# AtlasGS: Atlanta-world Guided Surface Reconstruction with Implicit Structured Gaussians

Scan the QR code for more details

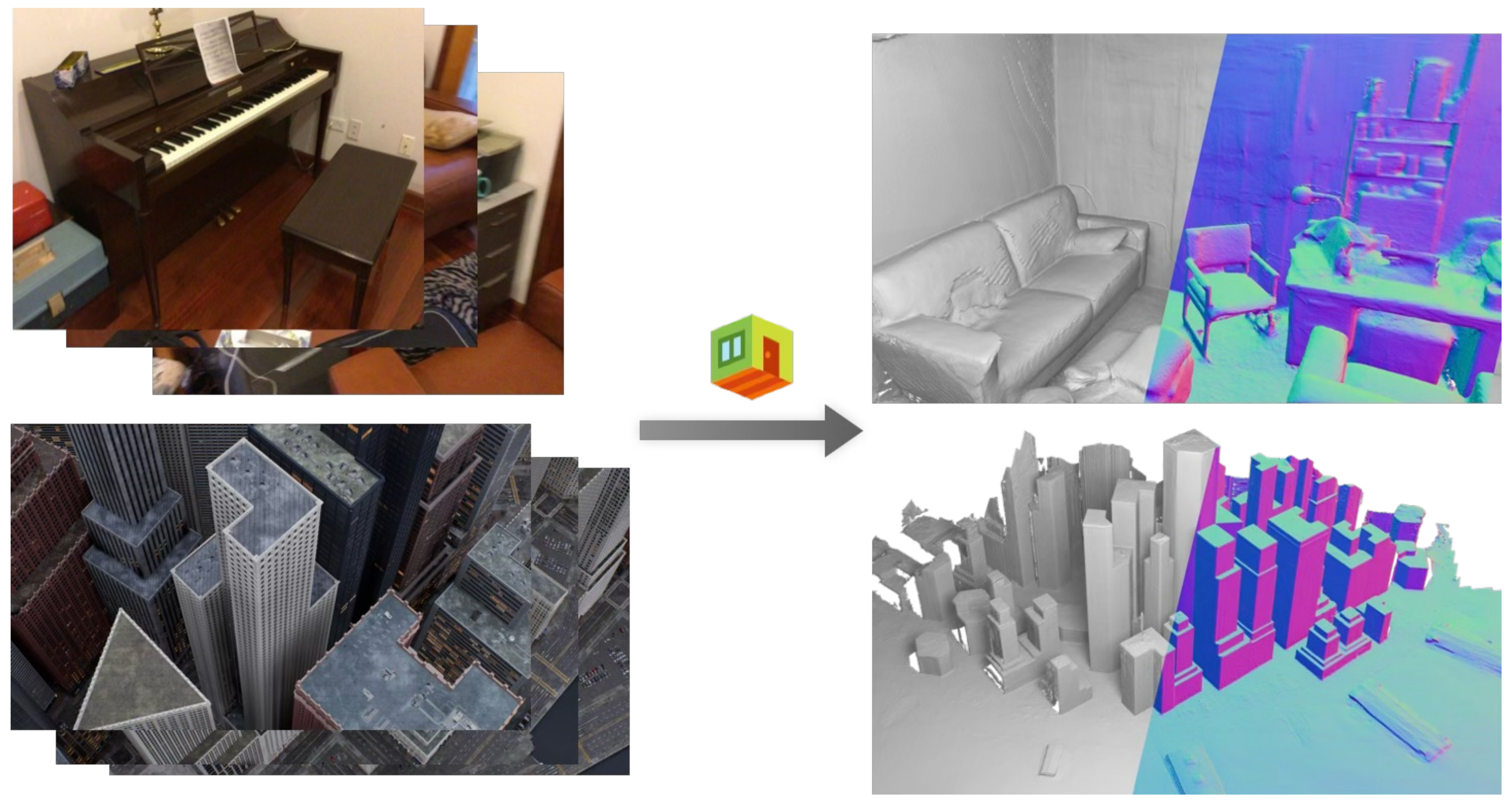
Project Page

Xiyu Zhang* · Chong Bao* · Yipeng Chen · Hongjia Zhai · Yitong Dong

Hujun Bao · Zhaopeng Cui · Guofeng Zhang†

State Key Lab of CAD & CG, Zhejiang University

## The Goal of AtlasGS



Dense posed Images of the **Structural** Scene, like indoor scenes and city scenes.

