OccluGaussian: Occlusion-Aware Gaussian Splatting for Large Scene Reconstruction and Rendering

Shiyong Liu[1] Xiao Tang[1] Zhihao Li[1] Yingfan He[2]
ChongJie Ye[2] Jianzhuang Liu[3] Binxiao Huang[4] Shunbo Zhou[5] Xiaofei Wu[1]

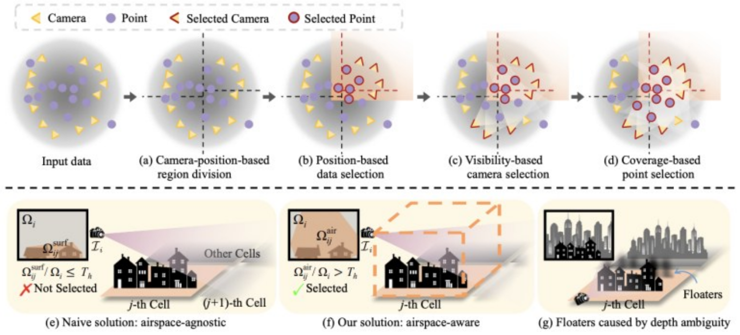
[1]Huawei Noah's Ark Lab [2]The Chinese University of HongKong (Shenzhen)
[3]Shenzhen Institute of Advanced Technology [4]The University of HongKong
[5]Huawei Embodied Intelligence Lab

NOAH'S ARK LAB
香港中文大學(深圳) The Chinese University of Hong Kong, Shenzhen
SIAT Shenzhen Institutes of Advanced Technology, Chinese Academy of Science
香港大學 THE UNIVERSITY OF HONG KONG
HUAWEI

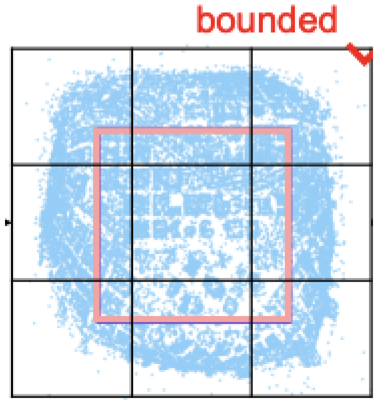
Introduction

- Divide-and-conquer strategies are widely adopted in large-scale scene reconstruction to address resource constraints.
- By splitting the scene into smaller and more manageable regions, each region is reconstructed independently and finally merged to a complete model.

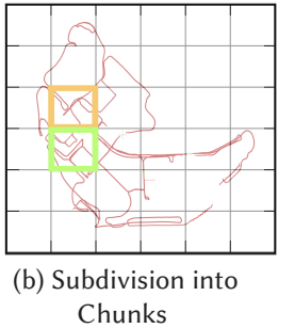
- **Existing scene division strategies are mainly based on camera positions or point clouds**



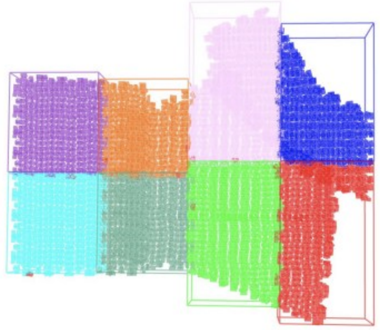
VastGaussian
(CVPR 2024)



CityGaussian
(ECCV 2024)



Hierarchical-GS
(SIGGRAPH 2024)



DOGS
(NeurIPS 2024)

