Survey: Large-scale 3DGS

刘峥

2024年11月29日

饮 水 思 源 • 爱 国 荣 校

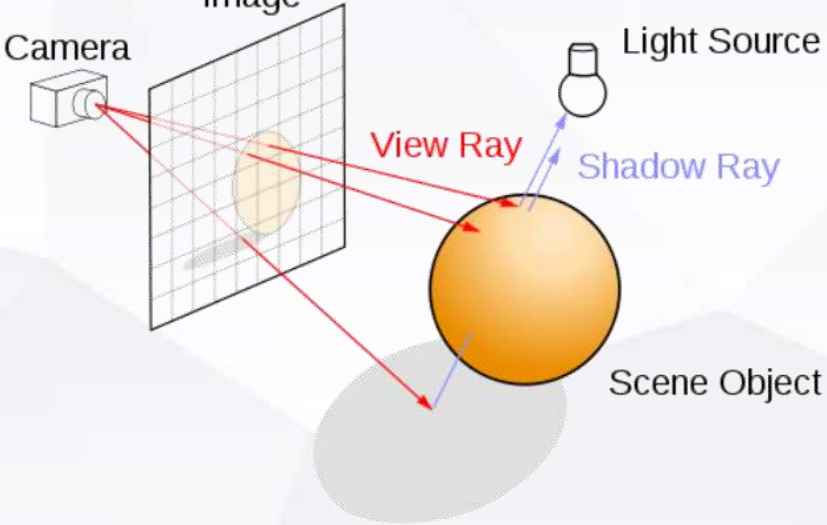
3D Gaussian Splatting: Background

- Traditional 3D Reconstruction
  - Local illumination
    - Lambert, Phong, Blinn-Phong, Cook-torrance
    - Simple, quick but unreal
  - Global illumination like ray-tracing
    - Realistic but compliate

