Exploring Fast Refinement of Monocular Depth Estimates with Sparse Point Clouds for Novel View Synthesis

(Masters-level internship)

Supervised by Laura Fink and George Drettakis
GRAPHDECO, Inria Sophia Antipolis, France
laura.fink@inria.fr george.drettakis@inria.fr

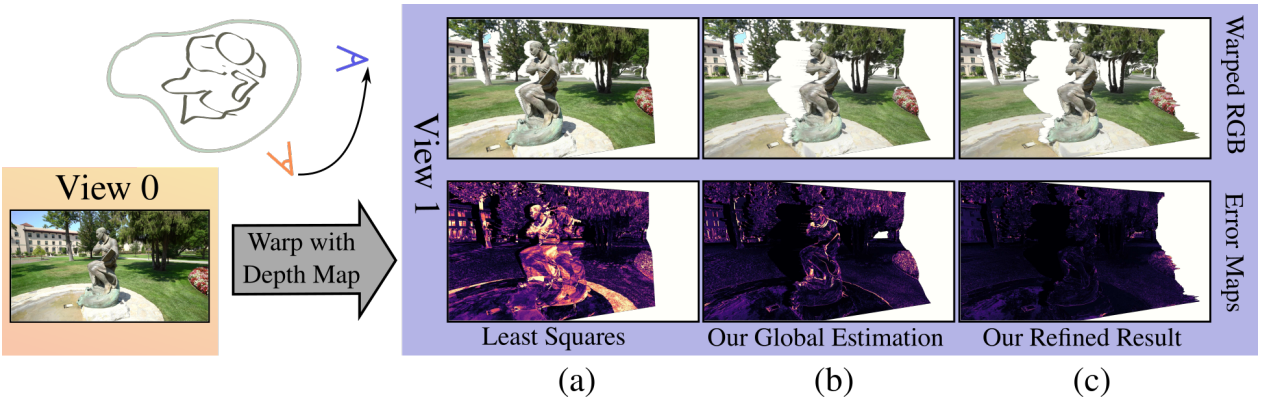


Figure 1: *Single-image depth estimation is underconstrained: local distortions and misplaced depth boundaries can produce incorrect geometry and visible rendering artifacts [7].*

Motivation

Many tasks use depth maps from monocular estimators. Novel view synthesis (NVS) methods, like Gaussian splatting, increasingly incorporate monocular depth predictions to initialize scene geometry and accelerate reconstruction [1]. Many image editing effects, like depth of field or motion blur, rely on such depth maps. Despite recent progress, single-image depth estimation is underconstrained: local distortions and misplaced depth boundaries can produce incorrect geometry and visible rendering artifacts (see figure). Meanwhile, sparse geometric observations are often available from structure-from-motion, SLAM, or LiDAR. Combining these observations with dense monocular predictions offers an opportunity to improve initialization while retaining the efficiency needed for downstream tasks.

Research question

Can a single estimator-agnostic refiner improve monocular depth using just a sparse point set for downstream tasks, without estimator-specific retraining or additional multi-view consistency cues?

Approach

Recent work demonstrates complementary strategies, using sparse point clouds for estimator-specific conditioning [2, 3, 4]. Likewise, SLAM systems and multi-view consistency based refinement provide useful ideas for compact correction and propagation from sparse anchors [5, 6, 7]. Building on these directions, the internship will investigate a lightweight, edge-aware refinement module that accepts an RGB image, its monocular depth estimate, and registered sparse depth observations with optional confidence values. A core requirement is plug-and-play compatibility with different monocular estimators: the refiner will consume their output maps without modifying their weights or requiring estimator-specific fine-tuning of the refiner. This will enable the direct integration of future released monocular foundation models into existing task pipelines, without the need

of fine-tuning for the user. Robust alignment, compact correction fields, and structure-aware propagation will be explored. Particular attention should be paid to avoiding foreground–background leakage, preserving thin structures, and handling noisy, clustered, or missing observations.

Work environment and requirements

The internship will take place at Inria Sophia-Antipolis, Université Côte d'Azur, in the GraphDeco group.

We are looking for a Master's or final-year engineering student with a strong interest in computer graphics, computer vision and machine learning. Experience with Python/PyTorch and optimization are desirable; proficiency in agentic coding is a plus.

References

[1] A. Meuleman, I. Shah, A. Lanvin, B. Kerbl, and G. Drettakis. *On-the-fly Reconstruction for Large-Scale Novel View Synthesis from Unposed Images*. ACM Transactions on Graphics, 2025. Paper.

[2] Z. Wang, S. Chen, L. Yang, J. Wang, Z. Zhang, H. Zhao, and Z. Zhao. *Depth Anything with Any Prior (Prior Depth Anything)*. 2025 preprint; ICLR 2026. Paper.

[3] Y. Zuo, W. Yang, Z. Ma, and J. Deng. *OMNI-DC: Highly Robust Depth Completion with Multiresolution Depth Integration*. ICCV, 2025. Paper.

[4] H. Lin et al. *Prompting Depth Anything for 4K Resolution Accurate Metric Depth Estimation*. CVPR, 2025. Paper.

[5] W. Zhang et al. *HI-SLAM2: Geometry-Aware Gaussian SLAM for Fast Monocular Scene Reconstruction*. IEEE Transactions on Robotics, 2025. Paper.

[6] E. Dexheimer and A. J. Davison. *COMO: Compact Mapping and Odometry*. ECCV, 2024. Paper.

[7] L. Fink, L. Franke, B. Egger, J. Keinert, and M. Stamminger. *Refinement of Monocular Depth Maps via Multi-View Differentiable Rendering*. Vision, Modeling, and Visualization (VMV), 2025. DOI / Paper.