MesonGS: Post-training Compression of 3D Gaussians via Efficient Attribute Transformation

Shuzhao Xie, Weixiang Zhang, Chen Tang, Yunpeng Bai, Rongwei Lu, Shijia Ge, Zhi Wang

paper code

Background

- Novel View Synthesis: Generating images of a specific subject or scene from a specific point of view, when the only available information are pictures taken from different point of views.
- 3D Gaussian Splatting (3D-GS) shows excellent quality and speed on this task.

3D Gaussian Splats

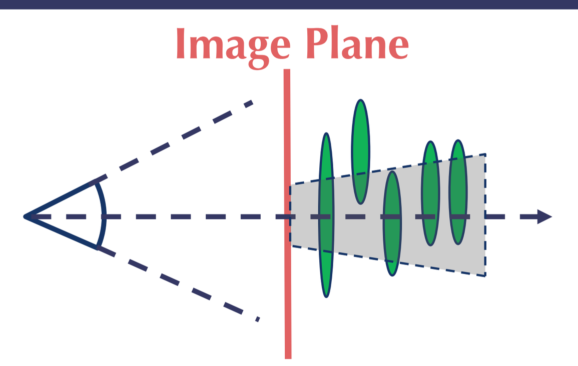
- 3D-GS reparameterizes the point in point clouds as a 3D Gaussian function:

$$G(\mathbf{X}) = \gamma_1 \exp\left(-\frac{1}{2}(\mathbf{X}-\mu)^\top \Sigma^{-1}(\mathbf{X}-\mu)\right)$$

- A 3D Gaussian splat contains:
- 3D center $\mu \in R^3$;
- Scale vector $s \in R^3$;
- Quaternion $q \in R^4$; } Covariance $\Sigma = \mathrm{RSS}^\top\mathrm{R}^\top$
- Spherical Harmonics $SH \in R^h$; → Color
- Opacity $o \in R$.

Volume Splatting

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



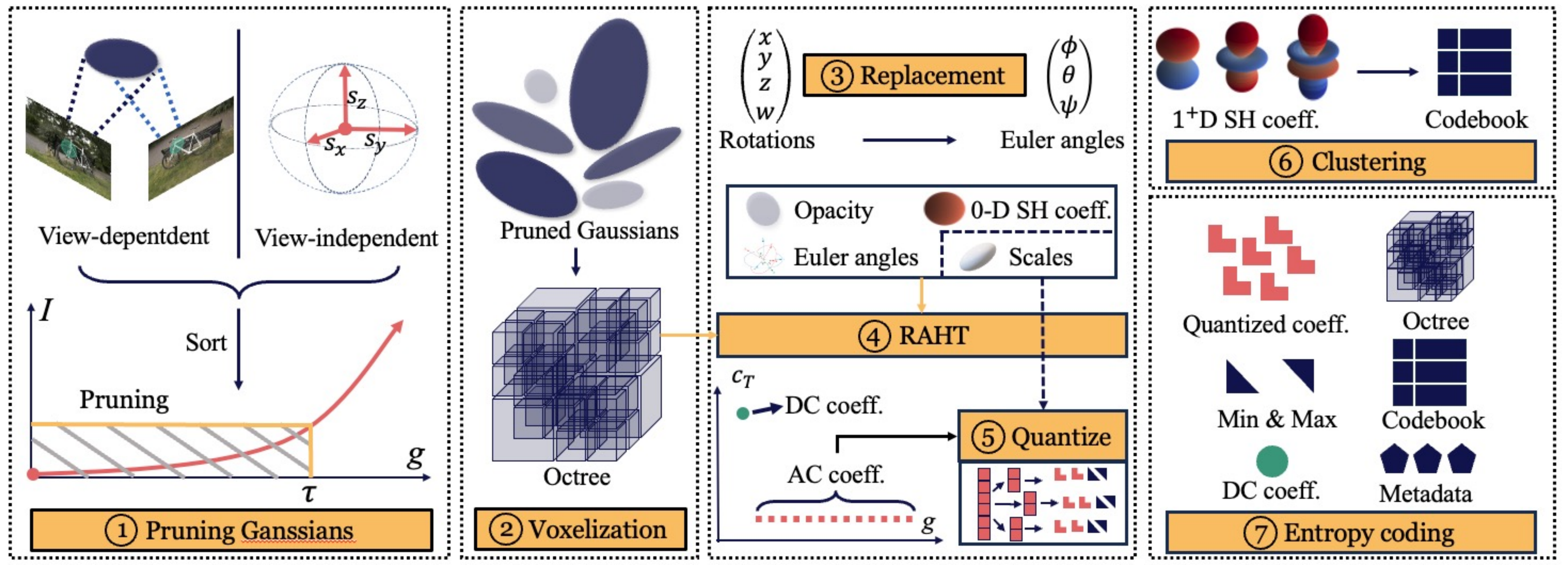
- Top Down View (N=5)

