DENSE 3D GAUSSIAN SPLATTING INITIALIZATION FOR SPRASE IMAGE DATA

Simon Seibt[1], Thomas Chang[1], Bartosz von Rymon Lipinski[1], March Erich Latoschik[2]

[1] Game Tech Lab, Faculty of Computer Science, Nuremberg Institute of Technology, Nuremberg, Germany
[2] Human-Computer Interaction Group, Institute of Computer Science, University of Wuerzburg, Wuerzburg, Germany

EUROGRAPHICS CONFERENCE 2024

INTRODUCTION

- **3D Gaussian Splatting** (3DGS) [KKLD23] offers a computationally efficient alternative to Neural Radiance Field (NeRF) methods for novel-view synthesis.
- 3DGS uses a splatting-based scene representation to achieve accelerated photorealistic rendering.

- The effectiveness of 3DGS depends on the **quality** and **quantity** of **initial Structure-from-Motion (SfM) points**.
- This is a challenge particularly for scenarios with a **limited number of input images**, where sparse and inaccurate point clouds can lead to suboptimal training convergence and result in visual artifacts.

→ This work advances novel-view synthesis, enhancing 3DGS results with improved Gaussian initialization via a dense and accurate point cloud reconstructed **by "Dense Feature Matching for SfM"** (**DFM4SfM**).

RELATED WORK

To enable 3DGS for scenes with **few training images**:

- [ZFJW23] proposed a proximity-based **re-distribution** of 3D Gaussians supported by **monocular depth-information** to optimize Gaussian training.
- It reduces **overfitting** and **visual artifacts** for sparse image data by relocating initial Gaussians, but it is not able to accelerate training convergence.

DENSE FEATURE MATCHING FOR SfM (DFM4SfM)

Seibt et al. [SVRLCL23] introduced a novel approach to enhance conventional SfM with robust and accurate dense feature matching.

The pipeline is based on a **homographic decomposition** of the image space through **iterative rematching**, which improves **precision** and **density** of the point cloud reconstruction using…

a) iterative rematching of remaining features,
b) positional refinement of matching feature points in the target image,
c) extrapolation of additional feature matches (in critical image areas).

Further pipeline steps, specifically for SfM:

d) **Global Refinement:** Extension of positional refinement to a multi-view approach, enhancing pose estimation and 3D reconstruction accuracy.
e) **Global Extrapolation:** Considering multiple neighboring views to increase matching recall and reconstruct even denser 3D structures.
f) Utilizing of a precomputed "sparse connectivity graph" for (d) and (e).

PIPELINE OVERVIEW

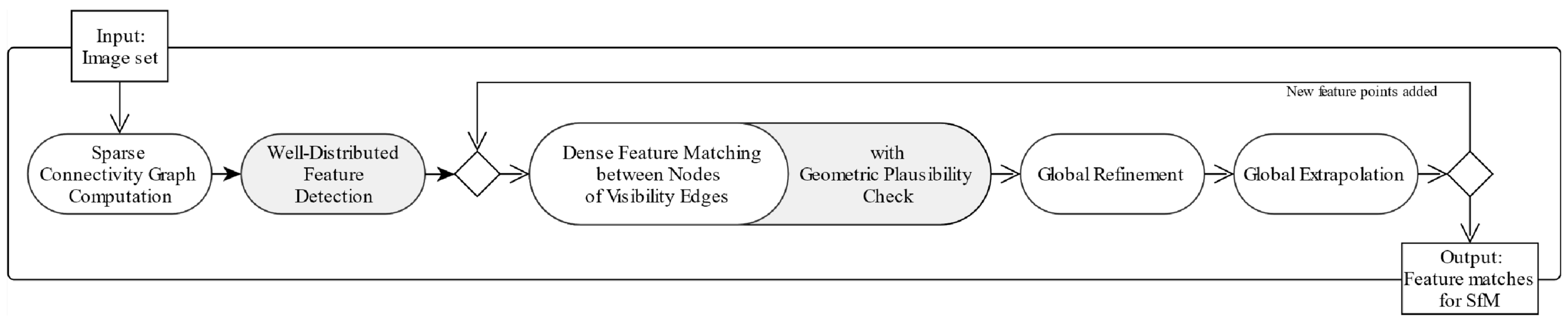

Fig. 1: UML-based activity diagram of enhanced DFM4SfM pipeline.