**Fast** and **Accurate** Surface Reconstruction

## Introduction

**TL;DR**: We propose a new Gaussian Splatting representation under Atlanta world assumption for indoor and urban surface reconstruction.

***Previous Methods***

Gaussian Based Methods (2DGS):

- Fast training and rendering (< 1h)
- Discrete representation leads to discontinuous Surfaces

NeRF Based Methods (MonoSDF):

- Smooth but low refenquency
- Hours for training (> 7h)

Both suffer from multiview inconsistency of monocular geometry priors

***Ours:***

**Implicit-Structured GS:** Combines MLP smoothness with Gaussian details.

**Structural Priors:** Atlanta-world guided regularization fixes Multiview inconsistency.

**SOTA Performance:** Fast training (< 30 mins) with accurate geometry.

## Methods

### *(a) Implicit Structured Gaussian Splatting*

We construct a sparse feature grid $\mathcal{V}$ and decode attributes for $K$ local Gaussians at voxel $v$ via geometry and semantic MLPs $\mathcal{M}_g, \mathcal{M}_s$

$$\{\alpha_k, s_k, q_k, c_k\} = \mathcal{M}_g(\mathcal{V}_g(v), d), \quad z_k = \mathcal{M}_s(\mathcal{V}_s(v))$$

$$p_k = v + l \cdot \Delta_k$$

Differentiable rasterization renders attributes (color $C$, depth $D$, semantics $Z$):

$$C = \sum_{i \in \mathcal{N}} c_i \alpha_i \prod_{j=1}^{i-1} (1 - \alpha_j)$$

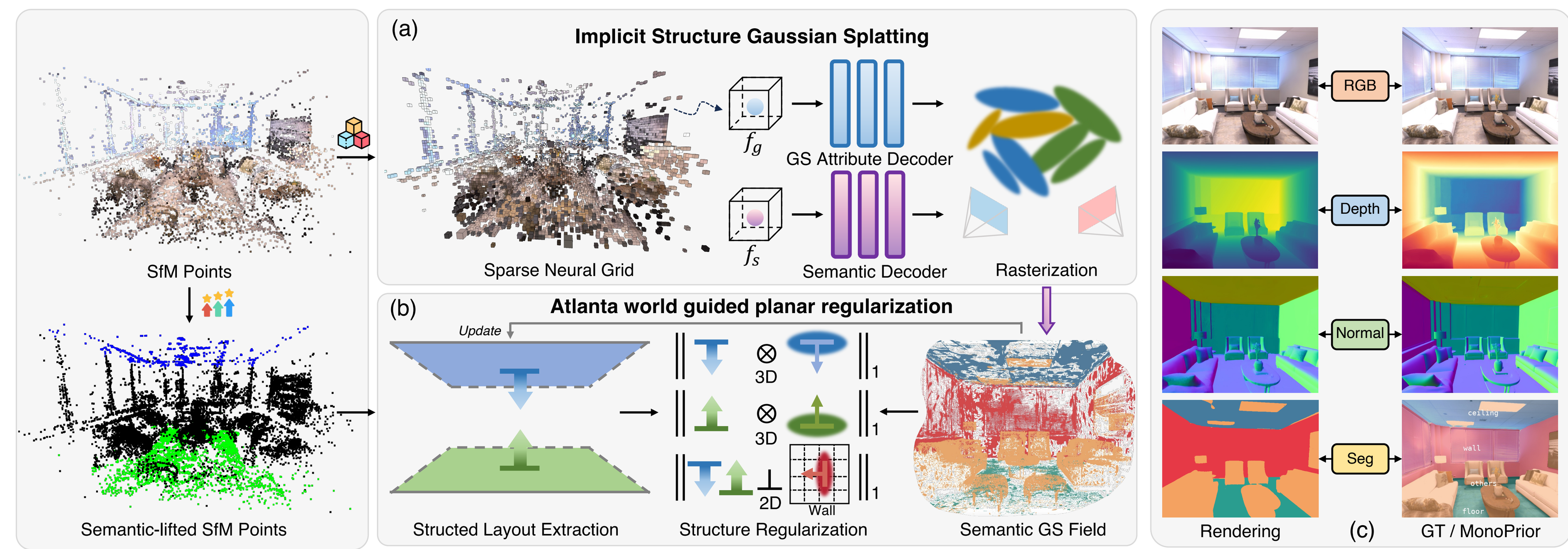



Figure 1: Overview of AtalsGS.

