# A Compact Dynamic 3D Gaussian Representation for Real-Time Dynamic View Synthesis

Kai Katsumata<sup>®</sup>, Duc Minh Vo<sup>®</sup>, and Hideki Nakayama<sup>®</sup>

The University of Tokyo, Japan {katsumata,vmduc,nakayama}@nlab.ci.i.u-tokyo.ac.jp

## Supplementary Material

### **1** Additional Qualitative Results

We can divide training inputs into subsets to improve rendering quality. Figure 1 shows the rendering results trained on first 60 frames and all 300 frames. In exchange for the gains in rendering quality, the models trained on 60 frames require extra memory size for the whole scene. The memory size is around 1.9GB and it is still three times smaller than storing Guassians parameters per timestep. The result shows our method can control the reconstruction quality by dividing frames. Figures 2 and 3 show further novel view synthesis results of our method, demonstrating that our method can reconstruct a wide range of scenes.

#### 2 Additional Quantitative Results

We show the ablation study of the flow loss ( $\mathcal{L}_{\text{flow}}$ ) in Tab. 1. The flow loss improves visual quality of some scenes out of benchmark scenes. We also investigate the effects of the number of harmonics in Fourier approximation L in Tab. 2. The method is robust to the hyperparameter except when L is extremely small (L = 1, 2).

#### References

 Park, K., Sinha, U., Hedman, P., Barron, J.T., Bouaziz, S., Goldman, D.B., Martin-Brualla, R., Seitz, S.M.: HyperNeRF: a Higher-Dimensional Representation for Topologically Varying Neural Radiance Fields. ACM TOG 40(6), 1–12 (2021) 3, 4

**Table 1:** Effects of flow reconstruction. The flow reconstruction enhance the performance of some cases and the average scores.

	SEAR STEAK		CUT ROASTED BEEF		COOK SPINACH		COFFEE MARTINI		FLAME SALMON 1		FLAME STEAK		Mean	
	PSNR↑	SSIM↑	<b>PSNR</b> ↑	SSIM↑	<b>PSNR</b> ↑	$\text{SSIM} \uparrow$	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	<b>PSNR</b> ↑	SSIM↑
Ours w/o flow	29.4	0.969	32.5	0.970	31.2	0.964	25.5	0.916	25.8	0.909	28.4	0.965	28.8	0.948
Ours	33.1	0.976	32.7	0.969	31.8	0.966	26.1	0.922	26.0	0.926	33.1	0.974	30.4	0.955

2 K. Katsumata et al.



Fig. 1: Reducing training frames improves rendering quality while increasing memory size. The moemory size is still compact.

Table 2: Per scene results on different L. The results on L = 5 achieve better trade-offs between performance and complexity (model size).

	SEAR STEAK		CUT RO.	ASTED BEEF	COOK SPINACH		COFFEE MARTINI		FLAME SALMON 1		FLAME STEAK		Mean	
	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	<b>PSNR</b> ↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑
L = 1	25.9	0.944	30.0	0.952	30.2	0.953	25.8	0.917	25.4	0.922	27.5	0.944	27.5	0.938
L = 2	29.5	0.965	24.2	0.923	30.7	0.957	25.8	0.919	24.4	0.910	24.5	0.932	26.5	0.934
L = 3	32.3	0.974	31.9	0.965	31.4	0.962	26.1	0.923	25.8	0.924	30.1	0.967	29.6	0.952
L = 4	33.1	0.976	22.3	0.909	31.5	0.964	26.1	0.924	26.2	0.931	29.3	0.969	28.1	0.945
L = 5	33.1	0.976	32.7	0.969	31.8	0.966	26.1	0.922	26.0	0.926	33.1	0.974	30.4	0.955
L = 6	31.9	0.975	32.6	0.970	31.9	0.966	26.0	0.920	26.4	0.930	31.4	0.972	30.0	0.955
L = 7	33.3	0.977	31.1	0.967	32.0	0.967	26.2	0.923	26.4	0.929	30.4	0.969	29.9	0.955
L = 8	31.9	0.975	32.8	0.970	32.0	0.966	26.2	0.920	26.1	0.928	30.8	0.972	30.0	0.955



Fig. 2: Qualitative examples on HyperNeRF [1]. Visual quality of our method depends on the complexity of target scenes.



Fig. 3: Qualitative examples on HyperNeRF [1]. Our method reconstruct scenes faithfully and struggle to model complex motions.