CompGS: Efficient 3D Scene Representation via Compressed Gaussian Splatting

Xiangrui Liu[1], Xinju Wu[1*], Pingping Zhang[1], Shiqi Wang[1†], Zhu Li[2], Sam Kwong[3]*

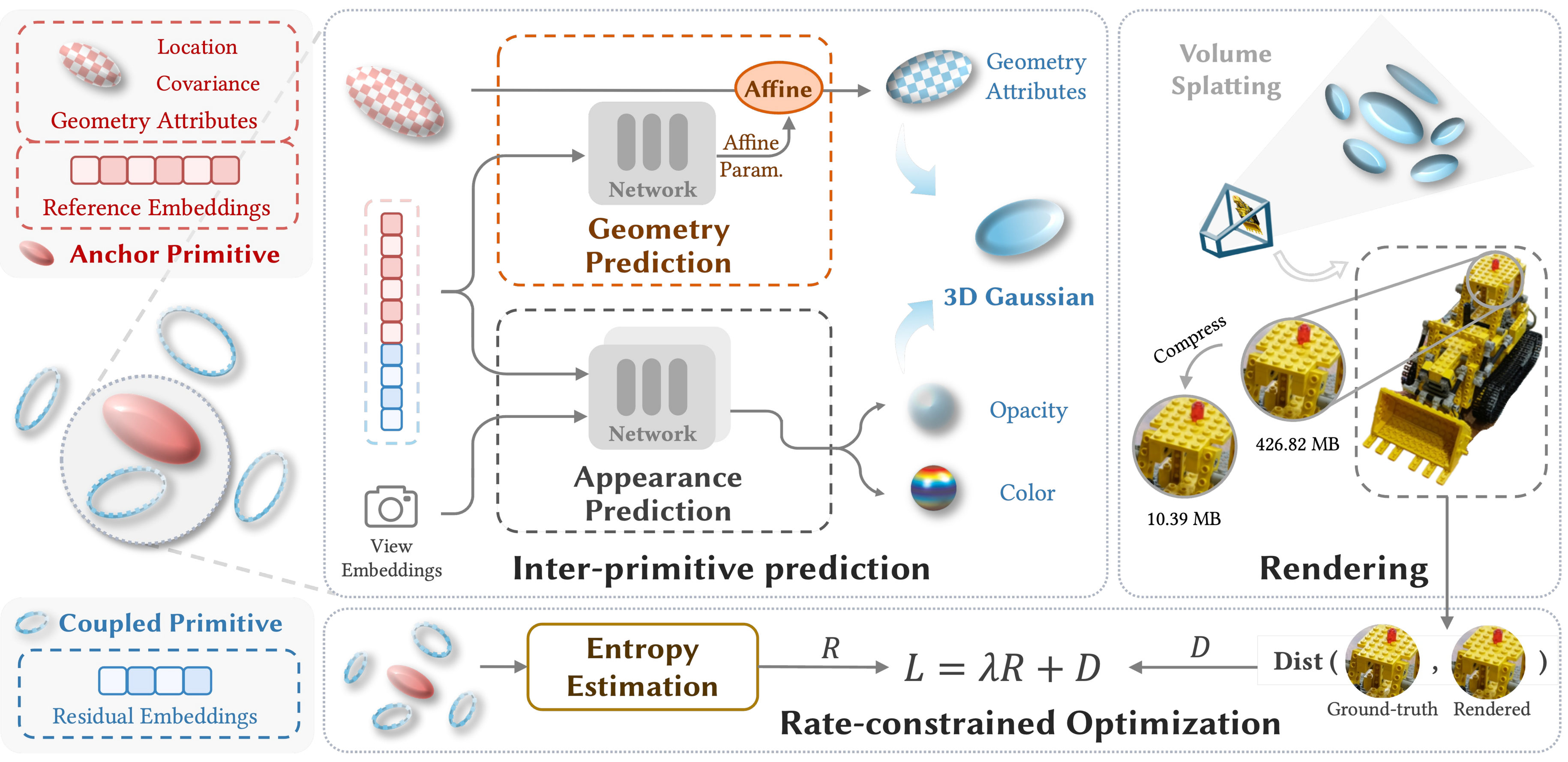
[1] City University of Hong Kong, [2] University of Missouri-Kansas City, [3] Lingnan University

