SuPerMVO : SuperPoint-Glued Pose Estimation for Monocular Visual Odometry

Aaron Sequeira, Aryaman Rao, Kunal Siddhawar, and Kush Patel

University of Michigan – Ann Arbor

Motivation

We aim to enhance monocular visual odometry by using transformer-based key-point matching like SuperPoint and SuperGlue, which provide robust descriptors even in low-texture scenes where traditional methods like ORB fail.
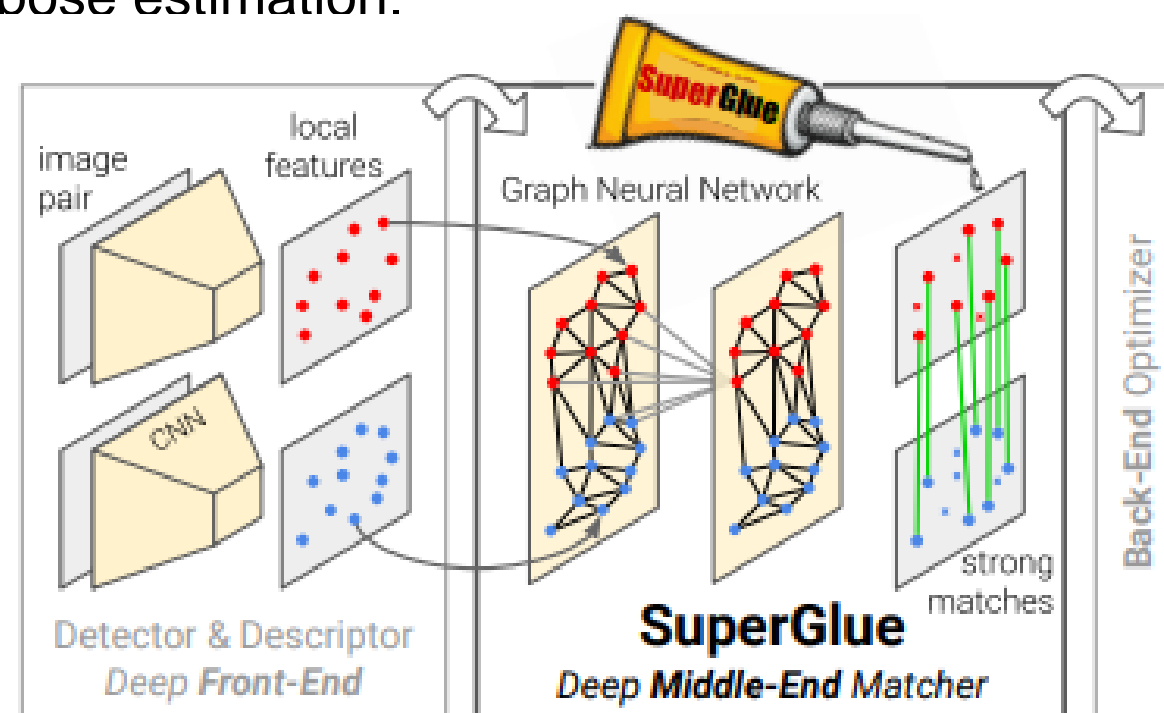
SuperPoint