Introduction

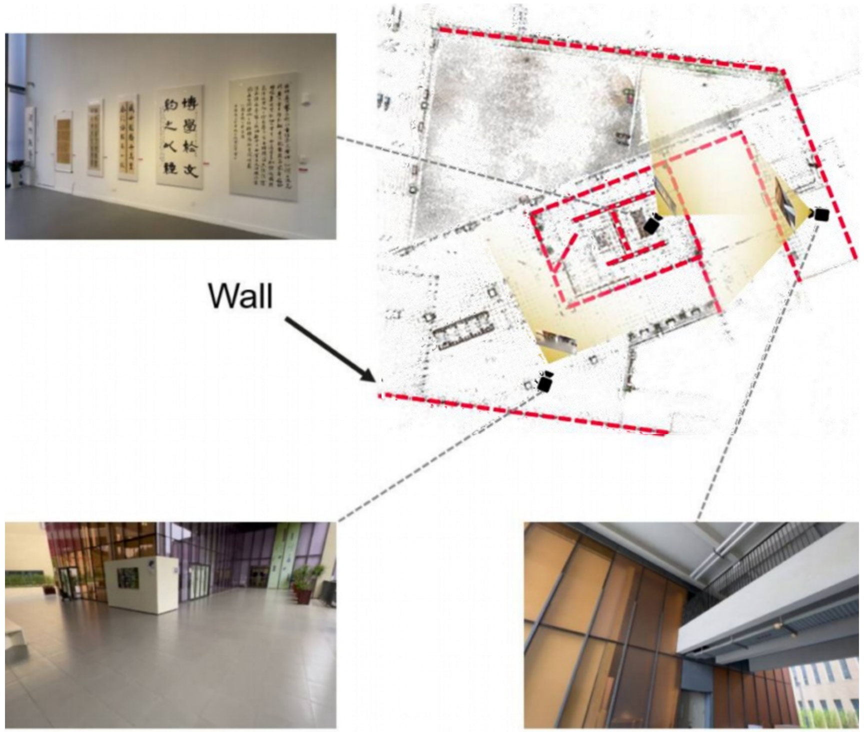
- **Current scene division strategies:**
  - Works for occlusion-free scenes like aerial imagery or open spaces.
  - Less effective in ground-level capture scenarios due to frequent occlusions.

Scene w/o occlusions

Scene w/ occlusions

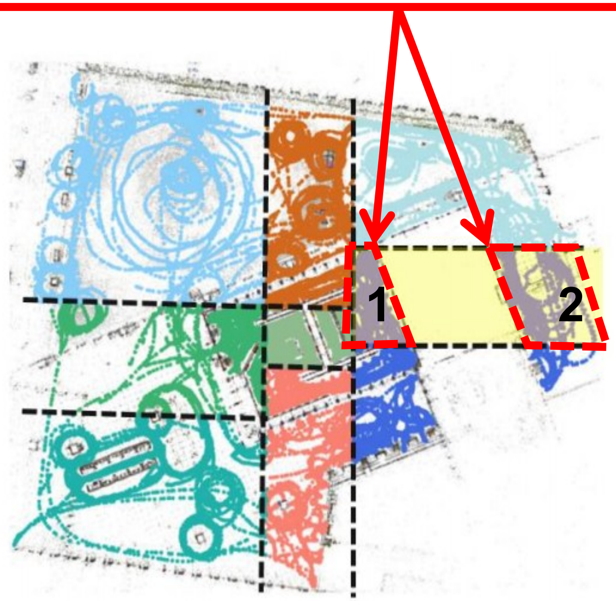
Challenges of large scene division

- **Existing scene division strategies fail to account for scene layout and occlusions.**
  - Occlusion-agnostic division splits scenes into mutually occluded regions.
  - Cameras consume training resources, but mutual occlusion reduces their average contribution to targets.



A large scene with frequent occlusions

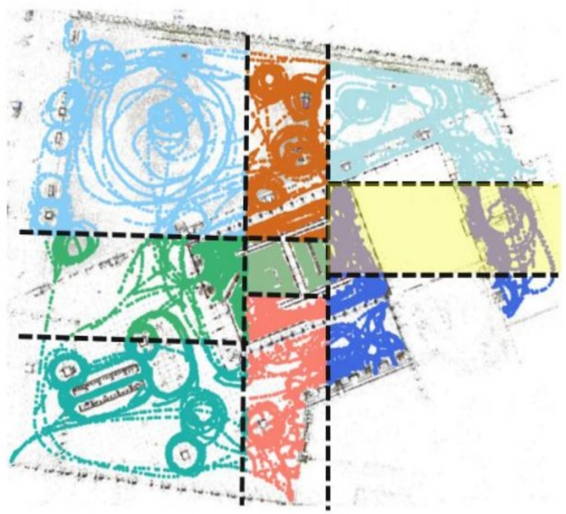
Mutual occlusion prevents cameras in area 1 from contributing to those in area 2.



