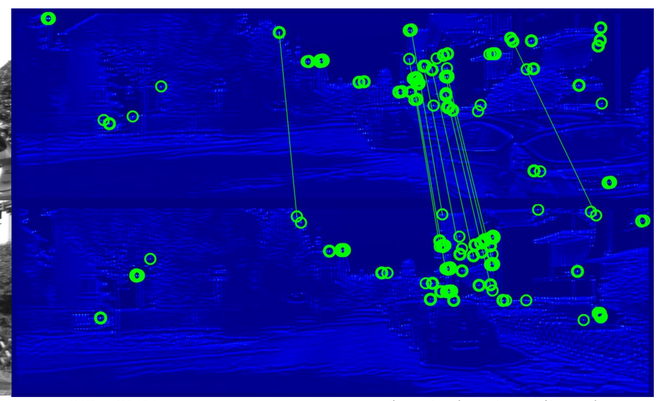
University of Zurich[UZH] ETH zürich Ryerson University ROBOTICS & PERCEPTION GROUP
Department of Informatics - Institute of Neuroinformatics

Sponsors: FNSNF Swiss National Science Foundation; robotics Swiss National Centre of Competence in Research; NVIDIA GPU Grant

SIPs: Succinct Interest Points from Unsupervised Inlierness Probability Learning

Titus Cieslewski, Konstantinos G. Derpanis, Davide Scaramuzza

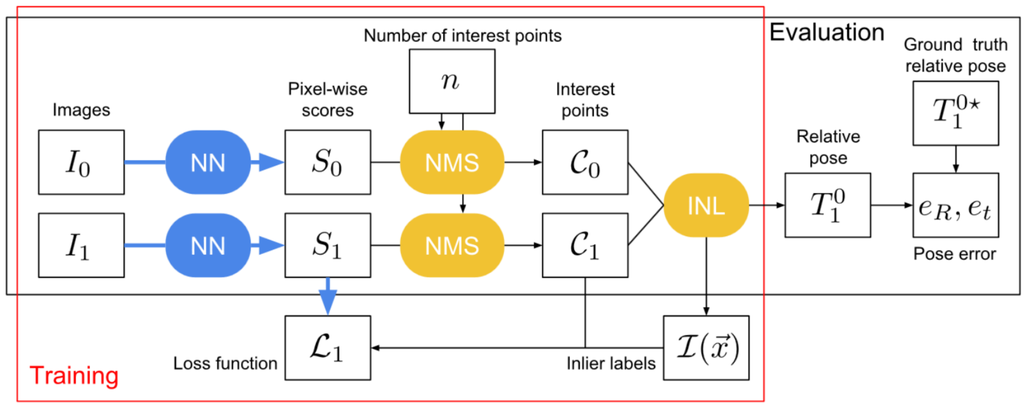
Input images

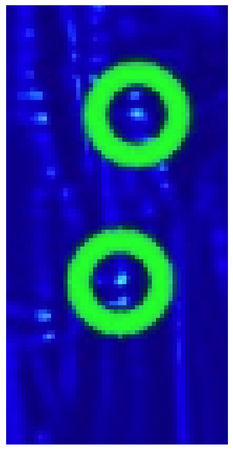


SIPs CNN output, interest points (circles), inliers (lines)

Given a detector, how many interest points are enough?
Can we train a detector to require as little interest points as possible?

Architecture





Peaked response without explicit training for peakedness

- Input: Image → Output: **Per-pixel** score
- **Non-maxima suppression** to obtain n interest points
- Match using a descriptor; here, SURF
- Loss: **Probabilistic classification of inliers**
- **Self-supervised** training, random initial weights, **converges to self-consistency**
- Trained on pairs of images
- Results in **peaked response** without explicitly training for this

Self-supervised Training

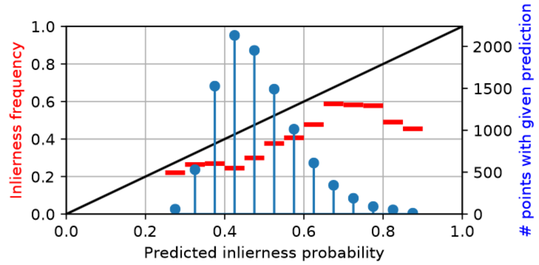
- From **uncalibrated** image sequences
- Select image pairs based on **visual overlap** (KLT track densely sampled points)
- **Ground truth** labels: **KLT track** interest point of one image into the other image
- **Any other labeling would also work** (e.g. inliers after RANSAC)

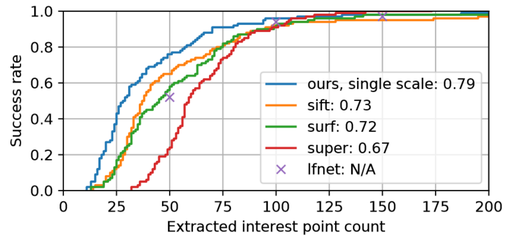
k-succinctness

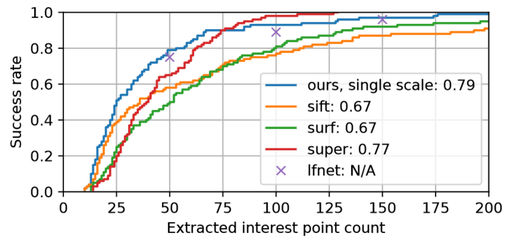
- A **novel metric** to **benchmark interest point** detectors
- **"How many points need to be detected to result in k inliers after matching and RANSAC?"**
- Plot **cumulative distribution** over set of image pairs; summarize with **area under curve**

Results

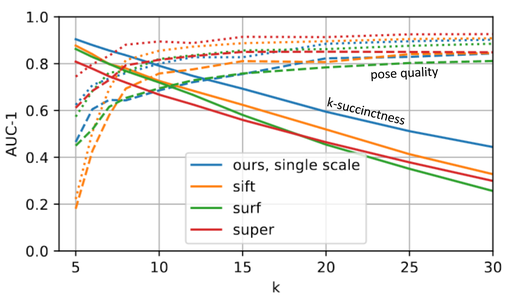
- Evaluated on **KITTI and EuRoC**, some **HPatches** results
- Relative pose estimation: Accuracy plateau reached with **10 inliers** or more → k = 10
- Our detector typically **requires less points** than baselines: **50 – 100 points often enough**
- Point score **predicts "inlierness" probability**

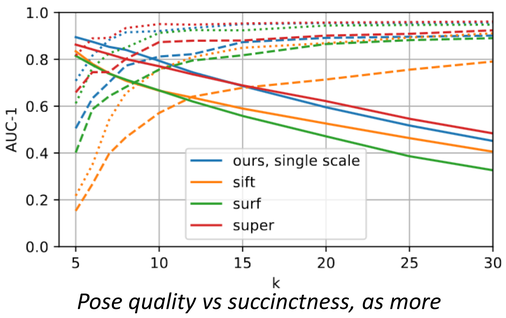






10-succinctness curves: KITTI (top), EuRoC (bottom)





Pose quality vs succinctness, as more inliers are required: KITTI (top), EuRoC (bottom)

Open source code:

https://github.com/uzh-rpg/sips2_open