0.023	0.047	26.632	0.868
LiDAR_RT [37]	0.103	0.938	2.254	0.303	0.019	0.048	26.498	0.861
GS_LiDAR [9]	0.096	0.940	2.150	0.300	0.024	0.050	25.431	0.840
LiDAR-GS	0.095	0.945	2.146	0.300	0.019	0.048	26.550	0.868

Table 4: Average training and inference costs. We reach real-time rendering for LiDAR re-simulation tasks.

Method	Train (min)↓	Inference (fps)↑
LiDARsim [22]	13	0.33
PCGen [14]	25	0.01
LiDAR-NeRF [28]	31	1.80
AlignMiF [29]	110	0.10
DyNFL [31](NFL [8])	464	0.15
LiDAR4D [36]	426	1.7
GS_LiDAR [9]	129	10.8
LiDARRT [36]	213	20.7
LiDAR-GS(w/o LBS)	17	15.9
LiDAR-GS	18	15.0

Table 5: Ablations for the proposed method on Waymo1067.

Module	Point				Intensity			
	CD↓	F-score↑	RMSE↓	MAE↓	MAE↓	RMSE↓	PSNR↑	SSIM↑
w/o NGR	3.907	0.243	12.1	6.64	0.041	0.074	22.57	0.60
w/o LBS	0.264	0.860	7.77	1.76	0.029	0.056	25.03	0.70
w/o \mathcal{L}_S	0.215	0.877	7.25	1.61	0.028	0.057	25.40	0.71
w/o AABB	0.219	0.879	7.42	1.65	0.029	0.057	25.43	0.71
w/o \mathcal{L}_{cd}	0.199	0.883	7.28	1.56	0.027	0.052	25.51	0.73
MLP w/o \mathbf{v}_ξ	0.221	0.871	7.52	1.66	0.029	0.058	25.35	0.69
MLP w/o d	0.205	0.878	7.35	1.61	0.028	0.058	25.38	0.70
LiDAR-GS	0.194	0.885	7.27	1.56	0.027	0.052	25.54	0.73

Ablations on Laser Beam Splatting. We conducted an ablation experiment to remove the LBS. Specifically, we referred to the camera projection method of the vanilla 3DGS [11] and projected Gaussian onto unified pano projection planes for rasterization. Table 5 demonstrates that LBS is more suitable for LiDAR applications. The improvements provided by LBS are illustrated in Figure 8, confirming our assertions in Section 4.2. Additionally, we implemented LBS using CUDA, resulting in only slight reduction in efficiency. The costs of projections are detailed in Table 4.

Ablations on AABB. The results in Table 5 demonstrate that avoiding unnecessary pixel coverage by using

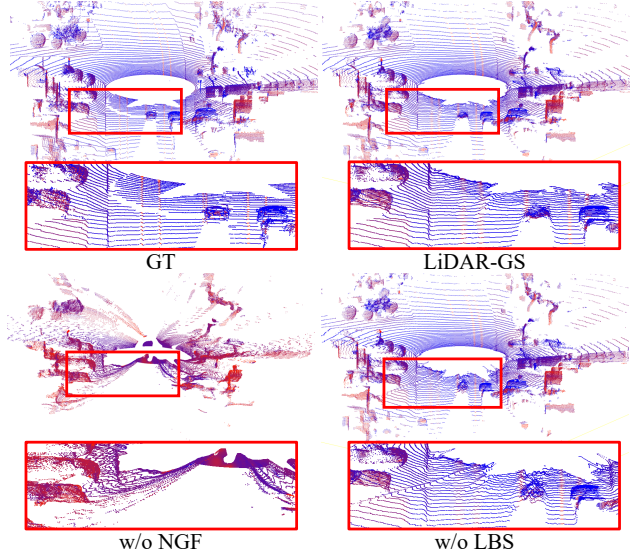


Figure 8: Ablations on Neural Gaussian Representation and Laser Beam Splatting.

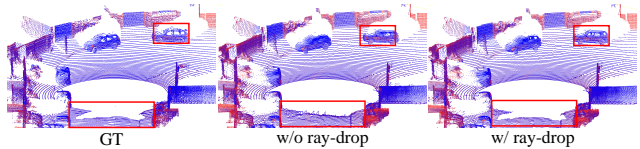


Figure 9: Ablations on the necessity of adding ray-drop estimation.

AABB for Gaussian projection is beneficial for further convergence.

Ablations on regularization terms. We also set up ablation experiments on the regularization term \mathcal{L}_S and loss \mathcal{L}_{cd} mentioned in Equation (10). The depth rendering is an important basis for reflecting the ability of re-simulation, and introducing these two constraints brings some benefits to the Neural Gaussian Representation.