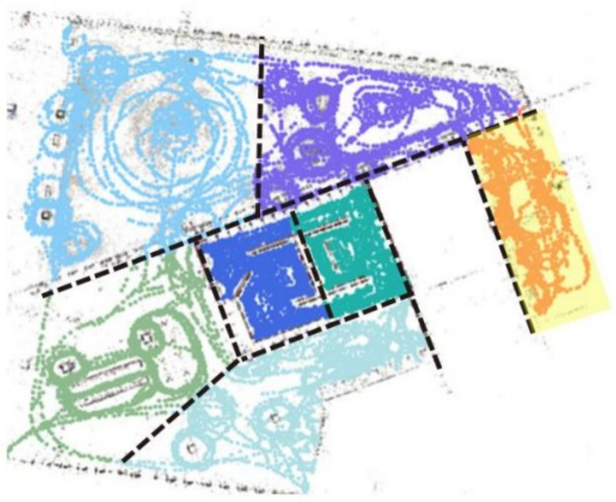
Occlusion-agnostic division

Motivation

- **An occlusion-aware scene division strategy that accounts for scene layout and occlusions:**
  - ➢ Stronger camera correlation within regions boosts mutual reconstruction quality.
  - ➢ Enhancing training efficiency and resource allocation.

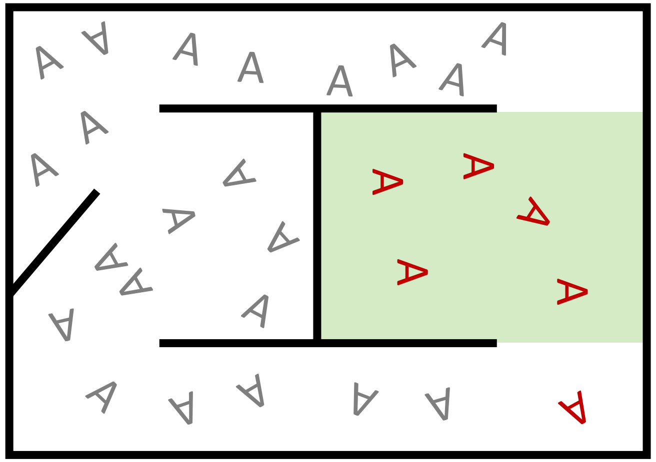


Occlusion-agnostic division



Occlusion-aware division

Enhanced average camera contribution within regions