- Neural Radiance Field

- 3D Gaussian Splatting

◆局部光照 ◆全局光照

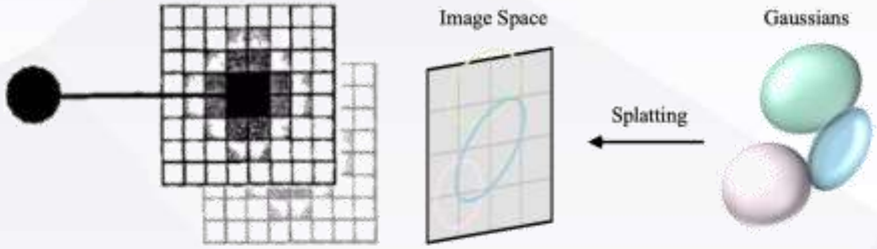


Source from

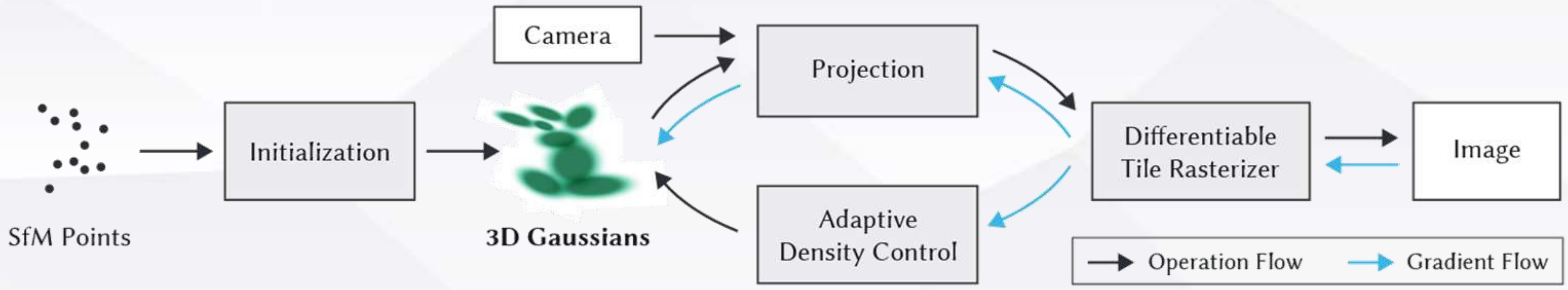
3D Gaussian Splatting: Background

3D Gaussian Splatting for Real-Time Radiance Field Rendering

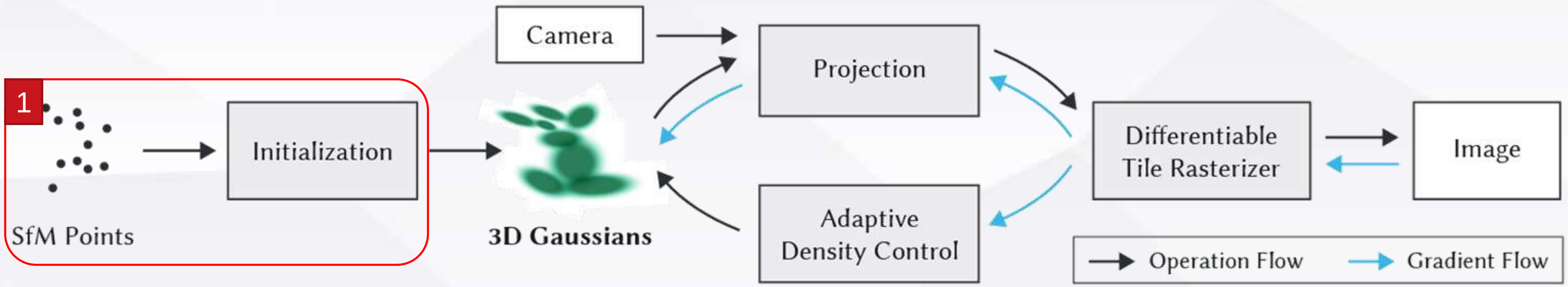
BERNHARD KERBL*, Inria, Université Côte d'Azur, France
GEORGIOS KOPANAS*, Inria, Université Côte d'Azur, France
THOMAS LEIMKÜHLER, Max-Planck-Institut für Informatik, Germany
GEORGE DRETTAKIS, Inria, Université Côte d'Azur, France