ENHANCING DFM4SfM FOR 3DGS

Main contributions of this work are enhancements for DFM4SfM to improve 3DGS rendering, especially for low-image scenes:

a) Grid-based feature detection for **well-distributed** point clouds, capturing both foreground and background details using a coarse-to-fine detection approach per grid cell.

→ Assures a more uniform and denser splat initialization for faster convergence by also considering image areas with visually less significant features.

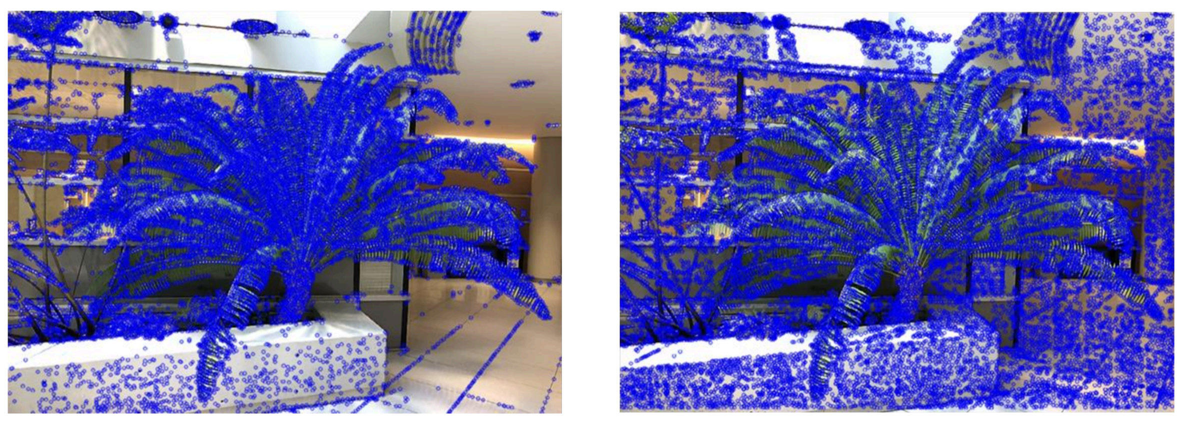

Fig. 2: Comparison of traditional feature detection (left) and proposed method (right).

b) Estimation of geometrically plausible feature matches (p, p_2') to supplement a multi-homography decomposition strategy via adjacent **sealed matches** (m, m') from previous rematching iterations.

→ Enhances homography estimation for **wide-baseline** image pairs with **complex visual structures** or **multiple depths.**

→ Minimizes **sparse** and **potentially incorrect** 3D Gaussian initializations of traditional SfM approaches, typically caused by fundamental matrix degeneration.

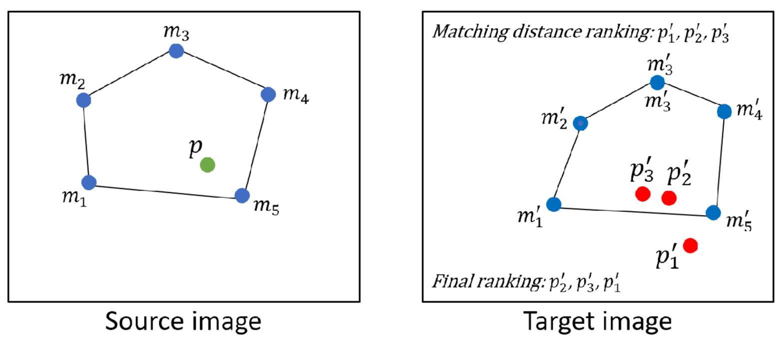

Fig. 3: Geometric plausibility check with adjacent sealed matches and adjusted candidate ranking.

