

S-TREK: Sequential Translation and Rotation Equivariant Keypoints for local feature extraction

Emanuele Santellani ^{*}, Christian Sormann ^{*}, Mattia Rossi [†], Andreas Kuhn [†], Friedrich Fraundorfer ^{*}

^{*} Graz University of Technology, [†] Stuttgart Laboratory 1, SSS-EU, Sony Europe B.V.

^{*} name.surname@icg.tugraz.at, [†] name.surname@sony.com

1. Supplementary material

We include additional qualitative comparisons on *Image Matching Benchmark* [2] and our rotated versions with all the tested methods in Figure 1, Figure 2 and Figure 3.

In Figure 4 we include an extended version of the main

paper figure which shows the detector heatmaps evolution during training. We include the additional 32 and 96 *Peaky* loss window sizes and our training framework run using 100 serial samples.

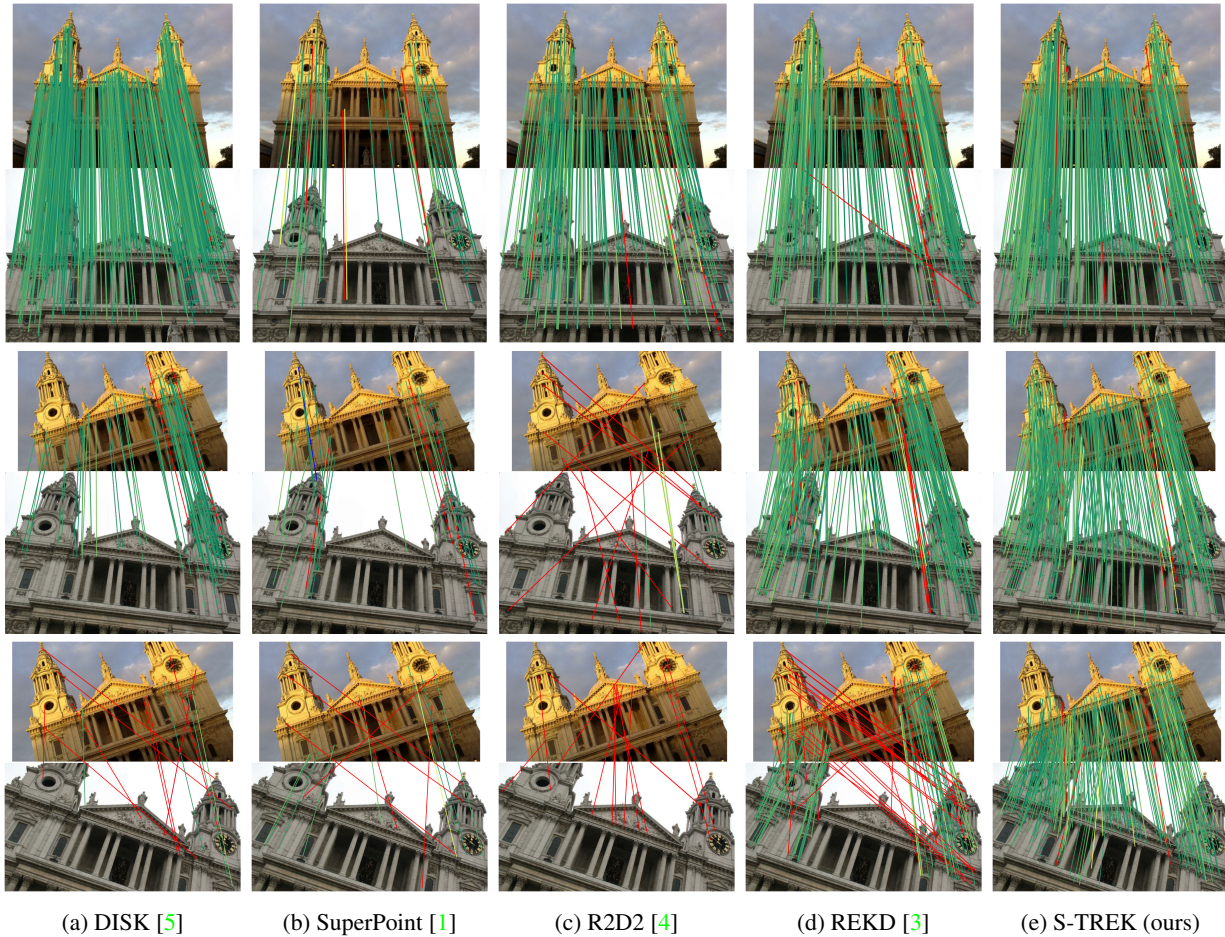
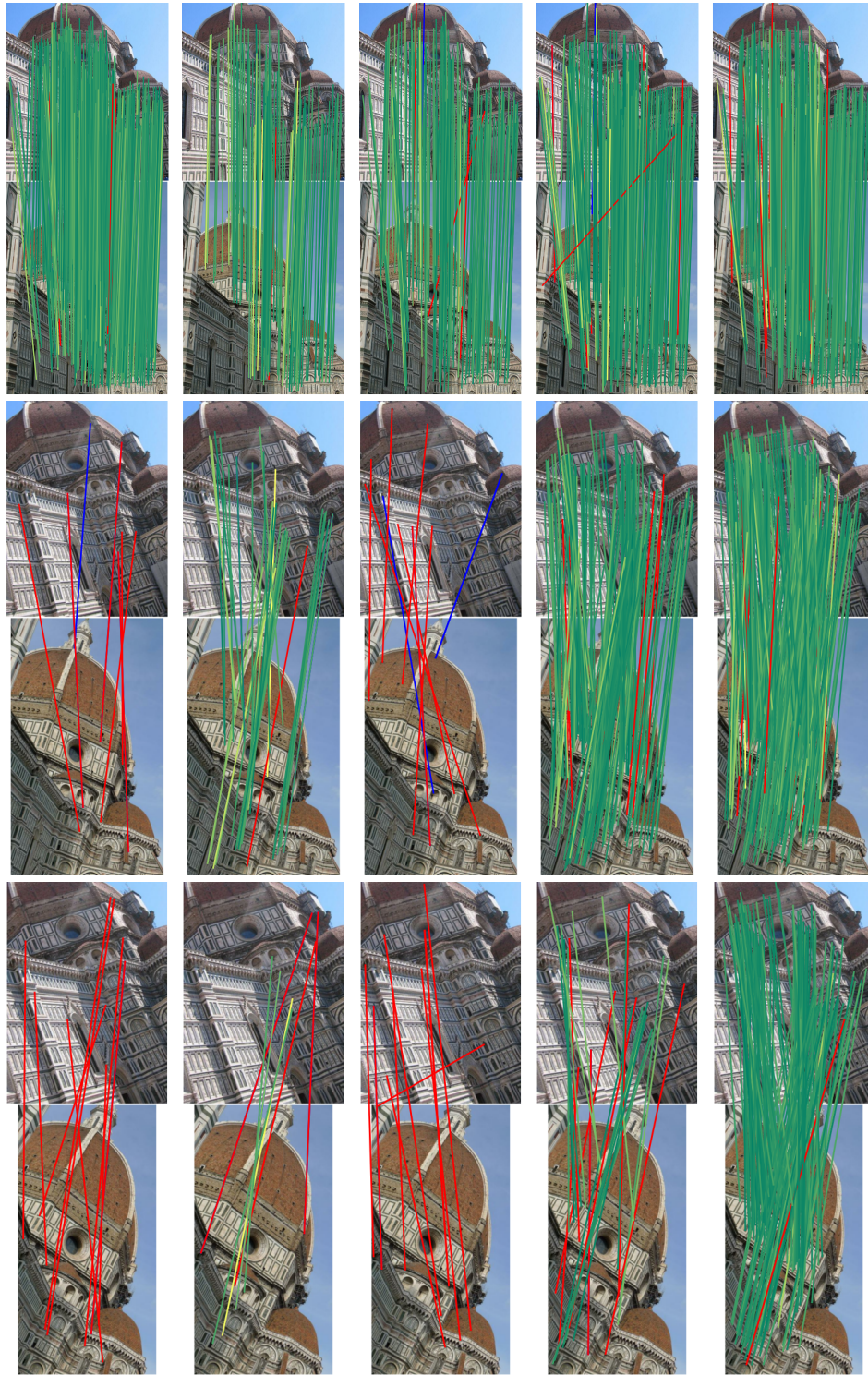


Figure 1: Qualitative comparison with state-of-the-art feature extraction methods on the Image Matching Benchmark [2] (top) and on our $\pm 20^\circ$ and $\pm 45^\circ$ rotated version of it. RANSAC inlier matches are color coded from green to yellow, representing reprojection errors equal to zero and 5px, respectively; the outlier matches are in red.