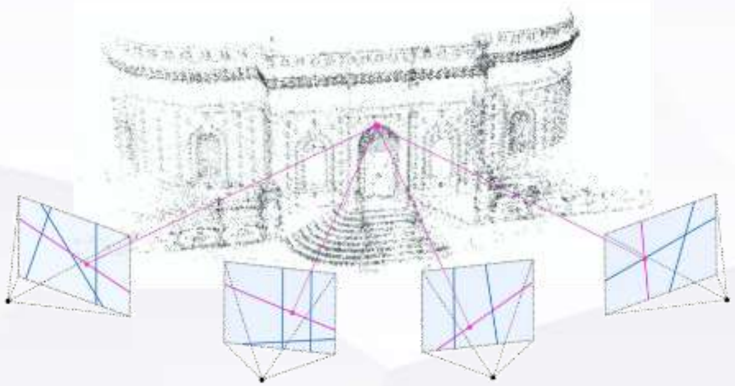
3D Gaussian Splatting: Method



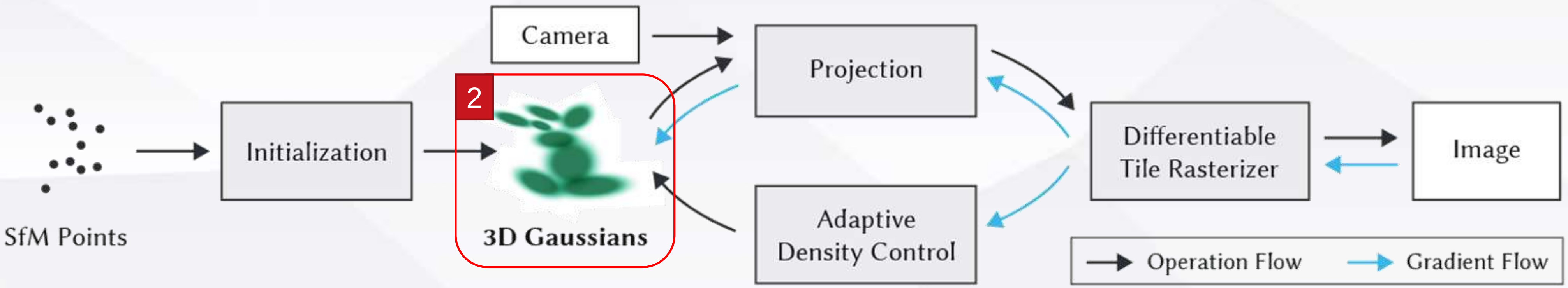
3D Gaussian Splatting: Method



- SfM： Structure from Motion



3D Gaussian Splatting: Method



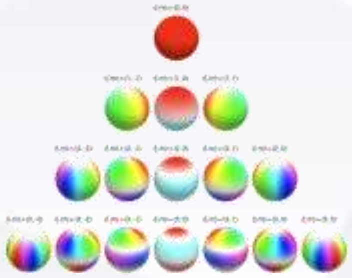
- 3D Gaussian's feature:
  - 3D Position $\rightarrow \mu$
  - Anisotropic Covariance Matrix $\rightarrow \Sigma$
    - Rotation Matrix
    - Scaling Matrix
  - Opacity $\rightarrow \alpha \in [0, 1)$
  - Spherical Harmonic Coefficients 16 * 3

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\{-\frac{(x-\mu)^2}{2\sigma^2}\}$$

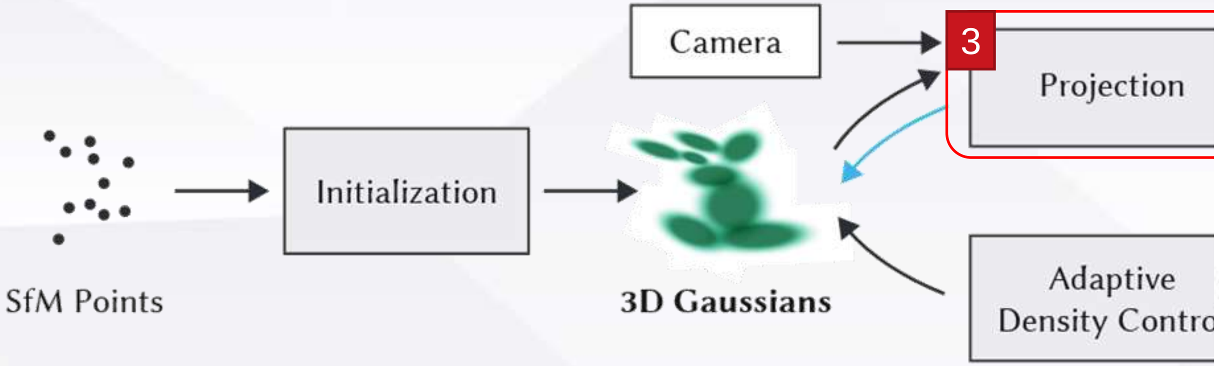
$$G(x, \mu) = \frac{1}{\sqrt{(2\pi)^3 \Sigma}} exp\{-\frac{1}{2}(x-\mu)^T \Sigma^{-1} (x-\mu)\}$$