ACM Multimedia 2024
Melbourne, Australia

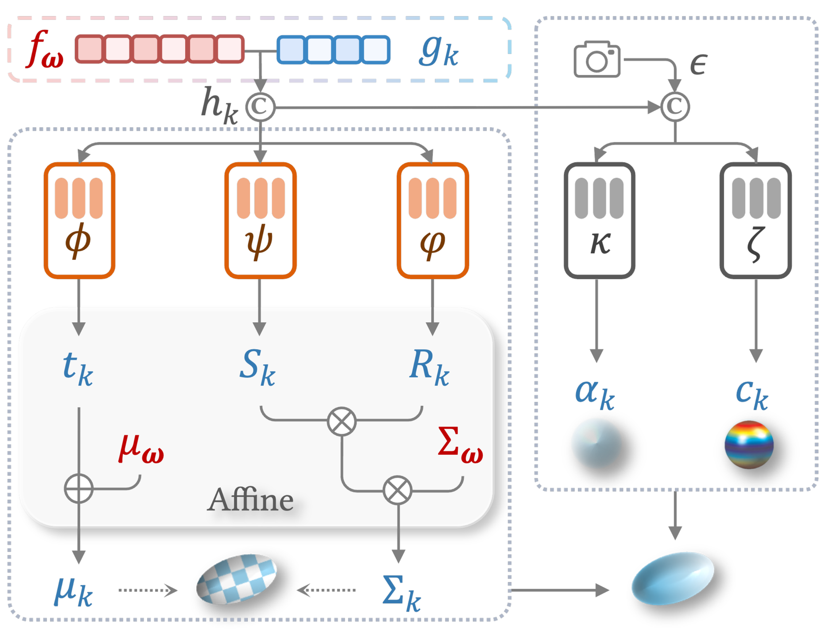
Code

Methodology

- We proposed a novel 3D scene representation method, Compressed Gaussian Splatting (**CompGS**), which utilizes compact primitives for efficient 3D scene representation with **remarkably reduced size**.



- We cultivate a **hybrid primitive structure** to facilitate compactness, wherein most primitives are adeptly predicted by a limited number of anchor primitives, thus allowing compact residual representations.



▸ Affine-based geometry prediction

location $\mu_k = \mu_\omega + t_k$ → k-th coupled

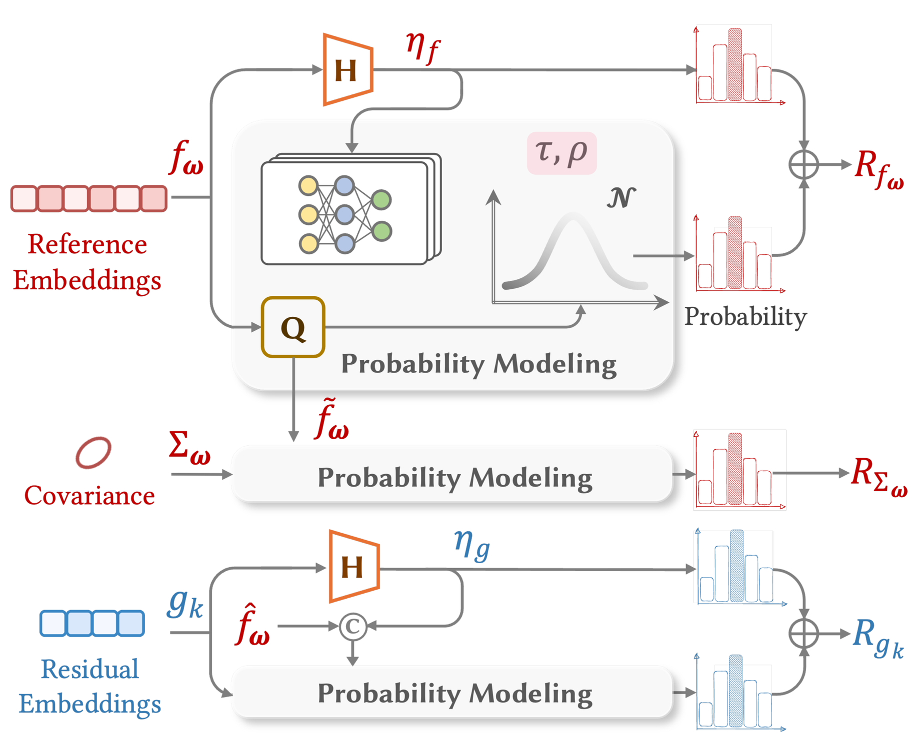
covariance $\Sigma_k = S_k R_k \Sigma_\omega$ → anchor