(a) DISK [5] (b) SuperPoint [1] (c) R2D2 [4] (d) REKD [3] (e) S-TREK (ours)

Figure 2: Qualitative comparison with state-of-the-art feature extraction methods on the Image Matching Benchmark [2] (top) and on our $\pm 20^\circ$ and $\pm 45^\circ$ rotated version of it. RANSAC inlier matches are color coded from green to yellow, representing reprojection errors equal to zero and 5px, respectively; the outlier matches are in red.

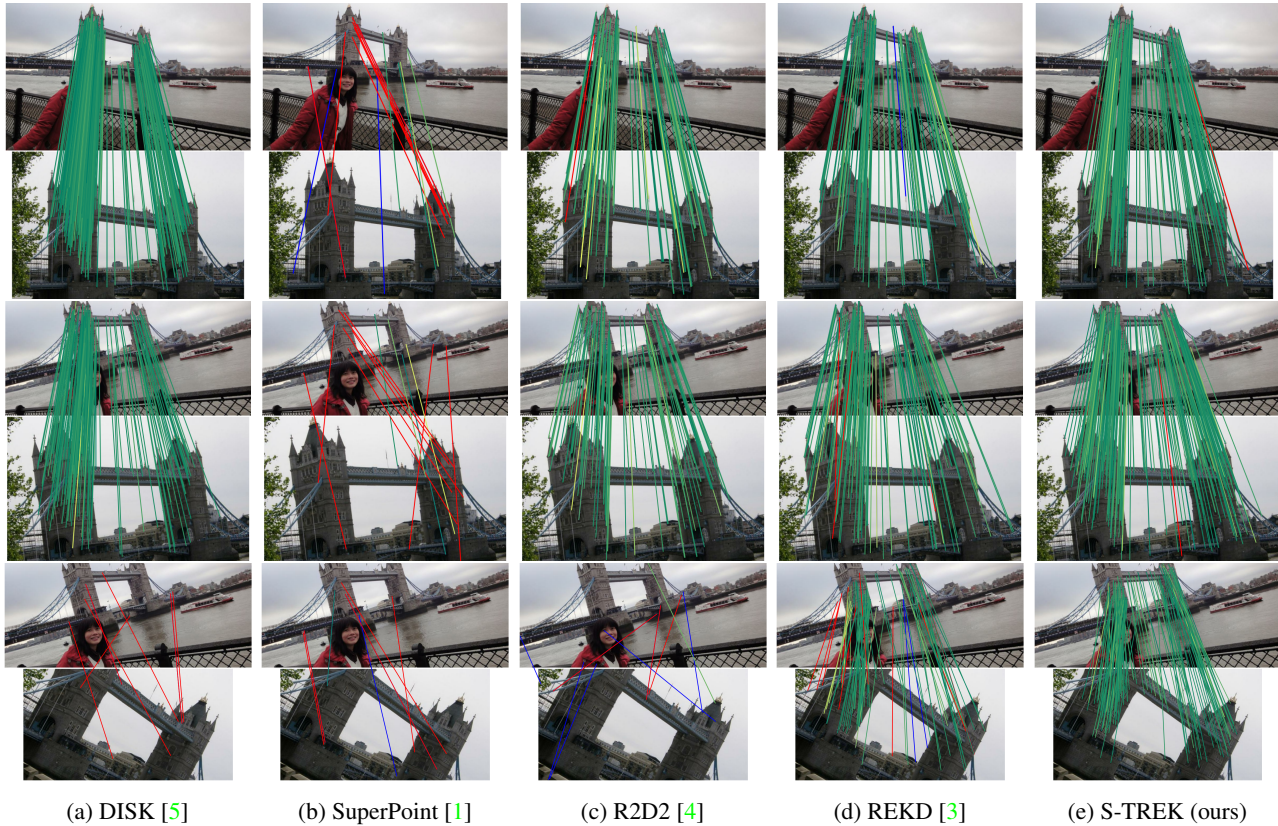


Figure 3: Qualitative comparison with state-of-the-art feature extraction methods on the Image Matching Benchmark [2] (top) and on our $\pm 20^\circ$ and $\pm 45^\circ$ rotated version of it. RANSAC inlier matches are color coded from green to yellow, representing reprojection errors equal to zero and 5px, respectively; the outlier matches are in red.

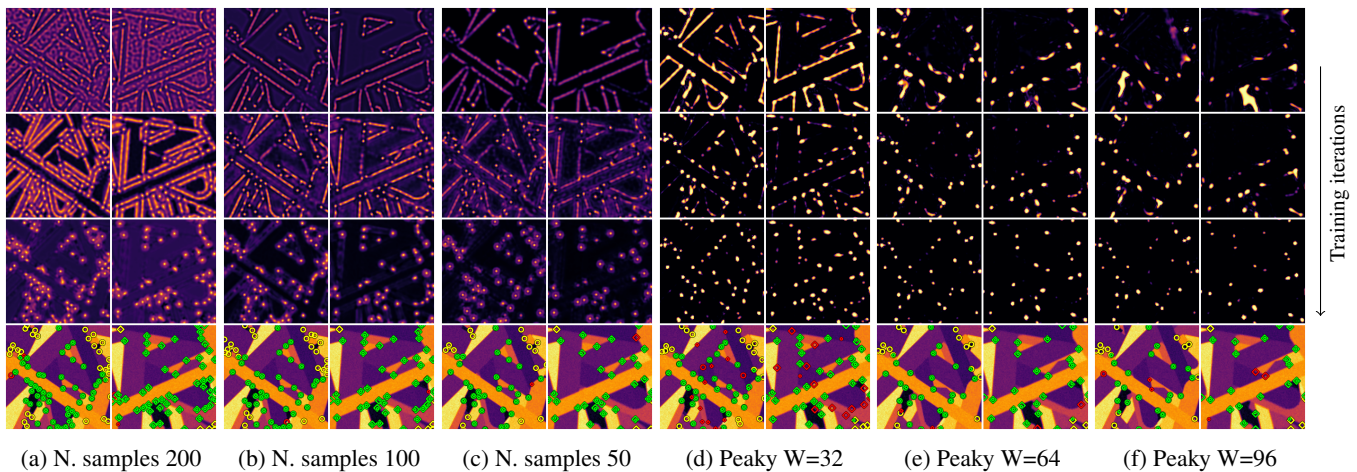


Figure 4: Visual comparison between different training methods on the validation set of the synthetic lines dataset. Last row: input images and detected keypoints (Green - repeatable, Red - non repeatable, Yellow - non overlapping). First three rows: detector heatmaps evolution during training. Our training framework excels at finding the repeatable points without any direct supervision.

References

- [1] Daniel DeTone, 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*, pages 224–236, 2018. [1](#), [2](#), [3](#)
- [2] Yuhe Jin, Dmytro Mishkin, Anastasiia Mishchuk, Jiri Matas, Pascal Fua, Kwang Moo Yi, and Eduard Trulls. Image Matching across Wide Baselines: From Paper to Practice. *International Journal of Computer Vision*, 2020. [1](#), [2](#), [3](#)
- [3] Jongmin Lee, Byungjin Kim, and Minsu Cho. Self-supervised equivariant learning for oriented keypoint detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2022. [1](#), [2](#), [3](#)
- [4] Jerome Revaud, Philippe Weinzaepfel, César Roberto de Souza, and Martin Humenberger. R2D2: repeatable and reliable detector and descriptor. In *NeurIPS*, 2019. [1](#), [2](#), [3](#)
- [5] Michał J Tyszkiewicz, Pascal Fua, and Eduard Trulls. Disk: Learning local features with policy gradient. *arXiv preprint arXiv:2006.13566*, 2020. [1](#), [2](#), [3](#)