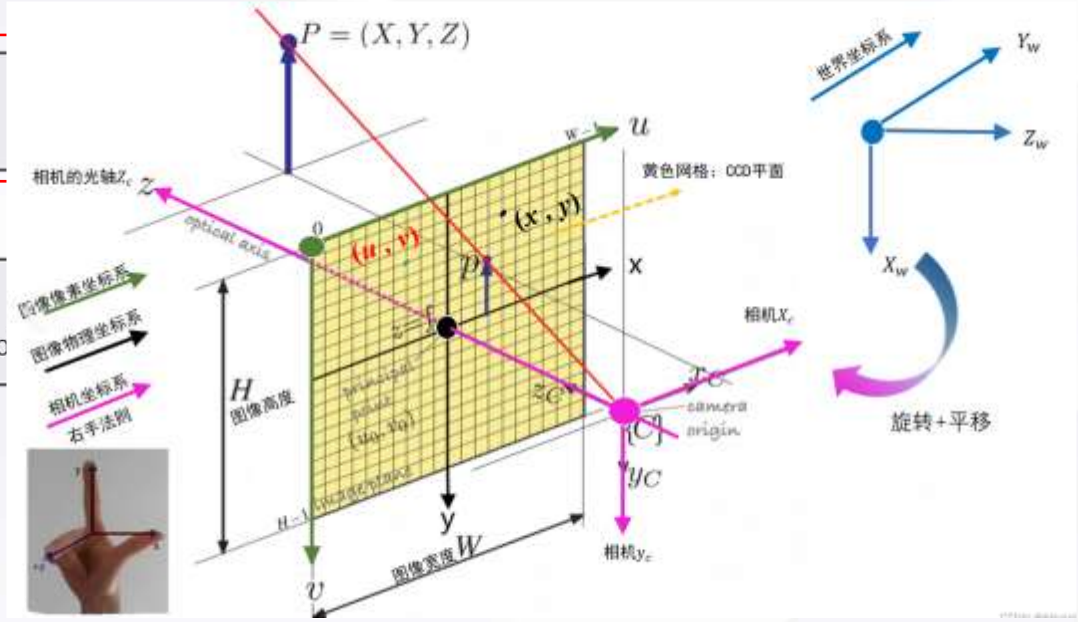
$$\Sigma = RSS^TR^T$$

$$\Sigma = \begin{bmatrix} \sigma_x^2 & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_y^2 & \sigma_{yz} \\ \sigma_{zx} & \sigma_{yz} & \sigma_z^2 \end{bmatrix}$$



3D Gaussian Splatting: Method



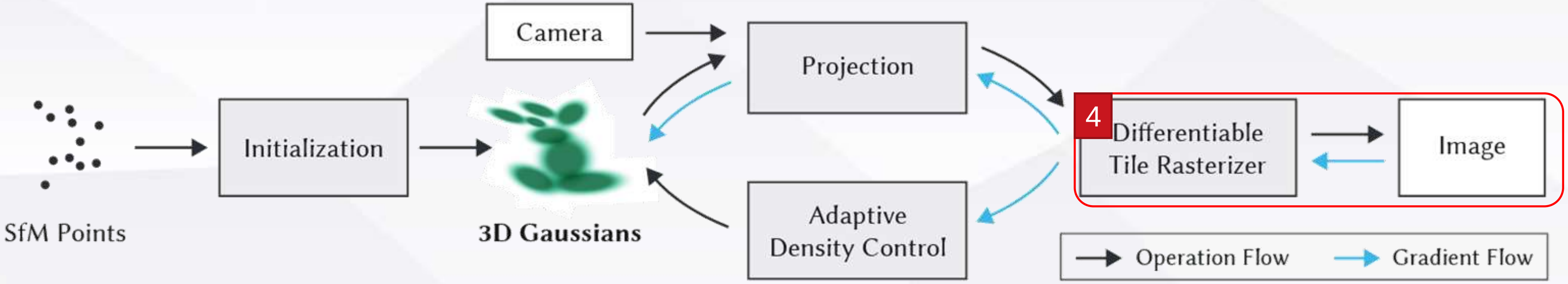


- Projection
  - View Transformation
  - Projection Transformation (perspective or orthographic ?)

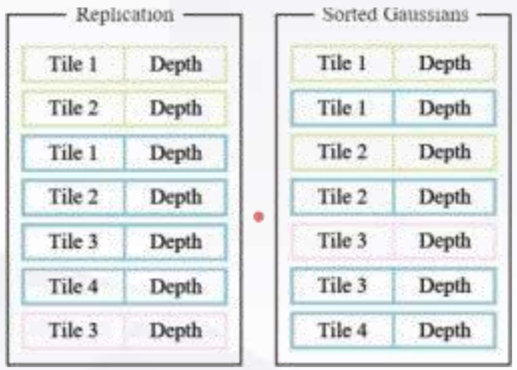
World Space $\xrightarrow{W = Ax + b}$ Camera Space $\xrightarrow{x_K = m(\mu_k)}$ Image Space (2D)

World Space	Camera Space	Image Space (2D)
μ	$\mu' = A\mu + b$	$\mu'' = A\mu + b$
Σ	$\Sigma' = W\Sigma W^T$	$\Sigma'' = JW\Sigma W^T J^T$