- Detects key-points using a shared encoder and interest point decoder for real-time performance.
- Outputs 2D key-points and 256D descriptors using a fully convolutional self-supervised network.
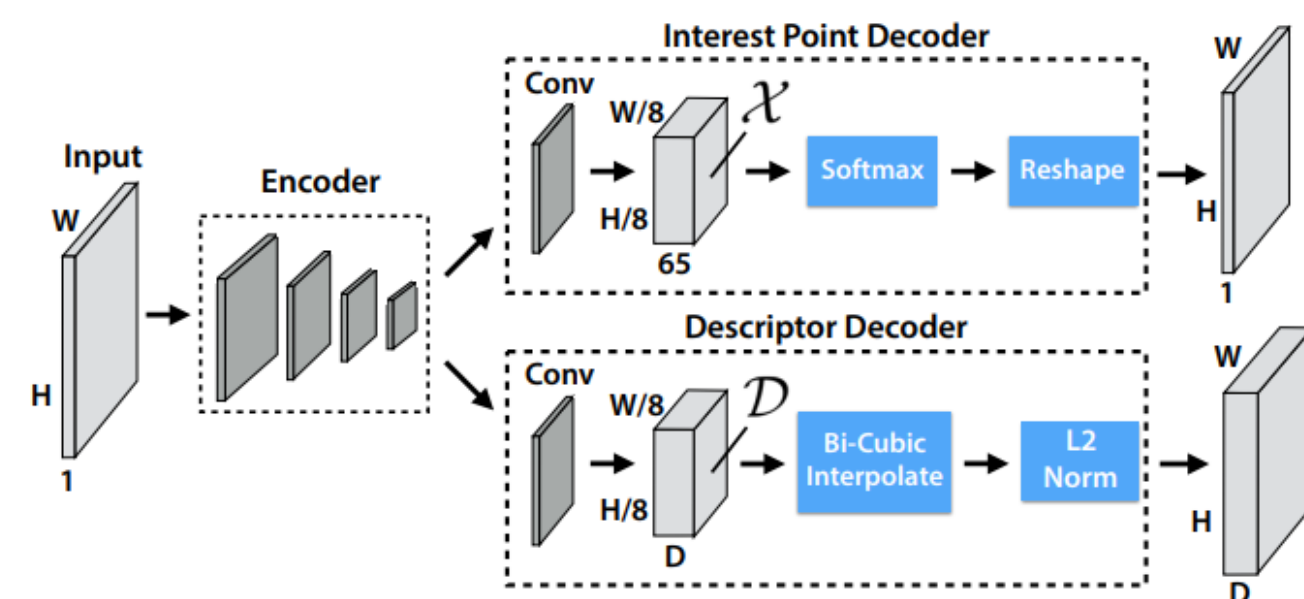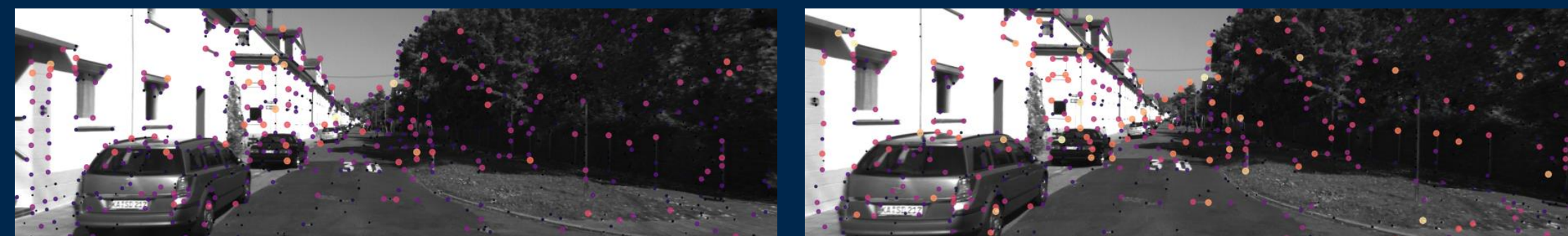

