Supplemental Material: Splat and Replace: 3D Reconstruction with Repetitive Elements

NICOLÁS VIOLANTE, Inria & Université Côte d'Azur, France and Adobe, USA
ANDREAS MEULEMAN, Inria & Université Côte d'Azur, France
ALBAN GAUTHIER, Inria & Université Côte d'Azur, France
FREDO DURAND, MIT, USA
THIBAULT GROUEIX, Adobe, USA
GEORGE DRETTAKIS, Inria & Université Côte d'Azur, France

ACM Reference Format:
Nicolás Violante, Andreas Meuleman, Alban Gauthier, Fredo Durand, Thibault Groueix, and George Drettakis. 2025. Supplemental Material: Splat and Replace: 3D Reconstruction with Repetitive Elements. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers (SIGGRAPH Conference Papers '25), August 10–14, 2025, Vancouver, BC, Canada.* ACM, New York, NY, USA, 1 page. https://doi.org/10.1145/3721238.3730727

1 ADDITIONAL DETAILS

In this supplemental document we present additional details on the method and its implementation.

We train 3DGS for 30K iterations using depth regularization and exposure adjustment [Kerbl et al. 2024], and also the opacity and scaling regularization. During this phase we also optimize the exposure adjustment for the test views. Following 3DGS Monte Carlo [Kheradmand et al. 2024], we disable the opacity reset.

For contrastive learning, we add a feature vector of dimension $d = 16$ to each Gaussian, and train them using the contrastive loss for 5K iterations while keeping the other attributes frozen. At each iteration, for an image with M masks (including the background mask), we sample $M_u = 4096$ pixels uniformly in the image so larger masks get more samples. We also sample $\min\{3, \frac{M_s}{M}\}$ pixels in each mask, for a total of $M_s = 4096$ pixels in the image, to ensure small masks get enough pixels to form positive and negative pairs. We show an example of masks and learned features in Fig. 1.

For the 2D matching, we use the recently introduced MASt3R [Leroy et al. 2024] due to its remarkable capabilities to handle strong viewpoint changes. After 3D segmentation, we render the segmented instances and apply MASt3R to the pair of images being considered for matching. We lift the 2D matches to 3D using the expected termination depth of the segmented instance. The expected termination for a ray $\mathbf{r}$ is defined as $\mathbf{x}_e = \mathbf{r}\left(\sum_{i=1}^{N} T_i \alpha_i\right)$, where α_i is the blending weight and T_i the transmittance [Kerbl et al. 2023]. Since the estimated depth is not fully reliable at edges, we apply a 5x5 dilation kernel and discard outliers with a median absolute deviation test based on z-scores with threshold 3.5 [Iglewicz and Hoaglin 1993]

ACM ISBN 979-8-4007-1540-2/2025/08.
https://doi.org/10.1145/3721238.3730727

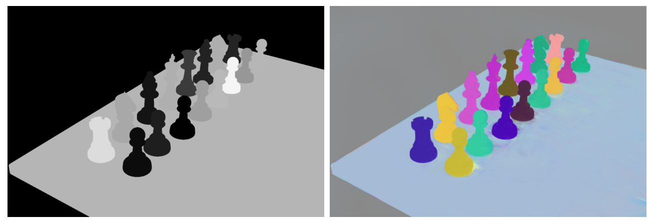

Fig. 1. **Contrastive features visualization.** Example masks (*left*) used to learn contrastive features, alongside the first three components of the resulting contrastive features (*right*).

We optimize our shared representation for a total of 7k iterations, gradually adding spherical harmonic bands every 1k iteration. Densification and pruning are enabled after 500 iterations until 5000 iterations. During finetuning, we also improve the rigid transformation by optimizing a 9D representation of the rotation [Zhou et al. 2019] and the translation vector using Adam [Kingma and Ba 2015] with learning rate of 10^{-3} annealed to 10^{-4}. More precisely, in the 9D representation, we predict the three columns of the rotation matrix and then apply Gram-Schmidt orthonormalization to ensure a proper rotation matrix.

REFERENCES

Boris Iglewicz and David C. Hoaglin. 1993. *How to detect and handle outliers*. Number 16 in The ASQC basic references in quality control: statistical techniques. ASQC Quality Press, Milwaukee, Wis.

Bernhard Kerbl, Georgios Kopanas, Thomas Leimkühler, and George Drettakis. 2023. 3D Gaussian Splatting for Real-Time Radiance Field Rendering. *ACM Transactions on Graphics* 42, 4. https://repo-sam.inria.fr/fungraph/3d-gaussian-splatting/

Bernhard Kerbl, Andreas Meuleman, Georgios Kopanas, Michael Wimmer, Alexandre Lanvin, and George Drettakis. 2024. A Hierarchical 3D Gaussian Representation for Real-Time Rendering of Very Large Datasets. In *ACM Transactions on Graphics*, Vol. 43. https://repo-sam.inria.fr/fungraph/hierarchical-3d-gaussians/

Shakiba Kheradmand, Daniel Rebain, Gopal Sharma, Weiwei Sun, Jeff Tseng, Hossam Isack, Abhishek Kar, Andrea Tagliasacchi, and Kwang Moo Yi. 2024. 3D Gaussian Splatting as Markov Chain Monte Carlo. In *NeurIPS*.

Diederik P. Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization.. In *ICLR (Poster)*, Yoshua Bengio and Yann LeCun (Eds.). http://dblp.uni-trier.de/db/conf/iclr/iclr2015.html#KingmaB14

Vincent Leroy, Yohann Cabon, and Jérôme Revaud. 2024. Grounding Image Matching in 3D with MASt3R. http://arxiv.org/abs/2406.09756 arXiv:2406.09756 [cs].

Yi Zhou, Connelly Barnes, Lu Jingwan, Yang Jimei, and Li Hao. 2019. On the Continuity of Rotation Representations in Neural Networks. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.