SIPs: Succinct Interest Points from Unsupervised Inlierness Probability Learning

Titus Cieslewski, Konstantinos G. Derpanis, Davide Scaramuzza

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

Architecture

- **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

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 \( \rightarrow k = 10 \)
- Our detector typically requires less points than baselines: 50 – 100 points often enough
- Point score predicts “inlierness” probability

Open source code:
https://github.com/uzh-rpg/sips2_open