***(b) Atlanta-world Guided Planar Regularization*** We model the scene with Atlanta-world assumption (gravity $n_g$ and offsets $d$):

$$\pi_{floor} : d_f + n_g \cdot p = 0, \quad \pi_{ceiling} : d_c - n_g \cdot p = 0$$

We enforce global structure via **3D Global** $\mathcal{L}_{3D}$ and **2D Local** $\mathcal{L}_{2D}$ regularization:

$$\mathcal{L}_{3D} = \underbrace{\sum p_{\parallel}(1 - |n_g^\top n_i|) + \sum p_{\perp}|n_g^\top n_i|}_{\text{Normal Alignment}} + \underbrace{\sum p_c|d_c - n_g^\top p_i| + \sum p_f|d_f + n_g^\top p_i|}_{\text{Planar Constraint}}$$

$$\mathcal{L}_{2D} = \sum p_w |N_d \cdot n_g| + \sum p_{c,f}(1 - |N_d \cdot n_g|)$$

where $p_{\parallel}, p_{\perp}, p_c, p_f, p_w$ are probabilities for parallel, perpendicular, ceiling, floor, and wall regions.

***(c) Optimization*** The framework is jointly optimized with photometric loss, semantic supervision, and monocular geometry priors (depth $\widehat{D}$ and normal $\widehat{N}$):

$$\mathcal{L} = \mathcal{L}_{rgb} + \lambda_{sem}\mathcal{L}_{sem} + \lambda_{reg}\mathcal{L}_{reg} + \lambda_{prior}(\mathcal{L}_{depth} + \mathcal{L}_{normal})$$

## Experiments

### *Quantitative Results of Surface Reconstruction*

Table 1: Quantitative comparison of on indoor scenes.

| Category | Methods | Replica [52] | | | | | ScanNet++ [51] | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Acc↓ | Comp↓ | Prec↑ | Recall↑ | F-score↑ | Acc↓ | Comp↓ | Prec↑ | Recall↑ | F-score↑ |
| Implicit | ManhattanSDF [13] | 4.76 | 5.59 | 68.80 | 66.40 | 67.57 | 3.96 | 4.98 | 77.30 | 76.16 | 76.67 |
| | MonoSDF [14] | 4.14 | 5.38 | 75.50 | 70.89 | 73.08 | 4.05 | 5.57 | 76.25 | 75.16 | 75.65 |
| Explicit | Scaffold-GS [54] | 8.58 | 11.27 | 63.53 | 54.91 | 58.89 | 23.11 | 15.36 | 28.18 | 31.75 | 29.78 |
| | 2DGS [10] | 4.76 | 6.34 | 74.54 | 65.37 | 69.64 | 16.53 | 17.91 | 22.84 | 20.79 | 21.71 |
| | DN-Splatter [44] | 16.97 | 15.52 | 32.67 | 30.90 | 31.75 | 16.77 | 15.05 | 22.42 | 21.97 | 22.16 |
| | GSRec [17] | 4.90 | 6.89 | 73.32 | 67.69 | 70.37 | 9.37 | 9.13 | 47.12 | 53.81 | 50.14 |
| | Ours | 2.25 | 4.08 | 93.18 | 82.22 | 87.35 | 3.22 | 4.09 | 87.59 | 87.47 | 87.48 |

Table 2: Quantitative comparison of on ScanNet.

| Category | Methods | Acc↓ | Comp↓ | Prec↑ | Recall↑ | F-score↑ | Time ↓ | FPS ↑ |
|---|---|---|---|---|---|---|---|---|
| Implicit | ManhattanSDF [13] | 4.25 | 5.23 | 72.39 | 63.18 | 67.25 | > 7 h | < 10 |
| | MonoSDF [14] | 4.25 | 4.76 | 73.53 | 69.18 | 71.21 | > 7 h | < 10 |
| Explicit | Scaffold-GS [54] | 9.47 | 7.99 | 51.41 | 49.08 | 50.17 | 12 mins | 279 |
| | 2DGS [10] | 11.46 | 13.89 | 43.15 | 36.17 | 39.27 | 11 mins | 118 |
| | DN-Splatter [44] | 13.54 | 14.77 | 21.71 | 18.94 | 20.22 | 12 mins | 145 |
| | GSRec [17] | 6.71 | 5.40 | 60.36 | 66.63 | 63.30 | 35 mins | 261 |
| | Ours | 3.62 | 3.93 | 80.31 | 75.85 | 77.98 | 27 mins | 70 |