Motivation of Compressing 3D Gaussians

Method	Mip-NeRF	Instant-NGP	3DGS
Size (MB)	8.6	15-50	350-700
PNSR (dB)	24.3	22.1	23.6-25.2

- Sheer Volume: 5 million Gaussians are required to represent the bicycle scene in the Mip-NeRF 360 dataset, occupying 1.3 GB of storage.
- Complex Multi-channel Attributes: 1) scale; 2) quaternion; 3) Spherical harmonics; 4) opacity.

Overview of MesonGS

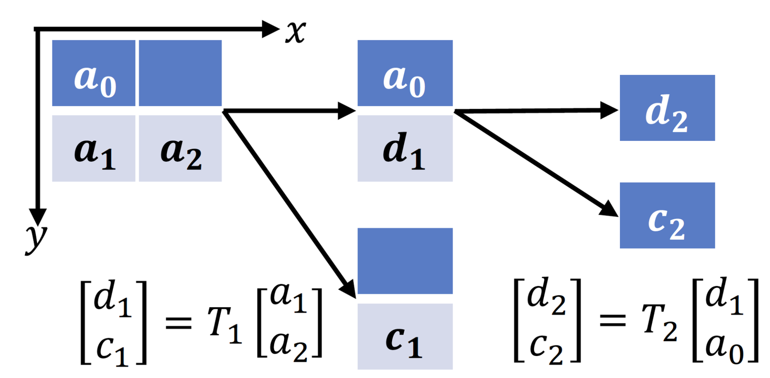


Pruning 3D Gaussians based on the Importance Score

- 30% of the 3D Gaussians contributes the 70% of the rendering result;
- Importance score of each Gaussian: $\mathrm{I}_g = V_{\mathrm{norm}}^{\beta} \cdot \sum_{p \in P} \alpha_i \prod_{j=1}^{i}(1-\alpha_j)$.

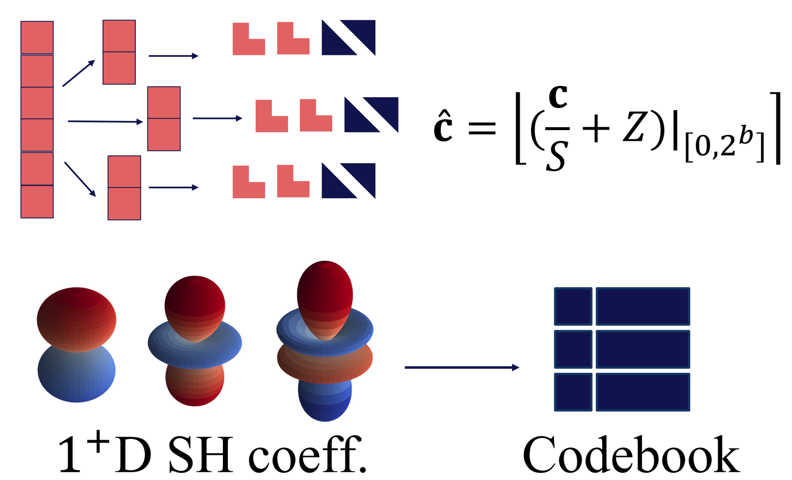
Attribute Transformation

- Transform Rotation to Euler angles;
- Voxelize 3D Center coordinates;
- Apply RAHT to Euler angles, Opacity, 0D-SH.



Quantization Strategies for Different Attributes

- RAHT produces DC and AC coefficients. For DC coefficients, we directly store them as 32-bit floating-point numbers; for AC coefficients, we perform block-wise quantization;
- For spherical harmonic coefficients in degrees greater than 0, we use a codebook for more compact compression.