3D Gaussian Splatting: Method



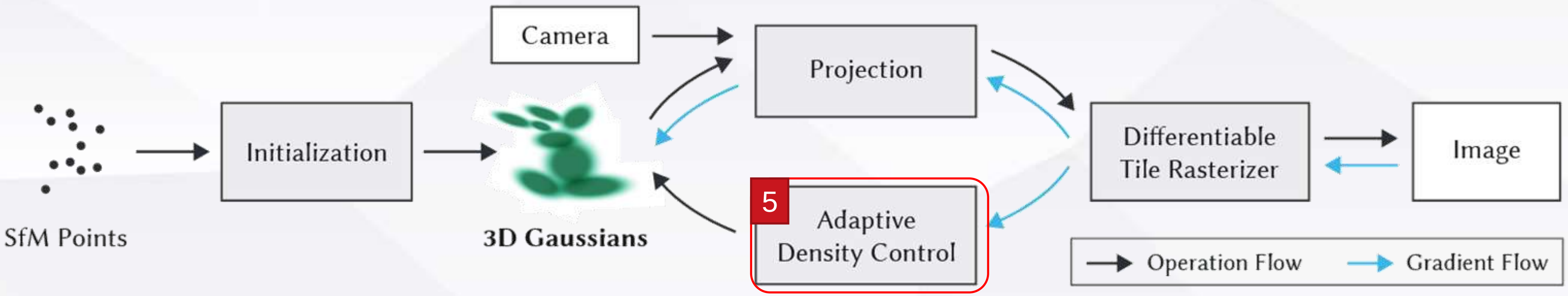
- Tile-based Rasterizer
    1. Pre-process
    2. Duplicate
    3. Radix sort
    4. α-blend
- Loss Function
    - $\mathcal{L} = (1-\lambda)\mathcal{L}_1 + \lambda\mathcal{L}_{\text{D-SSIM}}$



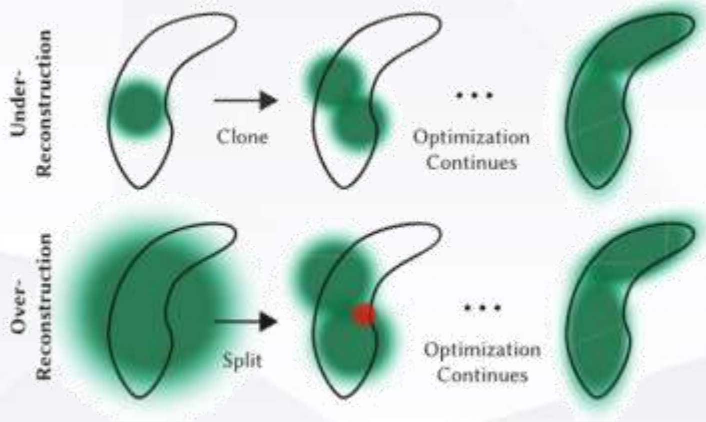
$$C = \sum_{i=1}^{N} T_i \alpha_i \mathbf{c}_i,$$

$$\alpha_i = (1 - \exp(-\sigma_i \delta_i)) \quad T_i = \prod_{j=1}^{i-1} (1 - \alpha_i).$$

3D Gaussian Splatting: Method



- Adaptive Desity Control
  - Clone: Under-reconstruction
  - Split: Over-reconstruction
  - Remove Gaussians
    - whose opacity are lower than threshold
    - which are very large in world space or have a big footprint in viewspace periodically



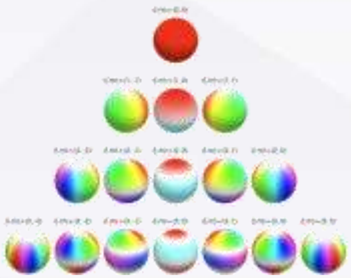
3D Gaussian Splatting: Evaluation Metrics & Results

1. PSNR (Peak Signal-to-Noise Ratio) ↑
2. SSIM (Structural Similarity Index Measure) ↑
3. LPIPS (Learned Perceptual Image Patch Similarity) ↓
4. FPS (Frame Per Seconds for inference rendering with A6000) ↑
5. Train (training time for A6000) ↓

Dataset	Mip-NeRF360						Tanks&Temples						Deep Blending					
Method|Metric	*SSIM*↑	*PSNR*↑	*LPIPS*↓	Train	FPS	Mem	*SSIM*↑	*PSNR*↑	*LPIPS*↓	Train	FPS	Mem	*SSIM*↑	*PSNR*↑	*LPIPS*↓	Train	FPS	Mem
Plenoxels	0.626	23.08	0.463	25m49s	6.79	2.1GB	0.719	21.08	0.379	25m5s	13.0	2.3GB	0.795	23.06	0.510	27m49s	11.2	2.7GB
INGP-Base	0.671	25.30	0.371	5m37s	11.7	13MB	0.723	21.72	0.330	5m26s	17.1	13MB	0.797	23.62	0.423	6m31s	3.26	13MB
INGP-Big	0.699	25.59	0.331	7m30s	9.43	48MB	0.745	21.92	0.305	6m59s	14.4	48MB	0.817	24.96	0.390	8m	2.79	48MB
M-NeRF360	0.792†	27.69†	0.237†	48h	0.06	8.6MB	0.759	22.22	0.257	48h	0.14	8.6MB	0.901	29.40	0.245	48h	0.09	8.6MB
Ours-7K	0.770	25.60	0.279	6m25s	160	523MB	0.767	21.20	0.280	6m55s	197	270MB	0.875	27.78	0.317	4m35s	172	386MB
Ours-30K	0.815	27.21	0.214	41m33s	134	734MB	0.841	23.14	0.183	26m54s	154	411MB	0.903	29.41	0.243	36m2s	137	676MB