Analysis of ray-drop properties. Figure 9 intuitively shows the effect of ray-drop properties on LiDAR re-simulation. Ray-drops are ubiquitous, some of which have certain regularities (such as blind spots), while others are affected by the incident direction and the environment, and it is difficult to find the distribution pattern intuitively. Therefore, parameterizing and optimizing ray-drops is helpful to improve LiDAR simulation capabilities.

Analysis of MLP. We conducted two additional experiments to demonstrate the necessity of incorporating additional inputs into the MLP network, including distance d and feature vectors \mathbf{v}_ξ . These inputs further assist the network in extracting features and producing accurate GS attributes.

Discussion on the gap between real and simulation.

To gauge the gap between real and simulated scans, we employed detection model, INT [33], on the same trajectory of dynamic data to evaluate the difference in detection results between two point cloud detection results relative to manually annotated GT. Table 6 presents that this difference is negligible in this task, where mAP(mean Average Precision), mF1(mean F1 Score) and mAVE(mean Average Velocity Error) are used as detection metric.

Table 6: Objects detection using simulation data.

Data	mAP↑	mF1↑	mAVE_vx↓	mAVE_vy↓
Real	87.701	87.512	0.314	0.165
LiDAR-GS	87.673	87.387	0.314	0.167

6. Conclusion

In this paper, we present the LiDAR Gaussian Splatting method for real-time high-fidelity re-simulation of sensor scans in real-world urban road scenes. LiDAR-GS choose range-view representation to organize laser beams, and innovatively proposes a differentiable laser beam splatting based on micro cross-section plane. Furthermore, LiDAR-GS precisely simulates LiDAR properties using Neural Gaussian Representation and introducing some strategies dedicated to Gaussian laser beam, including AABB on projection constraints, and spatial scale regularization. LiDAR-GS has been verified in multiple dynamic and static scenes, and performs well. In summary, this GS-based rendering LiDAR re-simulation method will accelerate the progress of autonomous driving in data closure.

Limitation and Future Work. Due to the single training perspective and the point density that inversely proportional to the scanning distance, our proposed LiDAR-GS still exhibits degradation in occluded and extrapolated regions at far distances. Meanwhile, the introduced micro-plane assumption is not sufficiently expressive for modeling semi-transparent materials, and the engineering implementation relies on dynamic decomposition during preprocessing. Throughout our proposed LiDAR-GS pipeline, we also find several points worth further adjustments as a secondary priority: Leverage the ability of generative models, such as diffusion model, to address depth occlusion caused by sparse perspective. And by fusing multi-modal sensor information, such as cameras, model can enhance the understanding of scene features.

References

[1] Y.-P. Cao, L. Kobbelt, and S.-M. Hu. Real-time high-accuracy three-dimensional reconstruction with consumer rgb-d cameras. *ACM TOG*, 37(5):171:1–171:16, 2018. 2

[2] A. Dai, M. Nießner, M. Zollöfer, S. Izadi, and C. Theobalt. Bundlefusion: Real-time globally consistent 3d reconstruction

using on-the-fly surface re-integration. *ACM TOG*, 36(4):76:1–76:18, 2017. 2

[3] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun. CARLA: An open urban driving simulator. In *CoRL*, pages 1–16, 2017. 1, 2

[4] H. Fan, H. Su, and L. Guibas. A point set generation network for 3d object reconstruction from a single image. In *CVPR*, pages 2463–2471, 2017. 4

[5] Y. Hu, J. Yang, L. Chen, K. Li, C. Sima, X. Zhu, S. Chai, S. Du, T. Lin, W. Wang, et al. Planning-oriented autonomous driving. In *CVPR*, pages 17853–17862, 2023. 1

[6] B. Huang, Z. Yu, A. Chen, A. Geiger, and S. Gao. 2d gaussian splatting for geometrically accurate radiance fields. In *SIGGRAPH*. Association for Computing Machinery, 2024. 2

[7] L. Huang, J. Bai, J. Guo, Y. Li, and Y. Guo. On the error analysis of 3d gaussian splatting and an optimal projection strategy. In *ECCV*, pages 247–263, 2024. 1, 2, 4

[8] S. Huang, Z. Gojcic, Z. Wang, F. Williams, Y. Kasten, S. Fidler, K. Schindler, and O. Litany. Neural LiDAR fields for novel view synthesis. In *ICCV*, pages 18236–18246, 2023. 2, 6, 7, 8

[9] J. Jiang, C. Gu, Y. Chen, and L. Zhang. GS-LiDAR: Generating realistic lidar point clouds with panoramic gaussian splatting. *arXiv:2501.13971*, 2025. 3, 6, 8