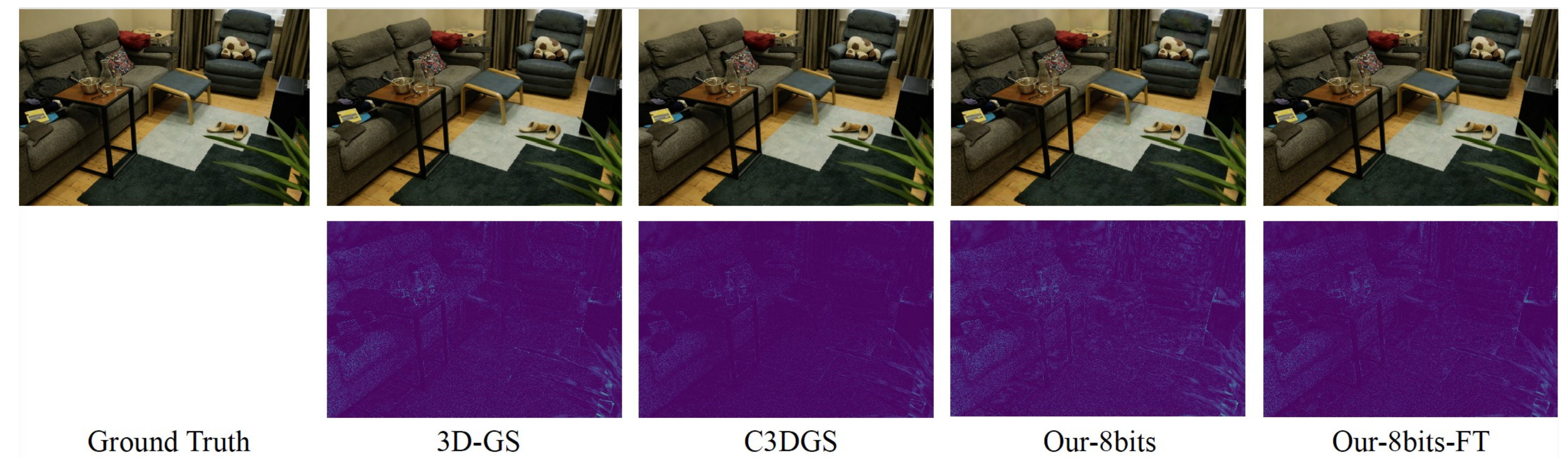
Finetune Strategy

- $\mu \in R^3$: Voxelization, fixed;
- $s \in R^3$: Quant. Simulation;
- $q \in R^4$: q -> Euler angle -> RAHT -> Quant. Simulation -> Euler angle -> Σ;
- $SH \in R^h$: 0D-SH -> RAHT + Quant. Simu. / 1^+D-SH -> Codebook;
- $o \in R$: Quant. Simulation;

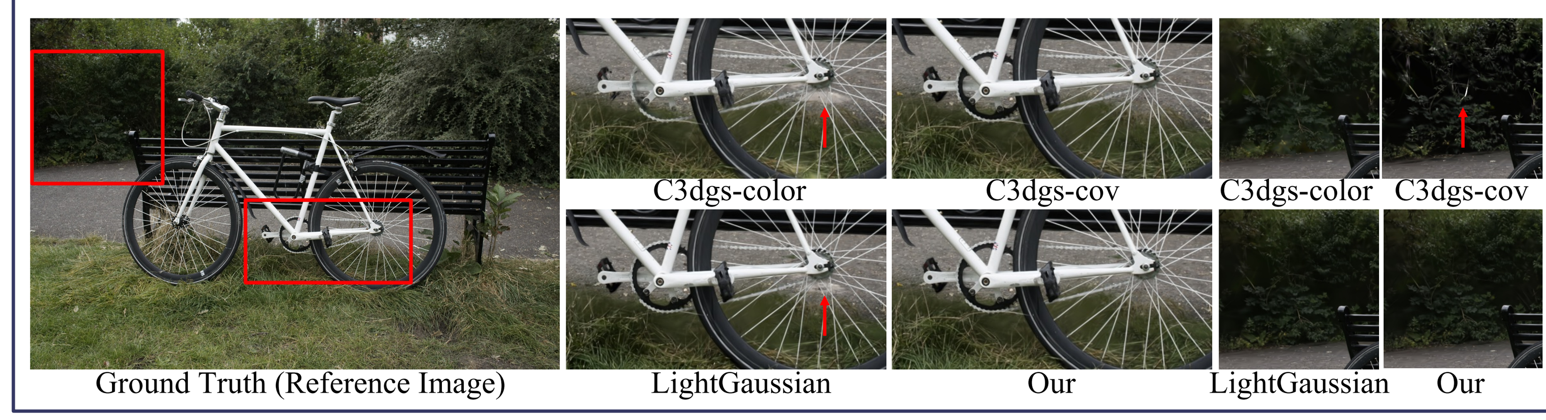
Results: Quantitative and Qualitative Comparison

Table 1: Quantitative comparison on Mip-NeRF 360, Tank&Temples, and Deep Blending. The best results overall are bolded in each metric, and the second-best results are underlined.