SuperGlue

- Matches key-points across image pairs using self- and cross-attention in a graph neural network.
- Predicts context-aware correspondences and match confidence scores, enabling more robust pose estimation.
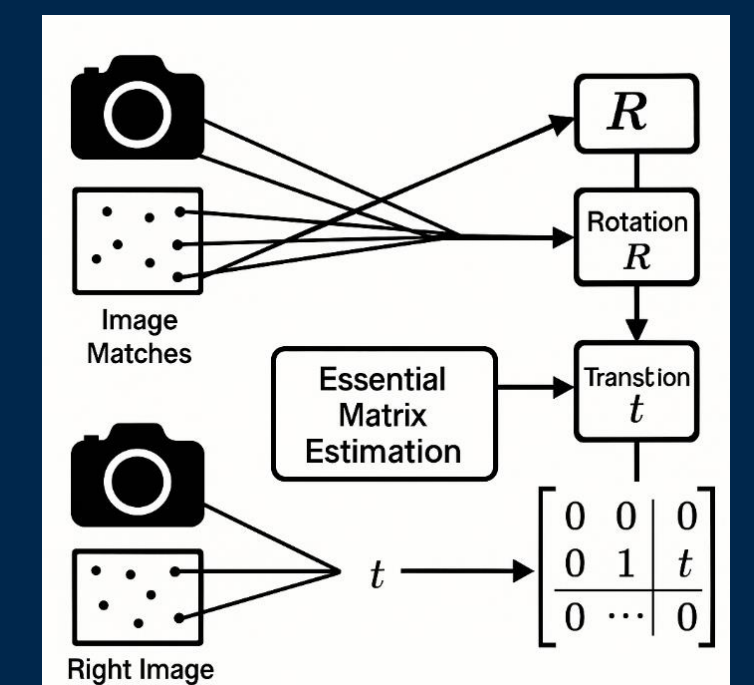

Monocular Grayscale KITTI Dataset

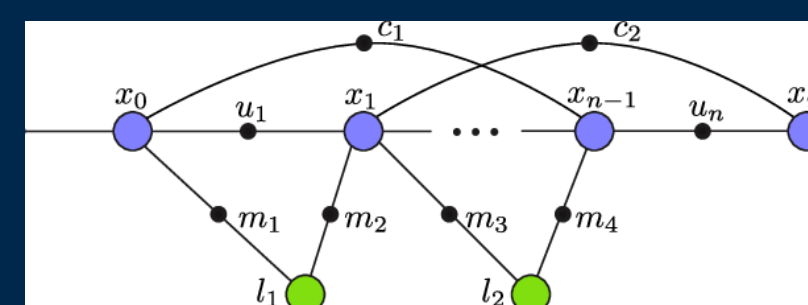### Keypoint Detection & Descriptor Generation (SuperPoint)