[10] M. Kazhdan and H. Hoppe. Screened poisson surface reconstruction. *ACM TOG*, 32(3), 2013. 2

[11] B. Kerbl, G. Kopanas, T. Leimkuehler, and G. Drettakis. 3D gaussian splatting for real-time radiance field rendering. *ACM TOG*, 42(4):139:1–139:14, 2023. 1, 2, 3, 4, 5, 8

[12] L. Kong, Y. Liu, R. Chen, Y. Ma, X. Zhu, Y. Li, Y. Hou, Y. Qiao, and Z. Liu. Rethinking range view representation for lidar segmentation. In *ICCV*, pages 228–240, 2023. 2, 3

[13] J. Lawrence, J. Bernal, and C. Witzgall. A purely algebraic justification of the kabsch-umeyama algorithm. *Journal of research of the National Institute of Standards and Technology*, 124:1–6, 2019. 5

[14] C. Li, Y. Ren, and B. Liu. Pcggen: Point cloud generator for lidar simulation. In *IEEE ICRA*, pages 11676–11682, 2023. 2, 6, 7, 8

[15] Y. Li, W. Yuan, S. Zhang, W. Yan, Q. Shen, C. Wang, and M. Yang. Choose your simulator wisely: A review on open-source simulators for autonomous driving. *IEEE TIV*, 9(5):4861–4876, 2024. 2

[16] B. Liao, S. Chen, X. Wang, T. Cheng, Q. Zhang, W. Liu, and C. Huang. MapTR: Structured modeling and learning for online vectorized HD map construction. In *ICLR*, 2023. 1

[17] Y. Liao, J. Xie, and A. Geiger. Kitti-360: A novel dataset and benchmarks for urban scene understanding in 2d and 3d. *IEEE TPAMI*, 45(3):3292–3310, 2022. 2, 6

[18] C. Lindström, G. Hess, A. Lilja, M. Fatemi, L. Hammarstrand, C. Petersson, and L. Svensson. Are nerfs ready for autonomous driving? towards closing the real-to-simulation gap. In *CVPRW*, 2024. 2

[19] W. E. Lorensen and H. E. Cline. Marching cubes: A high resolution 3d surface construction algorithm. *ACM TOG*, 21(4):163–169, 1987. 2

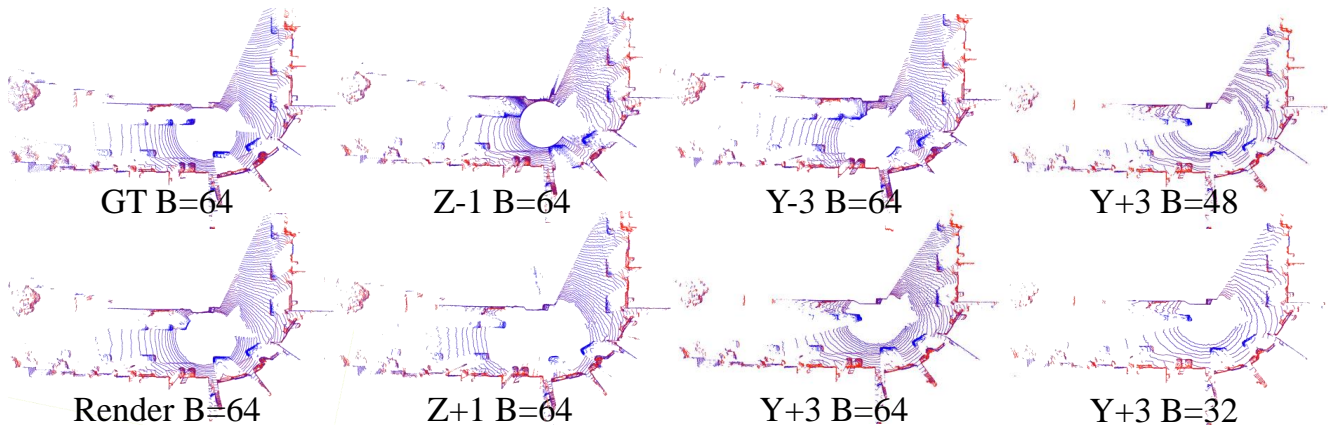


Figure 10: Novel View Synthesis application when re-simulating the LiDAR sensor. ‘Y’, ‘Z’ stands for left, and up, respectively. Positional offset unit is performed in meters. ‘B’ for number of laser beams to launch re-simulation.