Method	Mip-NeRF 360				Tank&Temples				Deep Blending			
	PSNR	SSIM	LPIPS	Size (M)	PSNR	SSIM	LPIPS	Size (M)	PSNR	SSIM	LPIPS	Size (M)
3DGS [23]	28.98	0.865	0.193	641.70	23.36	0.838	0.187	421.90	29.56	0.898	0.250	703.77
C3DGS [33]	28.49	**0.858**	**0.205**	<u>27.82</u>	**23.32**	<u>0.832</u>	<u>0.194</u>	<u>17.28</u>	29.38	0.898	**0.238**	<u>25.30</u>
Lee *et al.* [25]	<u>28.60</u>	<u>0.856</u>	0.209	46.98	**23.32**	0.831	0.201	39.40	**29.79**	**0.901**	0.258	43.20
Our	27.70	0.838	0.224	**27.62**	<u>22.85</u>	0.822	0.208	**16.99**	29.08	0.895	0.260	**24.76**
Our-FT	**28.61**	<u>0.856</u>	<u>0.206</u>	**27.62**	**23.32**	**0.837**	**0.193**	**16.99**	<u>29.51</u>	**0.901**	<u>0.251</u>	**24.76**

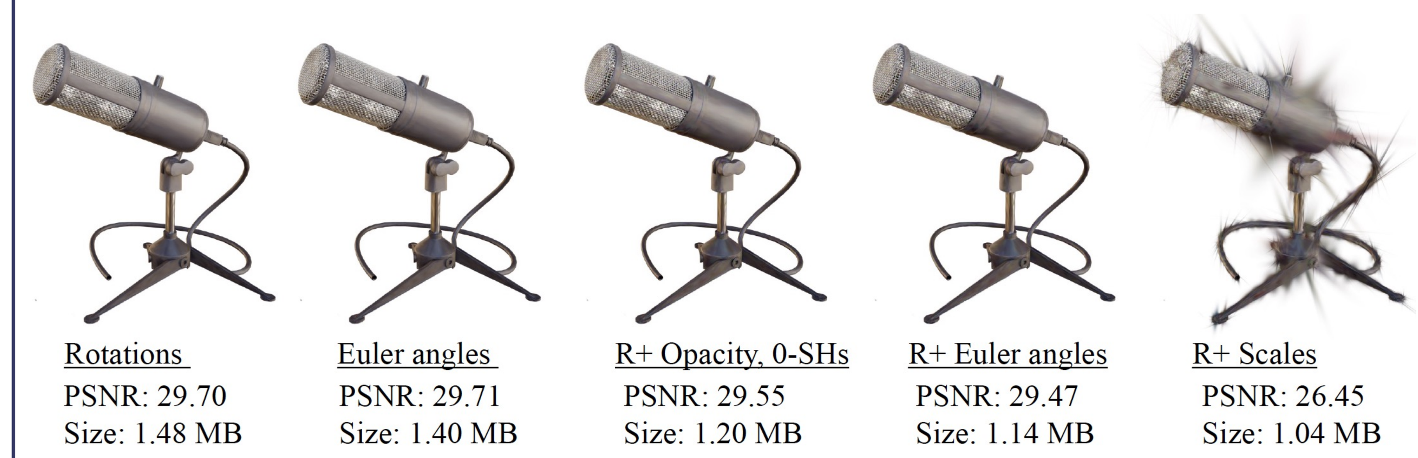


Performance of Pruning Strategy



- Our method is more universal.

Ablation Study: RAHT



- Applying RAHT to the scale vector can cause high quality degradation.

Ablation Study: Block-wise Quantization

Strategy	τ	Synthetic-NeRF				Mip-NeRF 360			
		PSNR(dB)	SSIM	LPIPS	Size(MB)	PSNR(dB)	SSIM	LPIPS	Size(MB)
Channel	66%	29.47	0.9476	0.0511	1.14	25.30	0.7533	0.3074	11.64
	50%	30.65	0.9529	0.0475	1.59	25.35	0.7461	0.3147	16.47
Block	66%	29.60	0.9494	0.0490	1.21	26.28	0.8035	0.2598	12.46
	50%	30.97	0.9560	0.0441	1.73	27.20	0.8238	0.2402	18.42