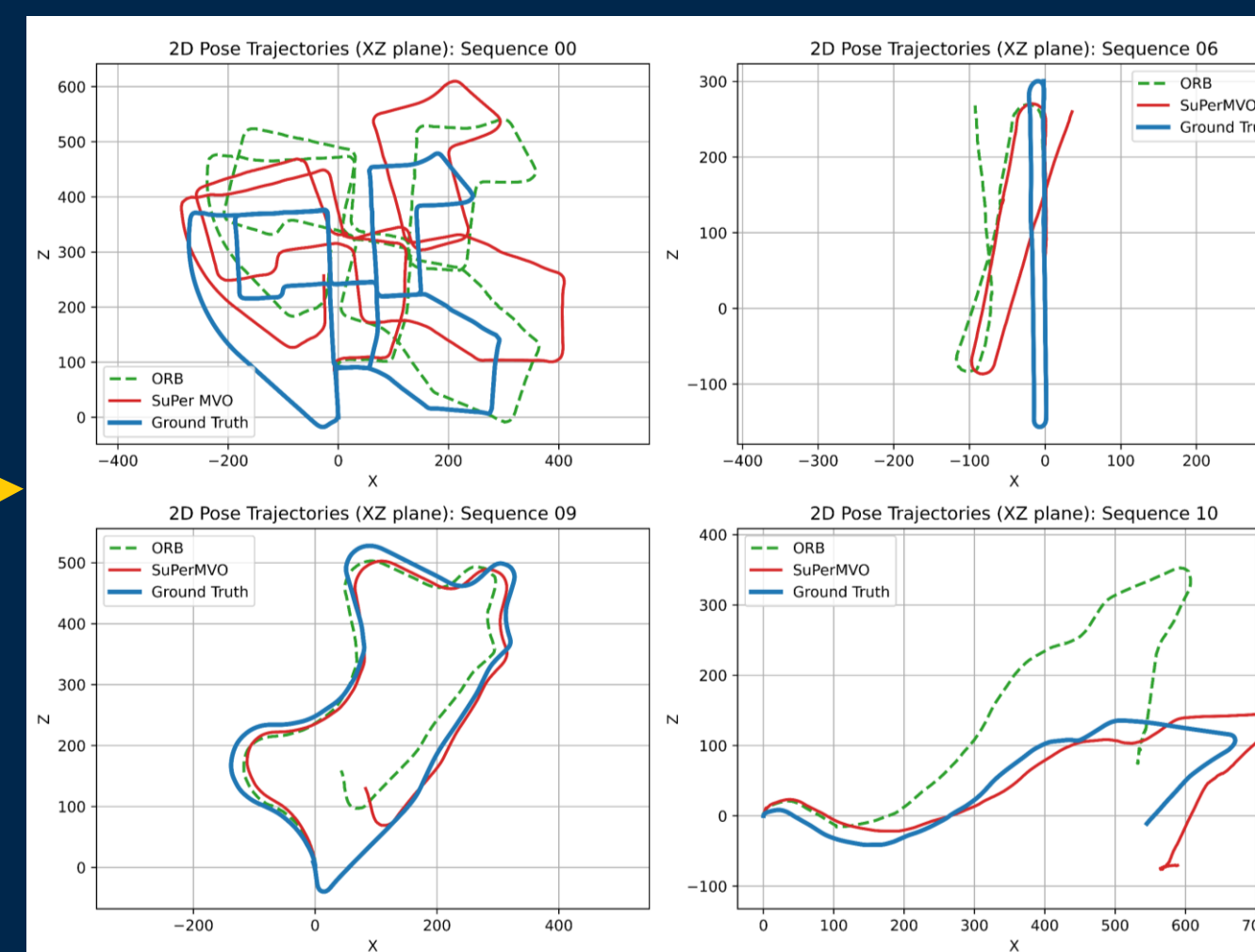### Feature Matching using SuperGlue
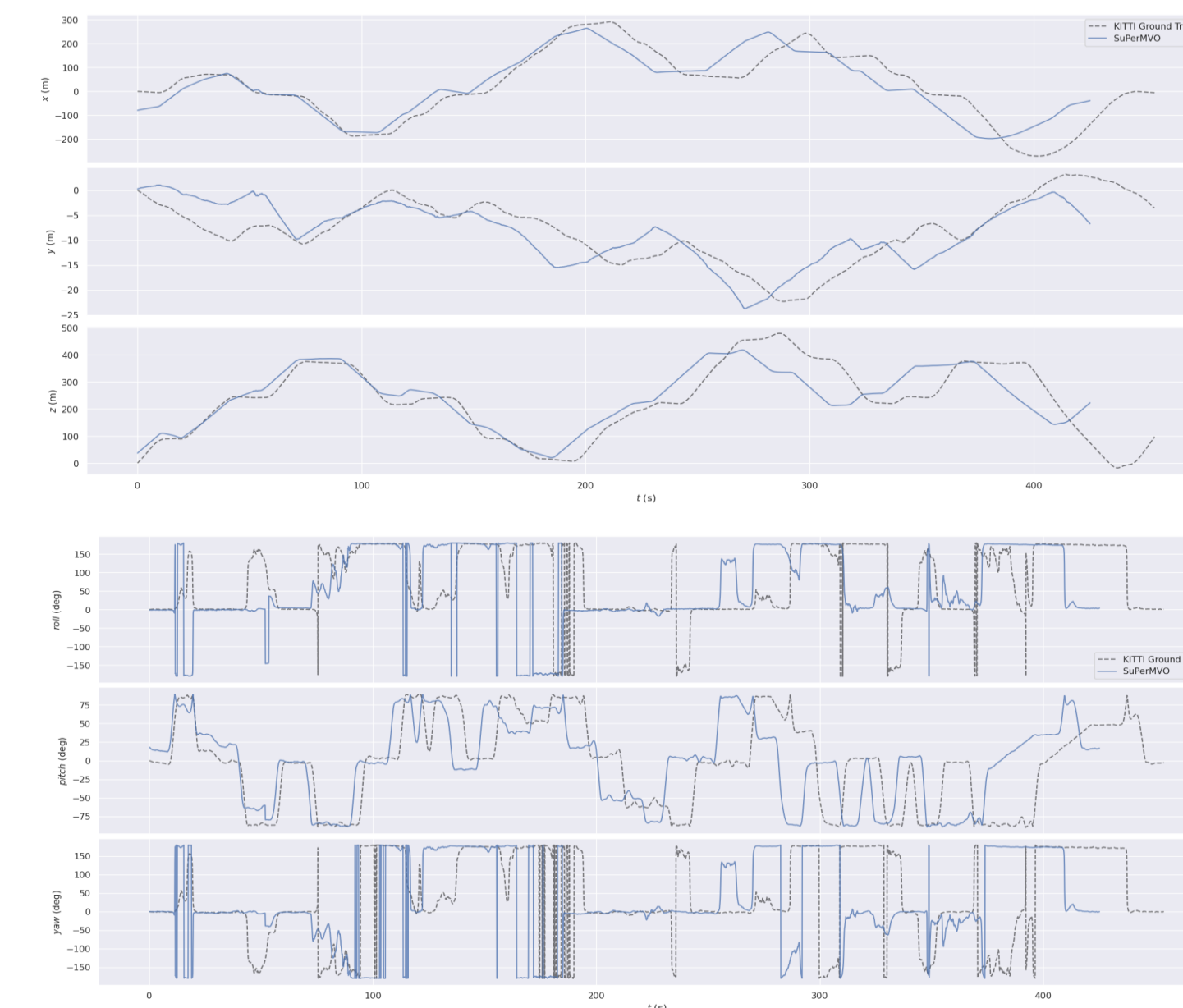

