LocalDyGS: Multi-view Global Dynamic Scene Modeling via Adaptive Local Implicit Feature Decoupling

Jiahao Wu, Rui Peng, Jianbo Jiao, Jiayu Yang, Luyang Tang, Kaiqiang Xiong, Jie Liang, Jinbo Yan, Runling Liu, Ronggang Wang*

https://wujh2001.github.io/LocalDyGS/

Peking University, Pengcheng Lab, University of Birmingham

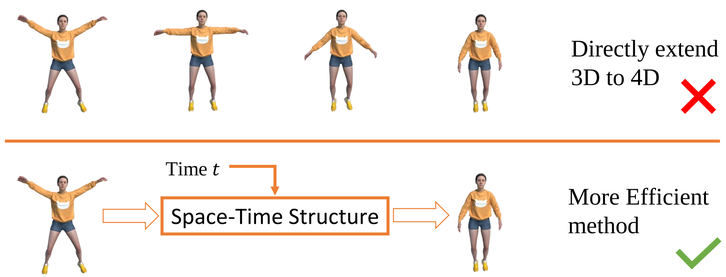
ICCV OCT 19-23, 2025 HONOLULU HAWAII

Background

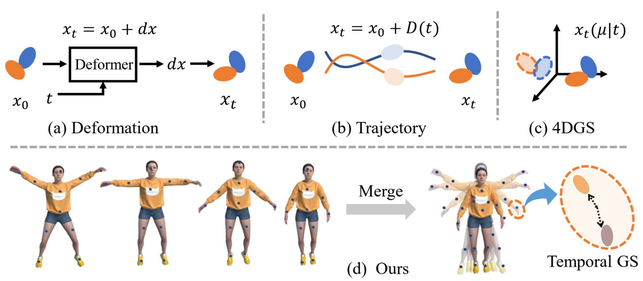
Since the real world is often in motion, achieving a more immersive visual experience requires addressing the critical problem of dynamic reconstruction. Previous approaches are mostly limited to small-motion scenarios and show significant shortcomings when applied to highly dynamic scenes such as basketball games. To tackle this challenge, we propose **LocalDyGS**.

Challenges of 3DGS for dynamic scene recon.

Unlike 3D static scene reconstruction, dynamic scene reconstruction requires the introduction of a temporal dimension, making it a 4D reconstruction. However, directly extending 3D to 4D results in significant computational overhead and storage costs. A more efficient approach is to fully leverage the temporal consistency within the scene and incorporate a compact 4D spatiotemporal structure to provide dynamic information, enabling faster and more compact 4D reconstruction.



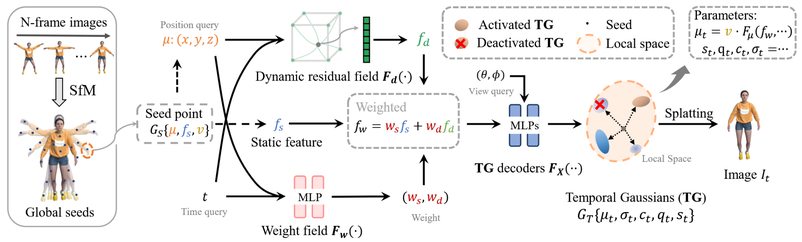
Dynamic scene NVS method compare.



(a) Deformation field based methods;
(b) Trajectory tracking-based solutions;
(c) 4DGS-Based methods.
(d) Our method

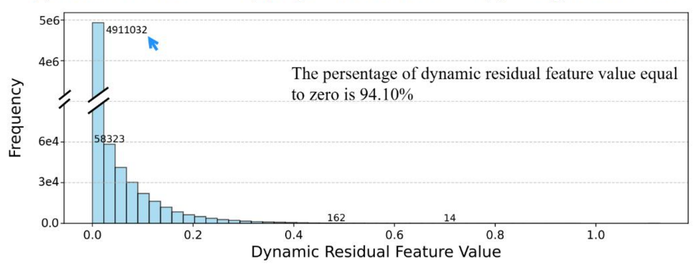
Method overview

In this paper we propose **LocalDyGS,** which consists of two parts to adapt our method to both large-scale and fine-scale motion scenes: 1) We decompose a complex dynamic scene into streamlined local spaces defined by seeds, enabling global modeling by capturing motion within each local space. 2) We decouple static and dynamic features for local space motion modeling. A static feature shared across time steps captures static information, while a dynamic residual field provides time-specific features. These are combined and decoded to generate Temporal Gaussians, modeling motion within each local space. As a result, we propose a novel dynamic scene reconstruction framework to **model highly dynamic real-world scenes more realistically**.