Table 3: Quantitative comparison of on Matrix City.

| Methods | Acc↓ | Comp↓ | CD↓ |
|---|---|---|---|
| GaussianPro [59] | 0.102 | 0.081 | 0.091 |
| Scaffold-GS [54] | 0.328 | 0.303 | 0.316 |
| GSRec [17] | 0.048 | 0.175 | 0.112 |
| 2DGS [10] | 0.115 | 0.098 | 0.106 |
| Ours | 0.022 | 0.034 | 0.028 |

### *Qualitative Results of Surface Reconstruction*

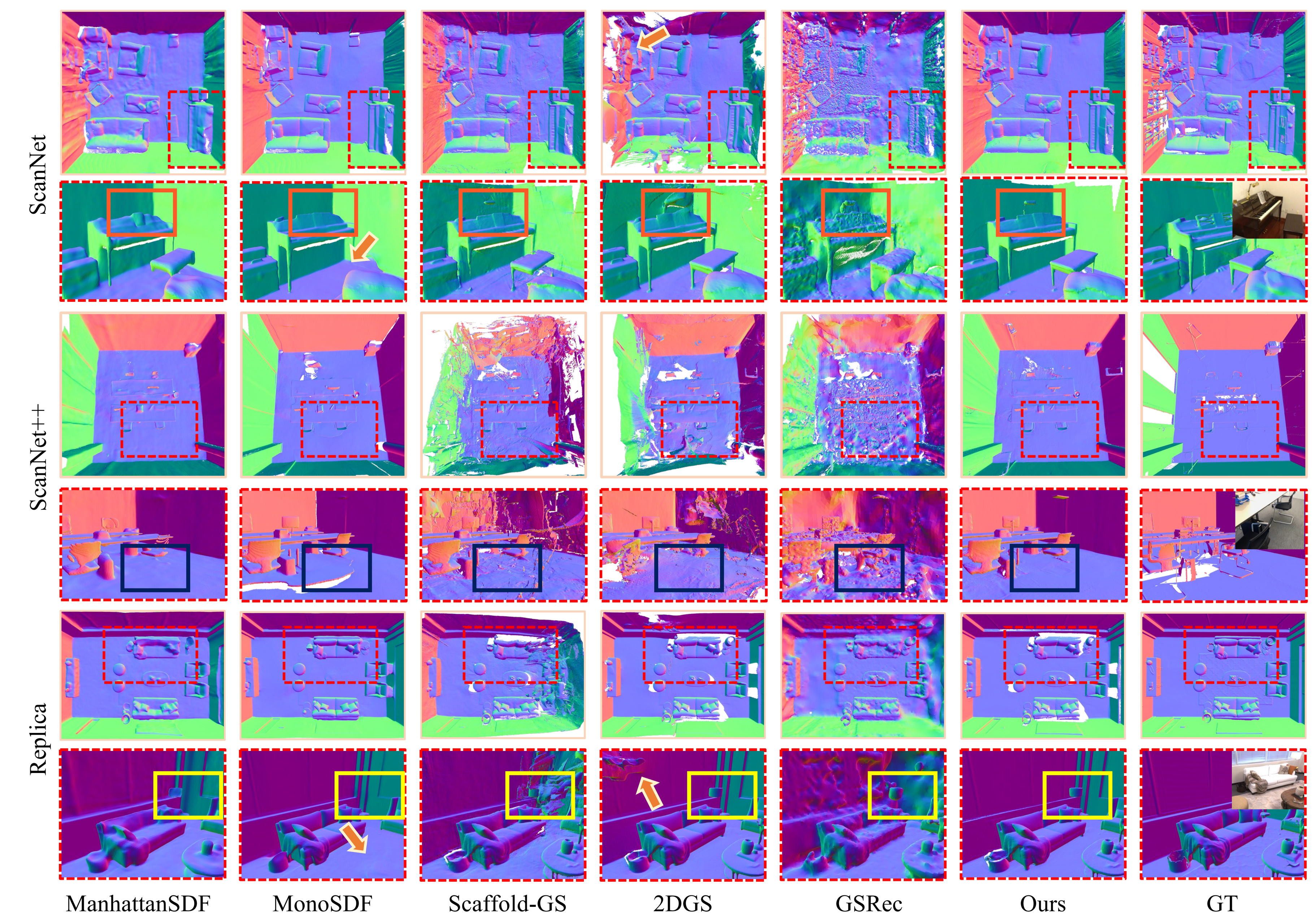



Figure 2: Qualitative comparison of indoor scene reconstruction.