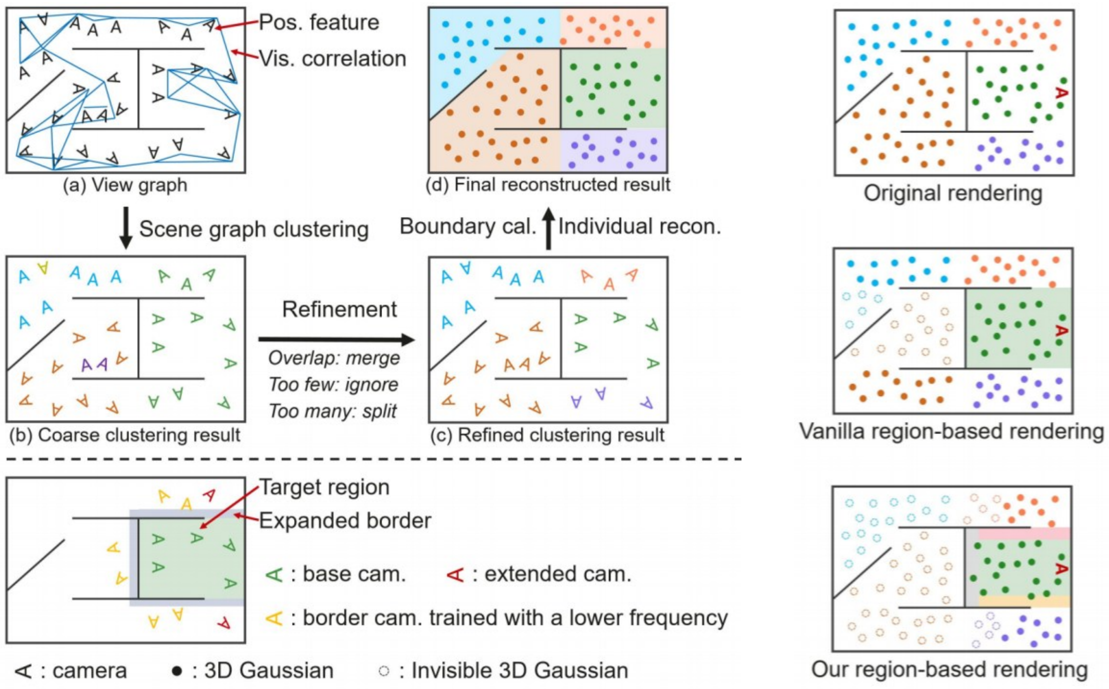


: Unselected camera

: Selected camera

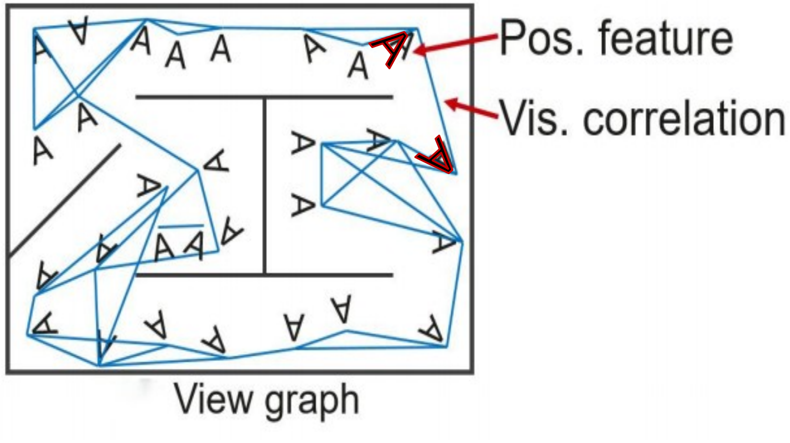
Technical contribution

1. **We propose an occlusion-aware scene division strategy that aligns with scene layouts, enhancing camera correlation within regions and improving average contribution to overall reconstruction.**
2. **We introduce a region-based rendering acceleration technique that significantly boosts rendering speed.**