▸ View-dependent appearance prediction

opacity $\alpha_k = \kappa(\epsilon \oplus h_k)$

color $c_k = \zeta(\epsilon \oplus h_k)$

- We devise a **rate-constrained optimization scheme** to further prompt the compactness of primitives.



▸ Rate modeling based on entropy model

$p(\tilde{f}_\omega) = \mathcal{N}(\tau_f, \rho_f)$, with $\tau_f, \rho_f = \mathcal{E}_f(\eta_f)$ (hyperpriors)

$p(\tilde{\Sigma}_\omega) = \mathcal{N}(\tau_\Sigma, \rho_\Sigma)$, with $\tau_\Sigma, \rho_\Sigma = \mathcal{E}_\Sigma(\tilde{f}_\omega)$

$p(\tilde{g}_k) = \mathcal{N}(\tau_g, \rho_g)$, with $\tau_g, \rho_g = \mathcal{E}_g(\tilde{f}_\omega \oplus \eta_g)$

$R_{f_\omega} = \mathbb{E}_\omega[-\log p(\tilde{f}_\omega) - p(\eta_f)]$

...

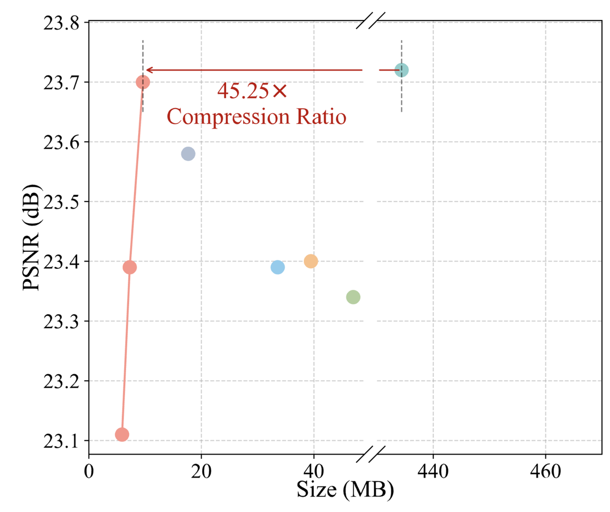
▸ Rate-constrained Optimization

$$\Omega^*, \Gamma^* = \arg\min_{\Omega, \Gamma} \lambda R + D$$

Performance

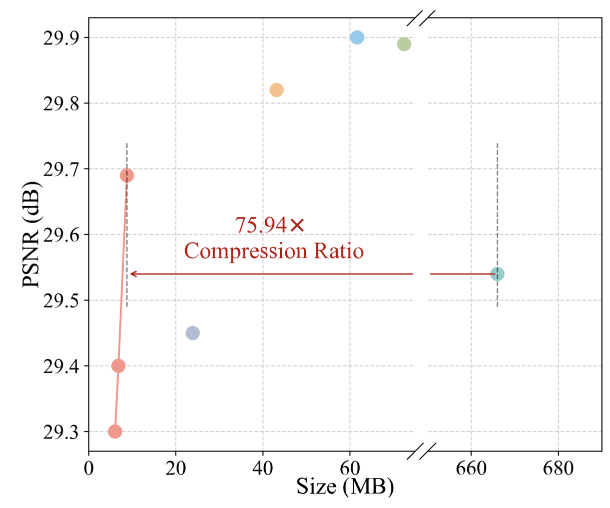
- Tanks & Temples

Methods	PSNR (dB)	SSIM	LPIPS	Size (MB)
Kerbl et al. [17]	23.72	0.85	0.18	434.38
Navaneet et al. [33]	23.34	0.84	0.19	47.01
Niedermayr et al. [34]	23.58	0.85	0.19	17.65
Lee et al. [20]	23.40	0.84	0.20	39.47
Girish et al. [10]	23.39	0.84	0.20	33.57
Proposed	23.70	0.84	0.21	9.60
	23.39	0.83	0.22	7.27
	23.11	0.81	0.24	5.89



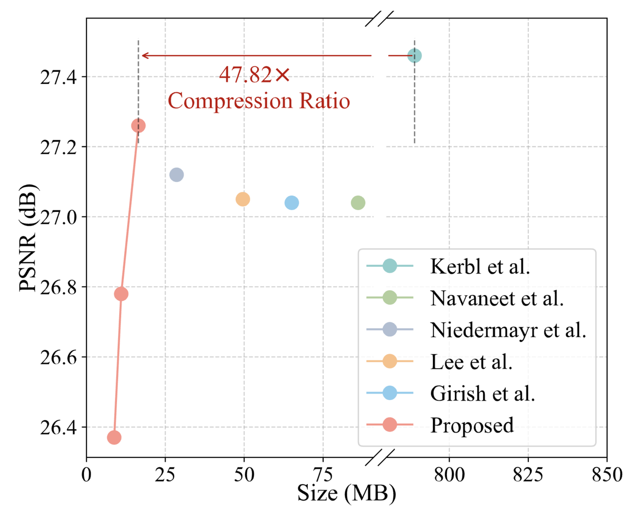
- Deep Blending

Methods	PSNR (dB)	SSIM	LPIPS	Size (MB)
Kerbl et al. [17]	29.54	0.91	0.24	665.99