→ Enhanced DFM4SfM results in a more robust multi-plane recovery in 3D space, generating a denser and more precise point cloud for 3DGS initialization, accelerating training convergence and improving rendering results.

RESULTS

- Benchmarks on Intel i9 14900KF CPU, 64GB RAM, NVIDIA RTX 4060 GPU.
- NeRF-LLFF datasets [MSOC*19] with ~10,000 initial keypoints per image.
- **COLMAP**'s default 3D reconstruction and **fine-tuned feature matching** compared to **DFM4SfM's expanded matching**.
- DFM4SfM reconstruction contains 213% more 3D points than default COLMAP while tripling processing time.

Metrics	NeRF-LLFF (8 Scenes)				
	3DGS	Ours		DFM4SfM w/o Improvements	
SSIM ↑	0.77	**0.86**	+11.7%	0.83	+7.8%
PSNR ↑	23.14	**26.62**	+15.0%	25.94	+12.1%
LPIPS ↓	0.20	**0.13**	-35.0%	0.16	-20.0%
Time ↓	32:05	**27:44**	-13.6%	28:33	-11.0%

Tab. 1: Results for 3DGS and 3DGS initialized with improved DFM4SfM (Ours) and w/o improvements on NeRF-LLFF for 30k training iterations.

	N	Iters	SSIM ↑		PSNR ↑		LPIPS ↓		Time ↓	
			3DGS	Ours	3DGS	Ours	3DGS	Ours	3DGS	Ours
Fern (NeRF-LLFF)	16	30000	0.68	**0.83**	21.16	**24.40**	0.25	**0.16**	36:58	**34:47**
	12	30000	0.54	**0.75**	18.13	**22.60**	0.32	**0.21**	27:16	**21:55**
	9	30000	0.52	**0.69**	17.58	**20.33**	0.33	**0.25**	25:07	**20:35**
	6	30000	0.48	**0.65**	16.60	**18.86**	0.38	**0.27**	21:23	**19:19**
	3	30000	0.33	**0.47**	13.38	**14.75**	0.52	**0.44**	19:37	**18:34**
	16	15000	0.69	**0.82**	21.34	**24.17**	0.24	**0.16**	15:54	**15:43**
	16	5000	0.72	**0.84**	22.07	**24.50**	0.24	**0.17**	**03:43**	04:15
	16	1000	0.57	**0.76**	19.59	**23.66**	0.51	**0.26**	**00:37**	00:45
	16	500	0.50	**0.66**	17.68	**21.68**	0.61	**0.41**	**00:19**	00:25
	16	100	0.49	**0.53**	15.91	**19.01**	0.59	**0.57**	**00:04**	00:06

Tab. 2: Results with varying number of images (N) and 3DGS training iterations (Iters) for the scene Fern.
Underlined: Proposed method (Ours) surpassing 3DGS's best result with lower N or Iters.