SJTU

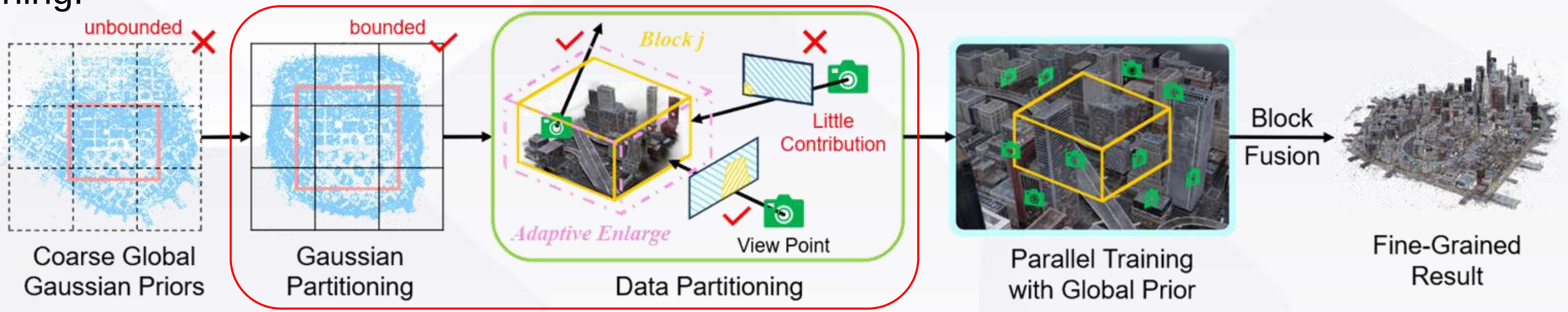
3D Gaussian Splatting: Research Directions



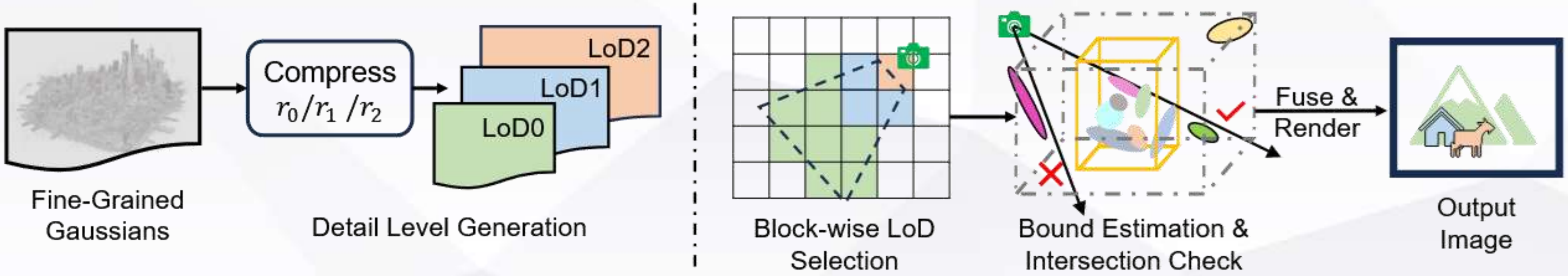
- Data efficient 3DGS
  - Sparse data leads to inaccuracies
- Memory efficient 3DGS
  - Reduce the number of 3D Gaussians
  - Compress the memory usage of 3D Gaussian properties
- Photorealistic 3DGS
- 3DGS with more properties
  - Scene understanding
- **Large-scale 3DGS**
  - Challenges Caused by Large-scale (over 1.5km^2)
    - Prohibitive overhead in GPU memory during training.
    - Rendering speed bottleneck lies in sorting.
  - Optimization
    - Divide-and-conquer training approach.
    - Level of detail strategy for rendering.