Navaneet et al. [33]	29.89	0.91	0.25	72.46
Niedermayr et al. [34]	29.45	0.91	0.25	23.87
Lee et al. [20]	29.82	0.91	0.25	43.14
Girish et al. [10]	29.90	0.91	0.25	61.69
Proposed	29.69	0.90	0.28	8.77
	29.40	0.90	0.29	6.82
	29.30	0.90	0.29	6.03



- MipNeRF 360

Methods	PSNR (dB)	SSIM	LPIPS	Size (MB)
Kerbl et al. [17]	27.46	0.82	0.22	788.98
Navaneet et al. [33]	27.04	0.81	0.23	86.10
Niedermayr et al. [34]	27.12	0.80	0.23	28.61
Lee et al. [20]	27.05	0.80	0.24	49.60
Girish et al. [10]	27.04	0.80	0.24	65.09
Proposed	27.26	0.80	0.24	16.50
	26.78	0.79	0.26	11.02
	26.37	0.78	0.28	8.83



- **Ablation Study**
  - The proposed **hybrid primitive structure** can effectively eliminate the redundancies among primitives.
  - The proposed method can learn compact primitive representations through **rate-constrained optimization**.

Hybrid	Rate-constrained	Train				Truck			
Primitive	Optimization	PSNR (dB)	SSIM	LPIPS	Size (MB)	PSNR (dB)	SSIM	LPIPS	Size (MB)
×	×	22.02	0.81	0.21	257.44	25.41	0.88	0.15	611.31
✓	×	22.15	0.81	0.23	48.58	25.20	0.86	0.19	30.38
✓	✓	22.12	0.80	0.23	8.60	25.28	0.87	0.18	10.61

- Proportion of coupled primitives

K	PSNR (dB)	SSIM	LPIPS	Size (MB)
5	22.04	0.80	0.24	7.87
10	22.12	0.80	0.23	8.60
15	21.90	0.80	0.24	8.28

- Effectiveness of Residual embeddings

	PSNR (dB)	SSIM	LPIPS	Size (MB)
w.o. Res. Embed.	20.50	0.73	0.31	5.75
Proposed	21.49	0.78	0.26	5.51

Visualization