Dynamic and static feature decomposition

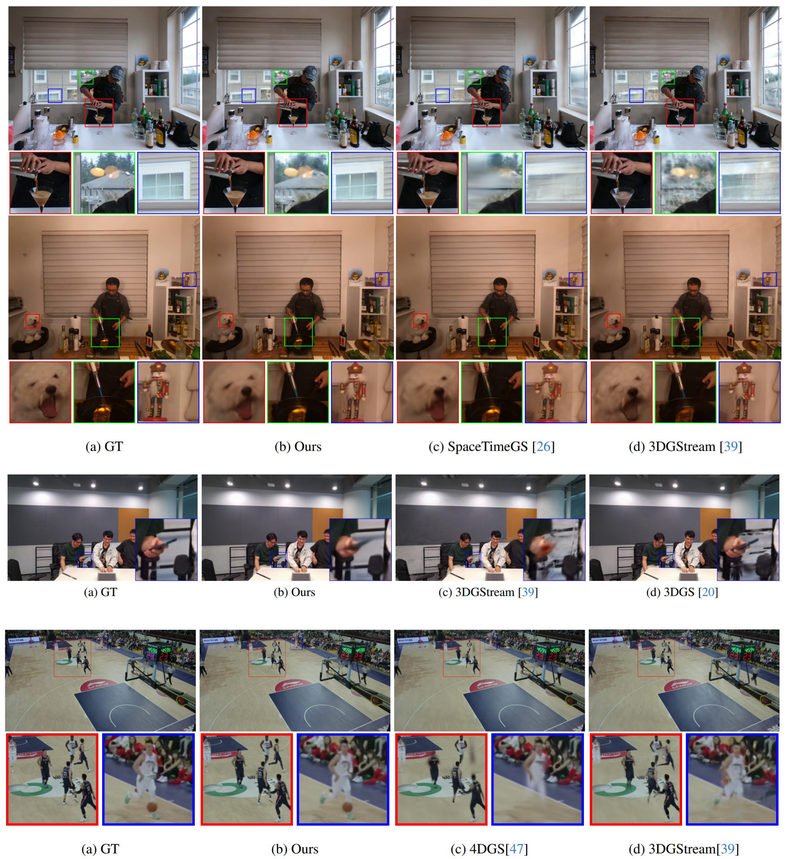
(a) Static feature (b) Dynamic feature (c) Weighted feature



(d) Distribution of dynamic residual feature value

We divide the features of dynamic scenes into static features and dynamic residual features. In this way, scene information is primarily represented by static features, while dynamic residual features capture only temporal residual details. This greatly reduces the difficulty of dynamic modeling and enables high-quality reconstruction of large-scale scenes.

Experiments



(a) GT (b) Ours (c) SpaceTimeGS [26] (d) 3DGStream [39]

(a) GT (b) Ours (c) 3DGStream [39] (d) 3DGS [20]

(a) GT (b) Ours (c) 4DGS[47] (d) 3DGStream[39]

Table 1. **Quantitative comparisons on the Neural 3D Video Dataset [25].** "Size" is the total model size for 300 frames. $DSSIM_1$ sets data range to 1.0 while $DSSIM_2$ to 2.0 [26]. * indicates online method.

Method	PSNR↑	$DSSIM_1$↓	$DSSIM_2$↓	LPIPS↓	FPS↑	Time↓	Size↓
StreamRF * [23]	28.26	-	-	-	10.9	-	5310 MB
NeRFPlayer [38]	30.69	0.034	-	0.111	0.05	6.0 h	5130 MB
HyperReel [1]	31.10	0.036	-	0.096	2	-	360 MB
K-Planes [14]	31.63	-	0.018	-	0.3	5.0 h	311 MB
HexPlane [9]	31.70	-	0.014	0.075	0.21	12.0 h	240 MB
MixVoxels [42]	31.73	-	0.015	0.064	4.6	-	500 MB
4DGaussian [47]	31.02	0.030	-	0.150	30	0.67 h	90 MB
3DGStream† [39]	31.67	-	-	-	215	1.0 h	1230 MB
RealTimeGS [55]	32.01	-	0.014	0.055	114	9.0 h	> 1000 MB
SpaceTimeGS [26]	32.05	0.026	0.014	0.044	140	> 5 h	200 MB
LocalDyGS(Ours)	32.28	0.028	0.014	0.043	105	0.58 h	100 MB

Table 3. **Quantitative comparison on VRU (GZ) basketball court dataset [40].** Static methods are tested on frame 0.

Method	PSNR ↑	SSIM ↑	LPIPS ↓
GOF [58]	30.39	0.949	0.141
2DGS [18]	30.78	0.949	0.187
3DGS [20]	30.50	0.949	0.171
4DGS [47]	28.32	0.930	0.186
SpaceTimeGS [26]	27.42	0.926	0.193
LocalDyGS(Ours)	**30.58**	**0.944**	**0.173**

In this paper, we first introduce a method that can adapt not only to fine-scale dynamic scenes but also to large-scale scenes. Extensive experiments show that LocalDyGS effectively adapts to dynamic scenes across various motion scales, performing well on both large-scale motion datasets, such as the basketball court, and fine-scale motion datasets, like N3DV and MeetRoom.