Large-scale 3D Gaussian Splatting: CityGaussian

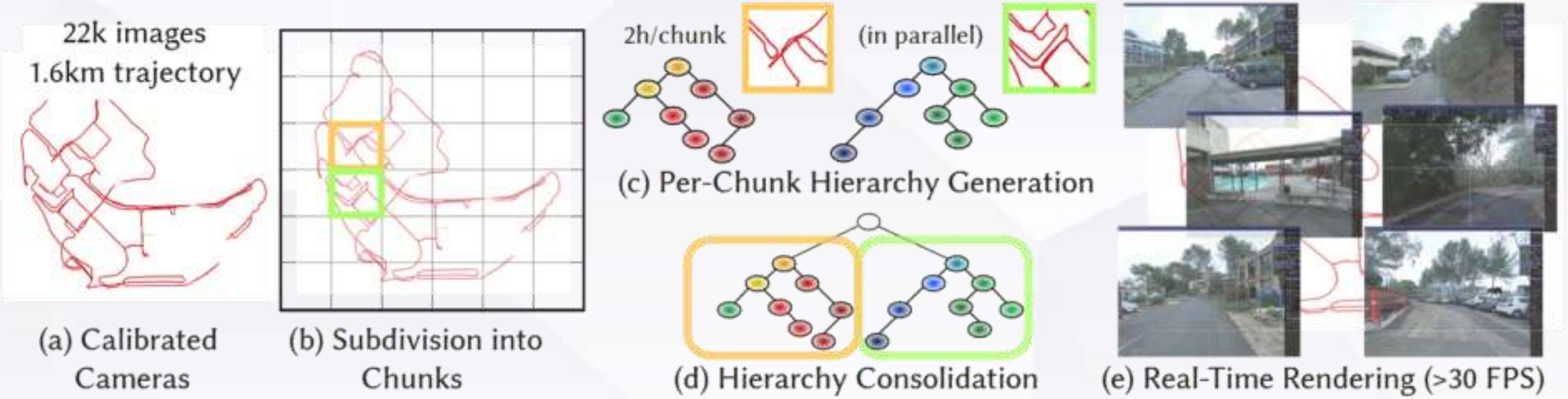
Training:

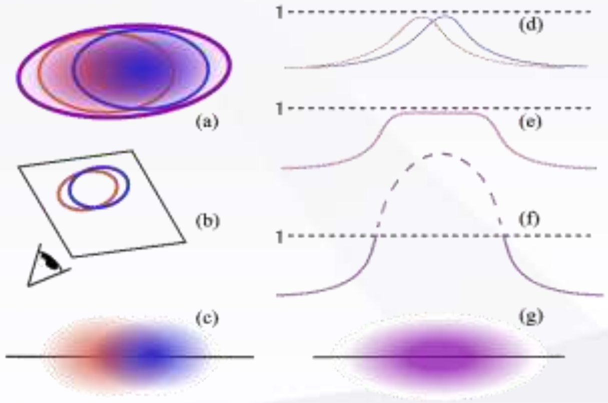


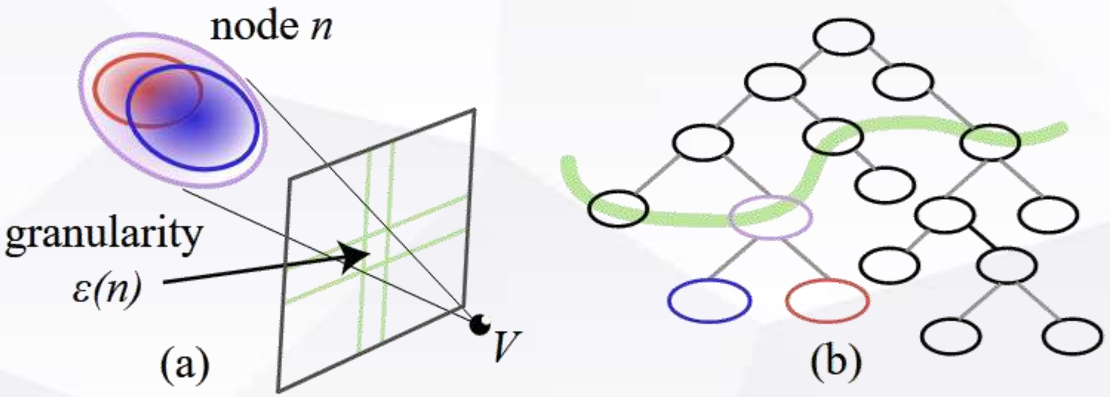
Rendering



Large-scale 3D Gaussian Splatting: Hierarchical 3DGS

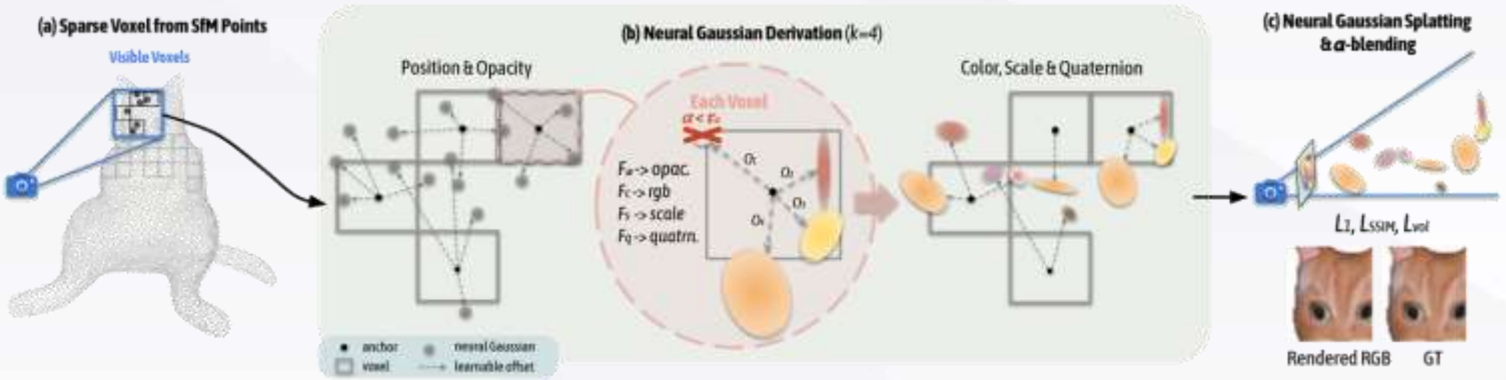




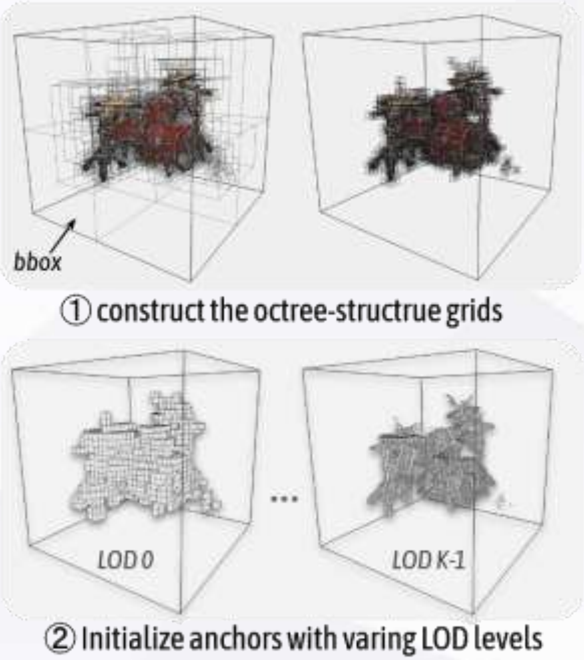


Large-scale 3D Gaussian Splatting: Octree-GS

- Anchor Point
  - MLP Feature
    - F_α, F_c, F_q, F_s



- Anchor-based Octree



References

1. Kerbl, Bernhard, et al. "3D Gaussian Splatting for Real-Time Radiance Field Rendering." ACM Trans. Graph. 42.4 (2023): 139-1.

2. Liu, Yang, et al. "Citygaussian: Real-time high-quality large-scale scene rendering with gaussians." European Conference on Computer Vision. Springer, Cham, 2025.

3. Kerbl, Bernhard, et al. "A hierarchical 3d gaussian representation for real-time rendering of very large datasets." ACM Transactions on Graphics (TOG) 43.4 (2024): 1-15.

4. Lu, Tao, et al. "Scaffold-gs: Structured 3d gaussians for view-adaptive rendering." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024.

5. Ren, Kerui, et al. "Octree-gs: Towards consistent real-time rendering with lod-structured 3d gaussians." arXiv preprint arXiv:2403.17898 (2024).

上海交通大学

SHANGHAI JIAO TONG UNIVERSITY

感谢聆听

饮 水 思 源　爱 国 荣 校