Pose Estimation using SE(3)

Factor Graph using GTSAM

Optimized Trajectory

Experimental Analysis

Results

Method (Sequence 10)	APME (↓)	Abs RMSE (↓)	RPME (↓)	Rel RMSE (↓)
SuPerMVO (Ours)	80.529	90.709	0.297	0.397
ORB-3	124.031	139.635	0.303	0.408
SuPerMVO-scaled (Evo)	51.086	58.392	0.260	0.363
ORB-3 scaled (Evo)	51.865	60.004	0.273	0.381

Future Work

- **Transformer Loop Closure** – Robust matching
- **Sensor Fusion** – Scale & drift correction
- **SLAM Integration** – Embed state-of-the-art SLAM

References

1. DeTone, Daniel, Tomasz Malisiewicz, and Andrew Rabinovich. "Superpoint: Self-supervised interest point detection and description." In Proceedings of the IEEE conference on computer vision and pattern recognition workshops, pp. 224-236. 2018.
2. Sarlin, Paul-Edouard, Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. "Superglue: Learning feature matching with graph neural networks." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 4938-4947. 2020.
3. Hess, Wolfgang, Damon Kohler, Holger Rapp, and Daniel Andor. "Real-time loop closure in 2D LIDAR SLAM." In 2016 IEEE international conference on robotics and automation (ICRA), pp. 1271-1278. IEEE, 2016.
4. https://www.cvlibs.net/datasets/kitti/eval_odometry.php.
5. Campos, Carlos, Richard Elvira, Juan J. Gómez Rodríguez, José MM Montiel, and Juan D. Tardós. "Orb-slam3: An accurate open-source library for visual, visual–inertial, and multimap slam." IEEE transactions on robotics 37, no. 6 (2021): 1874-1890.