Sticking Points and Lines for Improved Image Matching

Suárez Iago[1], Pautrat Rémi[2], Larsson Viktor[3], Pollefeys Marc[2,4]
[1]New Horizon Technologies, [2]ETH Zürich, [3]Lund University, [4]Microsoft

New Horizon Technologies · ETH zürich · POLITÉCNICA · graffter

ICVSS 2022
Sicily ~ 10-16 July
International Computer Vision Summer School

Abstract

Line segments are powerful features useful to complement points. They offer structural cues, robust to viewpoint and illumination changes. However, detecting, describing, and matching them is more challenging than points because of partial occlusions, lack of texture, or repetitiveness. This poster reviews the most common approaches to line matching. We show the potential of applying machine learning to stitch both types of features together, computing an improved matching.

Line Segment Detection

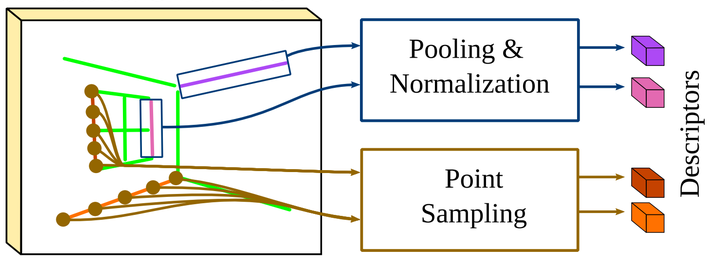
General-purpose approaches such as LSD [1], or ELSED [2] detect segments as elongated image regions with aligned gradients. There are also deep learning detectors that predict them, based of human-labels annotations or self-supervision, e.g. SOLD^2 [3].

ELSED SOLD^2

Line Segment Description

Next, methods encode the local appearance of the line segment in a distinctive way.

- LBD [4] uses a histogram of gradient orientations. Analogously, other methods use a CNN to extract deep features, followed by different kinds of pooling.
- SOLD^2 [3] samples points from each segment to extract its features.



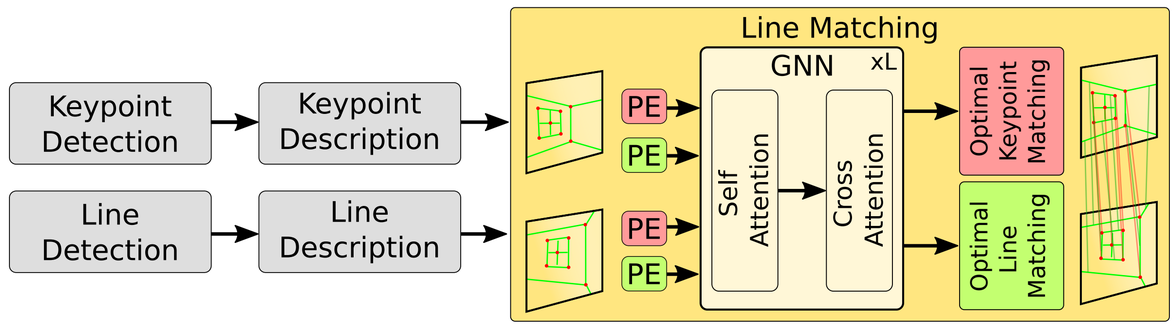
We choose Point Sampling because it can use the same description space as keypoints and is robust to partial occlusions when matched with Needleman-Wunsch algorithm.

Future Work

- Find a better way of representing line segments inside the GNN
- Learn a specific positional encoder for line segments
- Reduce the computational cost
- Make the line matching also contributes to improving the point matching results
- Evaluate localization performance

Line Segment Matching

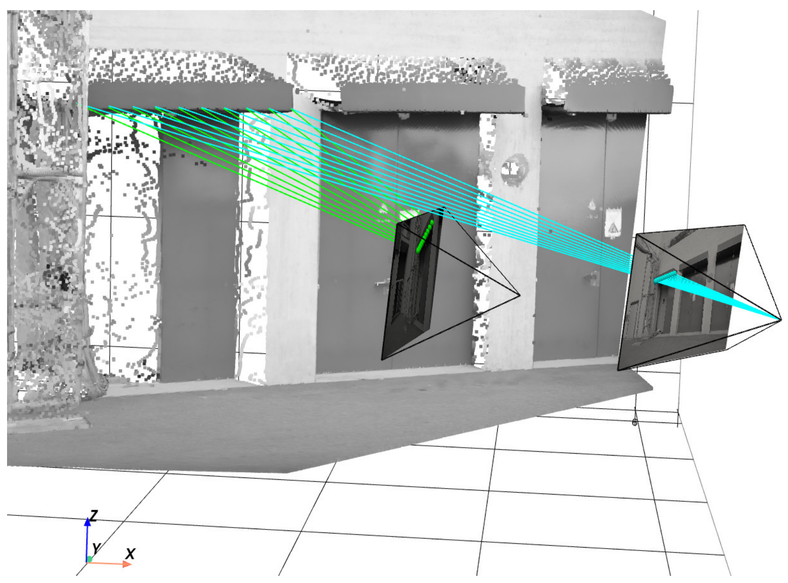
Line segments appear in texture-less repetitive scenes where appearance-based NN matching fails. Some works guide this matching with heuristics (assuming global rotation, planar scenes, etc), pre-computed point matches, or epipolar geometry. Inspired by point matching with Graph Neural Networks (GNN's) [5], we learn it. We match points and lines jointly, representing each line by its endpoints.