- [20] T. Lu, M. Yu, L. Xu, Y. Xiangli, L. Wang, D. Lin, and B. Dai. Scaffold-gs: Structured 3d gaussians for view-adaptive rendering. In *CVPR*, pages 20654–20664, 2024. 2, 3, 6
- [21] S. Manivasagam, I. A. Bârsan, J. Wang, Z. Yang, and R. Urtasun. Towards zero domain gap: A comprehensive study of realistic lidar simulation for autonomy testing. In *ICCV*, pages 8238–8248, 2023. 2
- [22] S. Manivasagam, S. Wang, K. Wong, W. Zeng, M. Sazanovich, S. Tan, B. Yang, W.-C. Ma, and R. Urtasun. LiDARsim: Realistic lidar simulation by leveraging the real world. In *CVPR*, pages 11167–11176, 2020. 2, 6, 7, 8
- [23] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng. Nerf: Representing scenes as neural radiance fields for view synthesis. In *ECCV*, page 405–421, 2020. 2, 3
- [24] T. Müller. tiny-cuda-nn, 4 2021. 6
- [25] J. Ost, F. Mannan, N. Thuerey, J. Knodt, and F. Heide. Neural scene graphs for dynamic scenes. In *CVPR*, pages 2856–2865, 2021. 5
- [26] K. Ren, L. Jiang, T. Lu, M. Yu, L. Xu, Z. Ni, and B. Dai. Octree-gs: Towards consistent real-time rendering with lod-structured 3d gaussians. *arXiv:2403.17898*, 2024. 2
- [27] P. Sun, H. Kretschmar, X. Dotiwalla, A. Chouard, V. Patnaik, P. Tsui, J. Guo, Y. Zhou, Y. Chai, B. Caine, et al. Scalability in perception for autonomous driving: Waymo open dataset. In *CVPR*, pages 2446–2454, 2020. 2, 6
- [28] T. Tao, L. Gao, G. Wang, Y. Lao, P. Chen, H. Zhao, D. Hao, X. Liang, M. Salzmann, and K. Yu. LiDAR-NeRF: Novel lidar view synthesis via neural radiance fields. In *ACM MM*, pages 390–398, 2023. 2, 3, 6, 7, 8
- [29] T. Tao, G. Wang, Y. Lao, P. Chen, J. Liu, L. Lin, K. Yu, and X. Liang. AlignMiF: Geometry-aligned multimodal implicit field for lidar-camera joint synthesis. In *CVPR*, pages 21230–21240, 2024. 2, 6, 7, 8
- [30] T. Whelan, S. Leutenegger, R. F. Salas-Moreno, B. Glocker, and A. J. Davison. Elasticfusion: Dense slam without a pose graph. In *Robotics: Science and Systems*, volume 11, page 3, 2015. 2
- [31] H. Wu, X. Zuo, S. Leutenegger, O. Litany, K. Schindler, and S. Huang. Dynamic lidar re-simulation using compositional neural fields. In *CVPR*, pages 19988–19998, 2024. 2, 6, 8
- [32] Z. Wu, T. Liu, L. Luo, Z. Zhong, J. Chen, H. Xiao, C. Hou, H. Lou, Y. Chen, R. Yang, Y. Huang, X. Ye, Z. Yan, Y. Shi, Y. Liao, and H. Zhao. Mars: An instance-aware, modular and realistic simulator for autonomous driving. In *CICAI*, 2023. 1, 2
- [33] J. Xu, Z. Miao, D. Zhang, H. Pan, K. Liu, P. Hao, J. Zhu, Z. Sun, H. Li, and X. Zhan. INT: Towards infinite-frames 3d detection with an efficient framework. In *ECCV*, pages 193–209, 2022. 9
- [34] M. Yang, Y. Du, K. Ghasemipour, J. Tompson, D. Schuurmans, and P. Abbeel. Learning interactive real-world simulators. *arXiv:2310.06114*, 2023. 1, 2, 6, 8
- [35] Z. Ye, W. Li, S. Liu, P. Qiao, and Y. Dou. AbsGS: Recovering fine details in 3d gaussian splatting. In *ACM MM*, 2024. 2, 6
- [36] Z. Zheng, F. Lu, W. Xue, G. Chen, and C. Jiang. Lidar4d: Dynamic neural fields for novel space-time view lidar synthesis. In *CVPR*, pages 5145–5154, 2024. 2, 6, 8
- [37] C. Zhou, L. Fu, S. Peng, Y. Yan, Z. Zhang, Y. Chen, J. Xia, and X. Zhou. Lidar-rt: Gaussian-based ray tracing for dynamic lidar re-simulation. *arXiv:2412.15199*, 2024. 3, 6, 8
- [38] H. Zhou, L. Lin, J. Wang, Y. Lu, D. Bai, B. Liu, Y. Wang, A. Geiger, and Y. Liao. HUGSIM: A real-time, photo-realistic and closed-loop simulator for autonomous driving. *arXiv:2412.01718*, 2024. 2