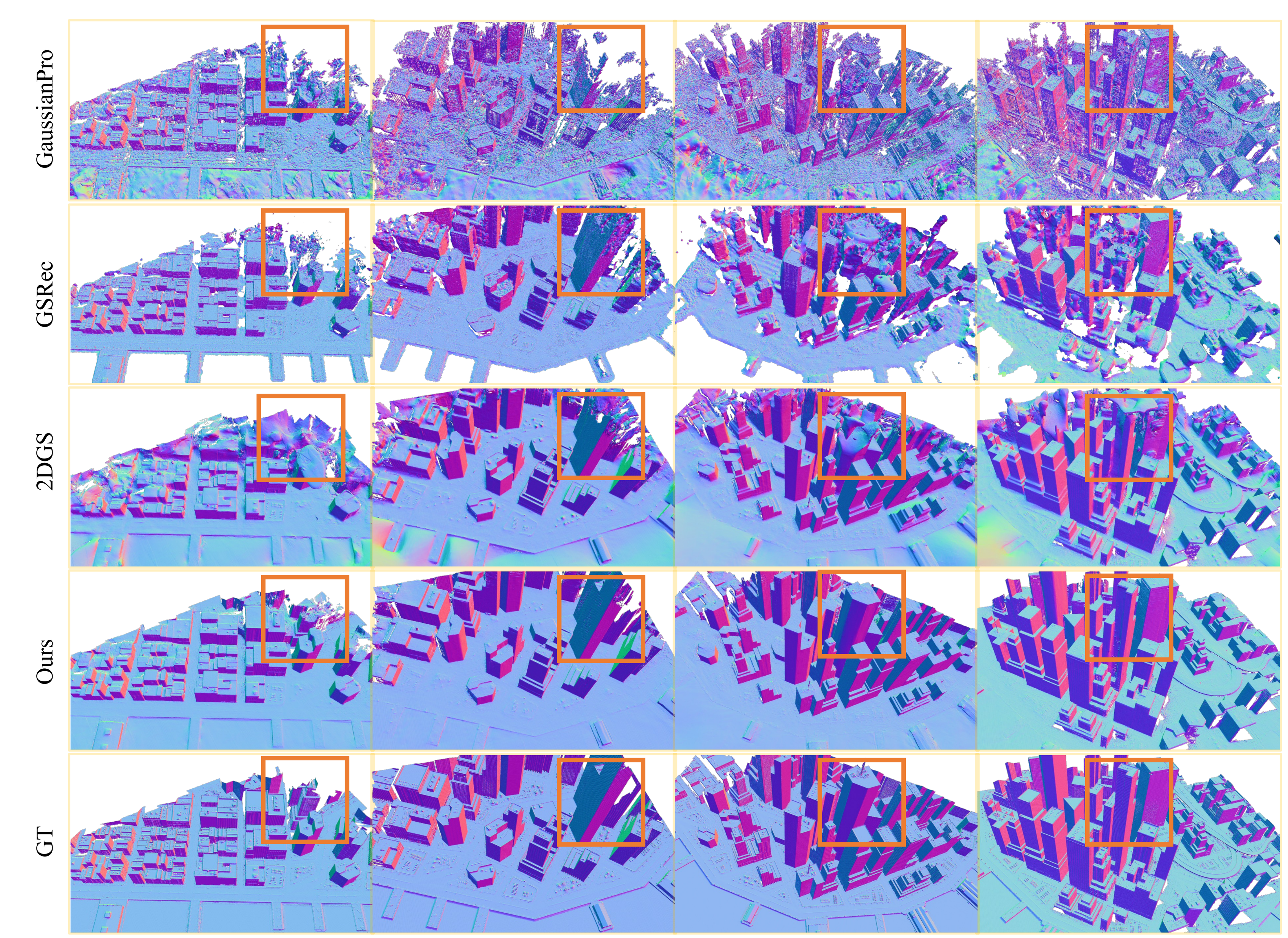



Figure 3: Qualitative comparison of outdoor scene reconstruction.