Occlusion-aware scene division

Building an occlusion-aware attributed scene graph



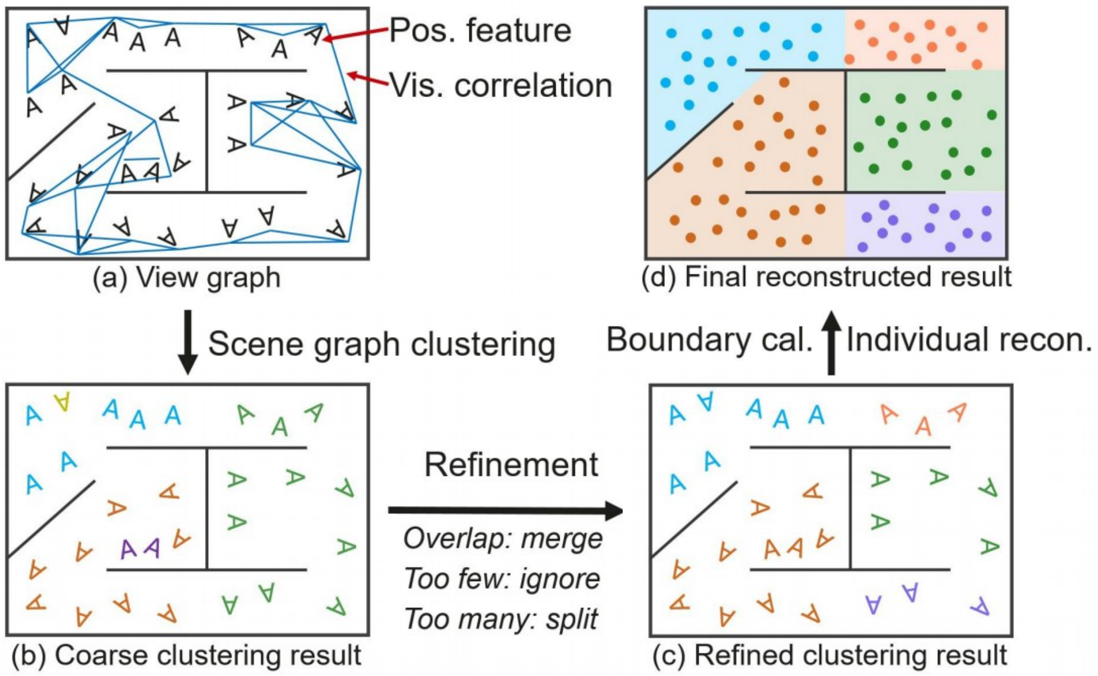
- **Node:** Camera with position features.
- **Edge:** Mutual visibility between two cameras.
- **Edge weight:** Number of feature point matches between two images.

Advantages: Node positional and edge image features are complementary.

1. **Distinguish visually similar but spatially distant cameras.**
2. **Distinguish spatially close but visually distinct cameras.**

Occlusion-aware scene division

Scene graph clustering



Adaptive clustering number determination:

Boundary calculation: Training a linear classification model to derive decision functions as boundary lines.

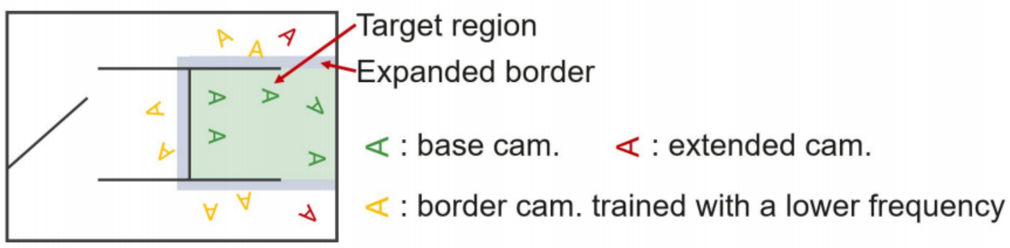
Setting an initial clustering number, then perform graph convolution-based clustering.

Refinement:

- **Too many:** Splitting clusters that contain too many cameras by further applying graph clustering.
- **Too few** Ignoring any cluster that either has too few cameras or whose convex hull is entirely covered by the convex hull of another cluster.
- **Overlap:** If two clusters' convex hulls overlap by more than half the area of the smaller hull, we merge them.

Individual region reconstruction

- **Training camera selection:**
  - ➢ Insufficient cameras cannot provide adequate supervision, leading to artifacts and floaters.
  - ➢ Excessive irrelevant cameras reduce average contribution to region reconstruction, leading to blurry details.



Training camera selection strategy:

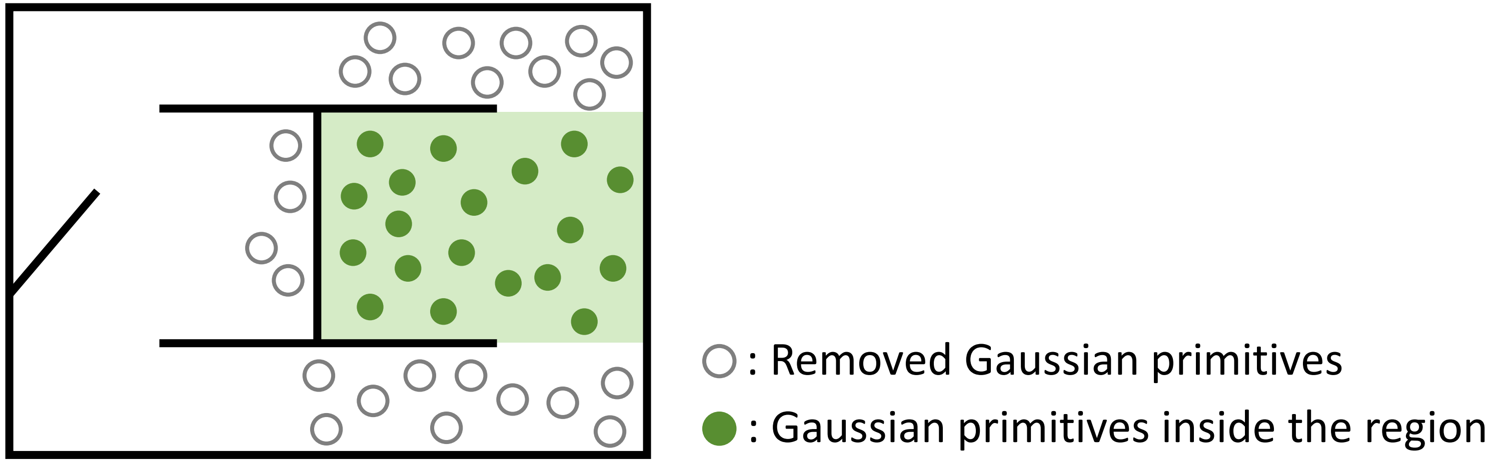
Base cameras: Located within the region.

Extended cameras: Located outside the region but acquiring sufficient visual information of it.

Border cameras: Facing the region but occluded. These cameras help constrain Gaussian primitives near the boundaries for the final seamless merging.

Seamless region merging

1. Removing Gaussian primitives outside the region to create sharp borders.
2. Finally merge regions to form a complete model.



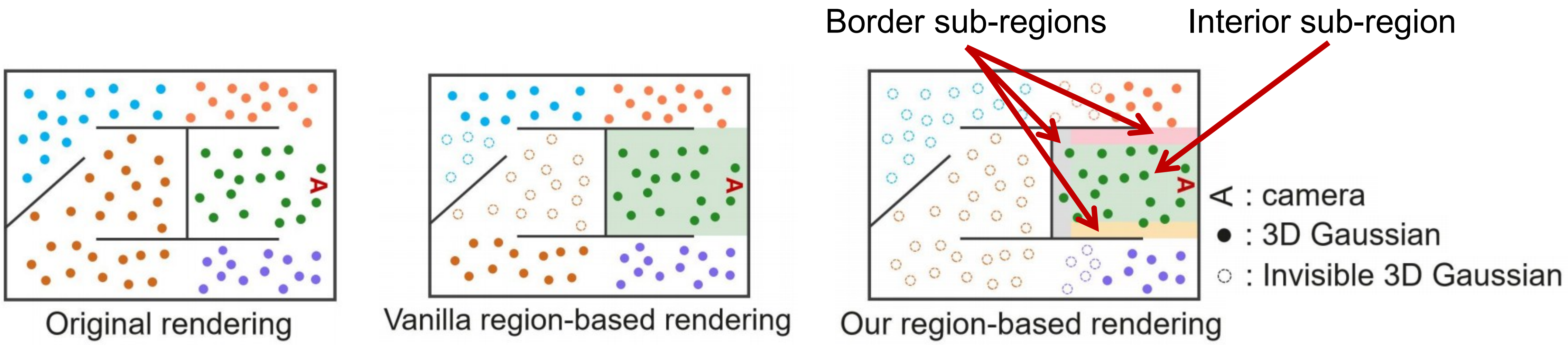
Region-based rendering

Massive 3D Gaussians incur high rendering costs. Pre-culling invisible primitives via occlusion/visibility reduces load while preserving visual fidelity.