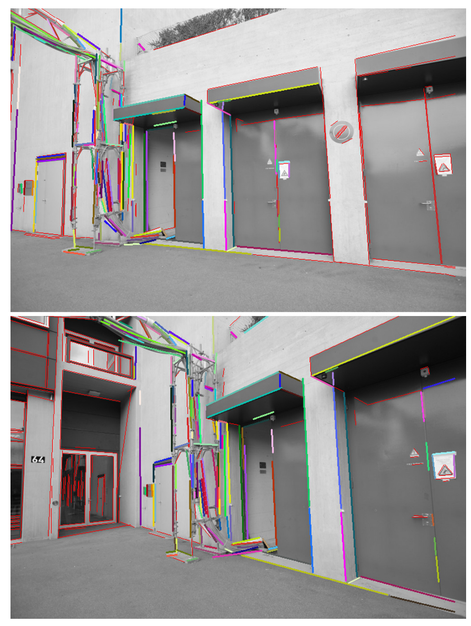
Letting line nodes attend to point nodes inside the GNN allows them to re-use the transformation between images, disambiguating between lines with similar appearance.

Evaluation

One of the more challenging tasks in line matching is generating high-quality ground truth (GT) that fairly treats line segment fragmentation, assignation, and partial visibility.

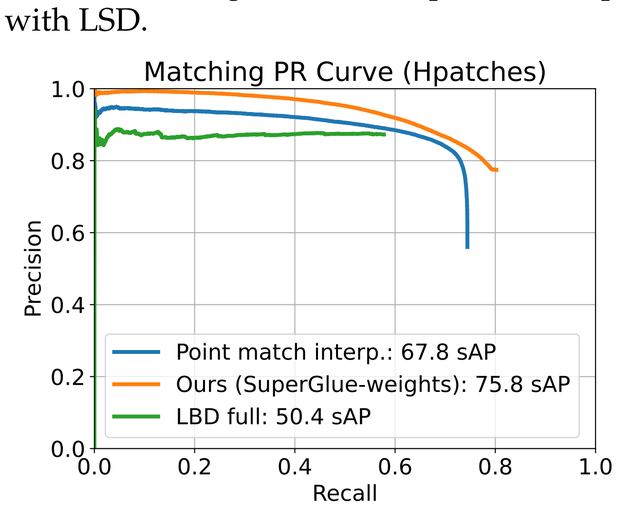


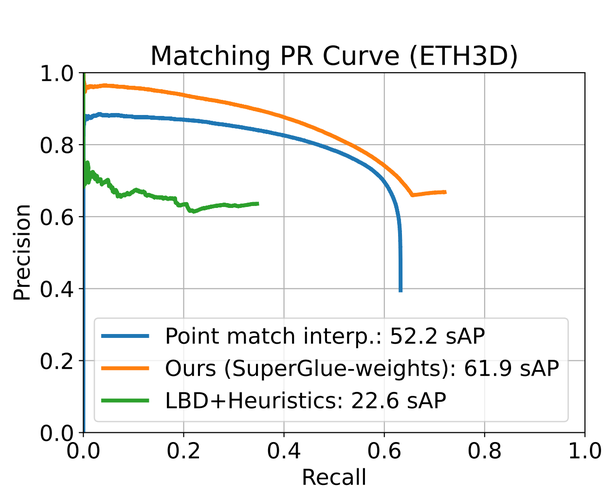
GT Generation: We sample K points along each segment, we project them to the other image and compute a cost matrix. By optimal transport we get the final assignations.



GT resulting assignations: Reference (top) and target (bottom) images

We compare the matching results for Hpatches [6] (planar) and ETH3D [7] (non-planar) datasets. Segments are detected with LSD.





References

[1] R. Grompone von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, "LSD: A fast line segment detector with a false detection control," *IEEE TPAMI*, vol. 32, no. 4, pp. 722–732, 2010.

[2] I. Suárez, J. M. Buenaposada, and L. Baumela, "ELSED: Enhanced line segment drawing," *Pattern Recognition*, vol. 127, p. 108619, 2022.

[3] R. Pautrat, J.-T. Lin, V. Larsson, M. R. Oswald, and M. Pollefeys, "SOLD^2: Self-supervised occlusion-aware line description and detection," in *CVPR*, 2021, pp. 11 368–11 378.

[4] L. Zhang and R. Koch, "An efficient and robust line segment matching approach based on lbd descriptor and pairwise geometric consistency," *JVCIR*, vol. 24, no. 7, pp. 794–805, 2013.

[5] P.-E. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich, "SuperGlue: Learning feature matching with graph neural networks," in *CVPR*, 2020.

[6] V. Balntas, K. Lenc, A. Vedaldi, and K. Mikolajczyk, "HPatches: A benchmark and evaluation of handcrafted and learned local descriptors," in *CVPR*, 2017.

[7] T. Schops, J. L. Schonberger, S. Galliani, T. Sattler, K. Schindler, M. Pollefeys, and A. Geiger, "A multi-view stereo benchmark with high-resolution images and multi-camera videos," in *CVPR*, 2017.