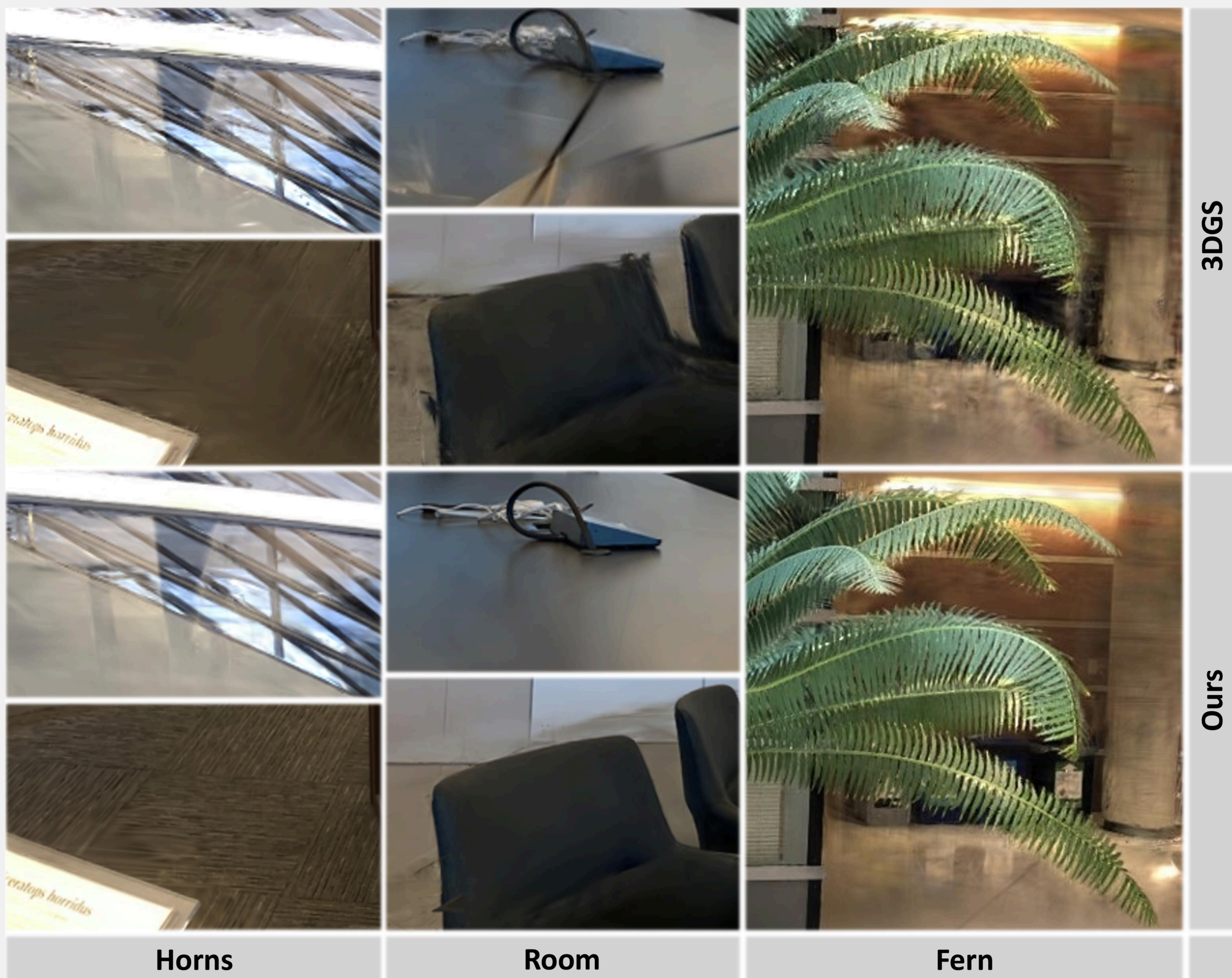

Fig. 4: Visual comparisons on sparse image data from NeRF-LLFF: 3DGS without (top) and with the proposed method (bottom).

Project page

Acknowledgement
This work is funded by the Federal Ministry of Education and Research (BMBF Germany, No. 01IS23007B).

REFERENCES

[KKLD23] Kerbl B., Kopanas G., Leimkuehler T., Drettakis G.: 3d gaussian splatting for real-time radiance field rendering. ACM Transactions on Graphics. (2023).

[MSOC*19] Mildenhall B., Srinivasan P. P., Ortiz-Cayon R., Kalantari N. K., Ramamoorthi R., Ng R., Kar A.: Local light field fusion: practical view synthesis with prescriptive lines. ACM Transactions on Graphics. (2019).

[SVRLCL23] Seibt S., Von Rymon Lipinski B., Chang T., Latoschik M. E.: Dfm4sfm - dense feature matching for structure from motion. In IEEE International Conference on Image Processing Workshops. (2023).

[ZFJW23] Zhu Z., Fan Z., Jiang Y., Wang Z.: Fsgs: Real-time few-shot view synthesis using gaussian splatting, 2023.