1. Region-based visibility calculation: Record the 3D Gaussians visible from all viewpoints in the region.

2. Region subdivision: Subdividing the region reduces 3D Gaussians from neighboring regions.

3. Rendering with region-based culling: Identify the viewpoint's region, cull invisible 3D Gaussians, then render.



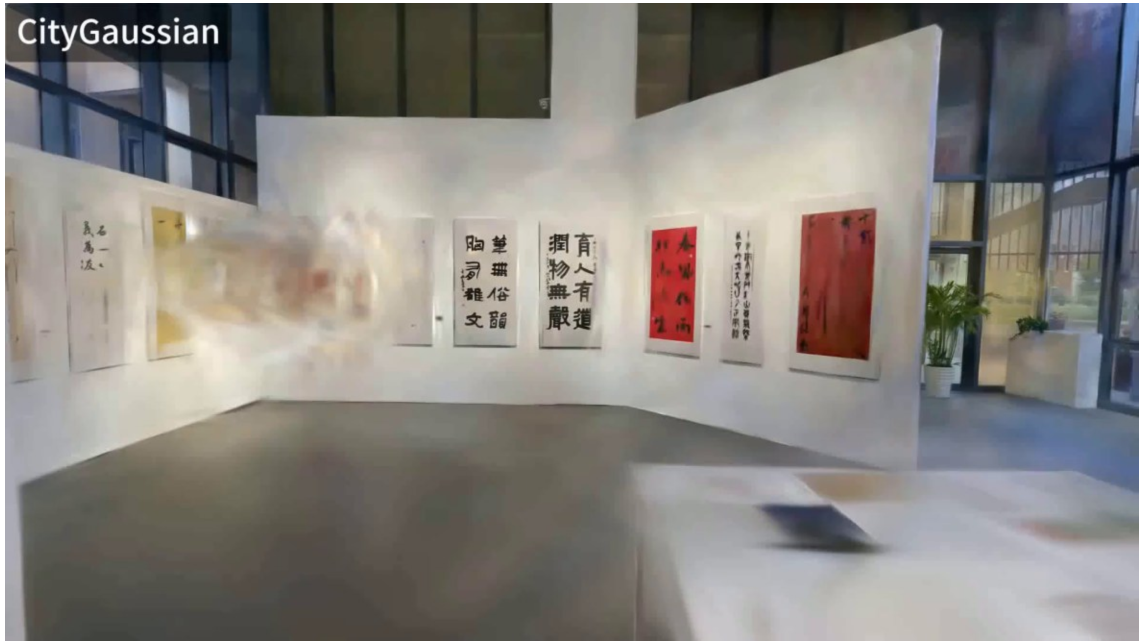
Experiments

- The rendering quality surpasses the existing SOTA methods on multiple datasets.
- Lossless rendering acceleration.

Scene	GALLERY				CANTEEN				CLASSBUILDING			
Metrics	PSNR	SSIM	LPIPS	FPS	PSNR	SSIM	LPIPS	FPS	PSNR	SSIM	LPIPS	FPS
VastGaussian* [28]	<u>25.09</u>	**0.903**	<u>0.095</u>	215.22	<u>24.60</u>	<u>0.890</u>	<u>0.105</u>	211.02	<u>24.05</u>	<u>0.884</u>	<u>0.111</u>	269.97
CityGaussian [32]	21.98	0.808	0.294	119.86	20.41	0.794	0.275	54.02	20.48	0.840	0.244	65.57
Hierarchical-GS [22]	22.23	0.800	0.182	216.00	22.71	0.825	0.178	199.33	23.87	0.881	0.128	198.58
3DGS [21]	21.36	0.843	0.213	**344.92**	21.86	0.847	0.183	**525.93**	19.41	0.871	0.186	**395.13**
OccluGaussian	**25.81**	**0.903**	**0.094**	<u>288.94</u>	**25.25**	**0.900**	**0.100**	<u>311.59</u>	**25.33**	**0.921**	**0.083**	<u>339.64</u>

Table 1. Quantitative comparison on the OccluScene3D dataset. We report SSIM ↑, PSNR ↑, LPIPS ↓ and FPS ↑ on the test views. The **best** and <u>second best</u> results are highlighted. * denotes that it is our re-implementation of VastGaussian.

Experiments



Thank You!

Link to project page!