ETH zürich

A Radial Distortion Invariant Features Detector and Descriptor

Sepideh Mamooler

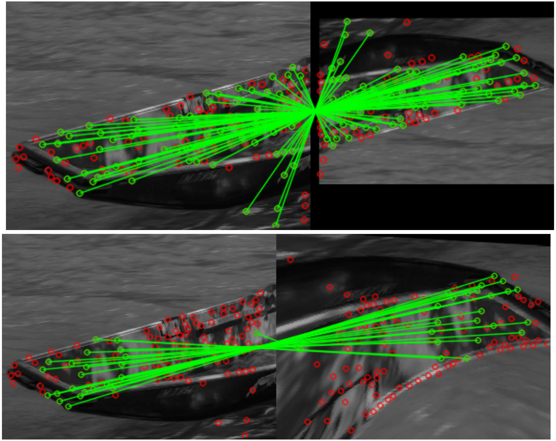
Supervisors: Rémi Pautrat, Dr. Viktor Larsson, Prof. Dr. Marc Pollefeys

Table of contents

Interest Point Detection and Description

- The first step in geometric computer vision tasks is to detect the interest points in the images.
- There exists many algorithms like SIFT, SURF, and SuperPoint that detect interest points and compute their corresponding descriptors.
- Many of these algorithms are invariant to image transformations like scale and rotation.
- But in the presence of radial distortion in the image, their performance downgrades considerably.

Interest Point Detection and Description



RD-Invariant Detectors and Descriptors

- sRD-SIFT [1]: proposes some modifications on SIFT to make it invariant to radial distortion.
- mdBRIEF [2]: a distorted and masked version of the BRIEF descriptor.

[1] M. Lourenco, J. P. Barreto , and A. Malt. sRD-SIFT: Keypoint Detection and Matching in Images With Radial Distortion. IEEE Transactions on Robotics, 2012

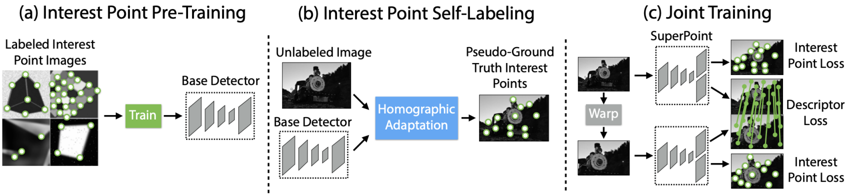
[2] S. Urban, and S. Hinz. mdBrief - A Fast Online Adaptable, Distorted Binary Descriptor for Real-Time Applications Using Calibrated Wide-Angle Or Fisheye Cameras. arXiv, 2016

Goal

- SuperPoint is a learned feature detector and descriptor.
- It is invariant to rotation, scale and illumination changes.
- But like other algorithms, it is not invariant to radial distortion.
- Our goal is to modify SuperPoint to make it invariant to radial distortion.

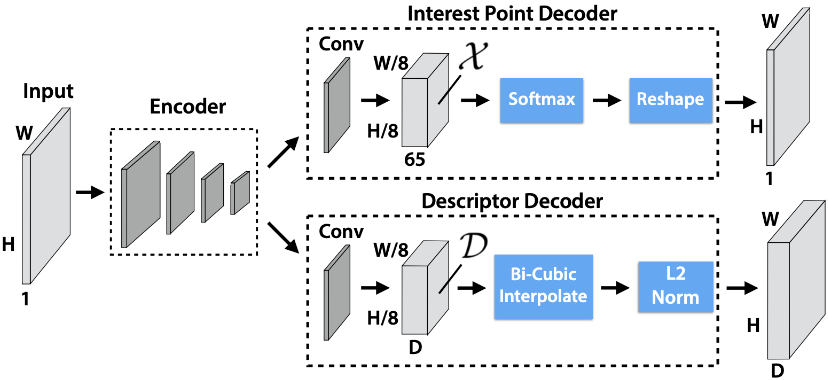
SuperPoint [3]

- SuperPoint is a self-supervised interest point detector and descriptor.
- Jointly computes pixel-level interest point locations and associated descriptors in a single forward pass.



[3] image taken from "SuperPoint: Self-Supervised Interest Point Detection and Description"

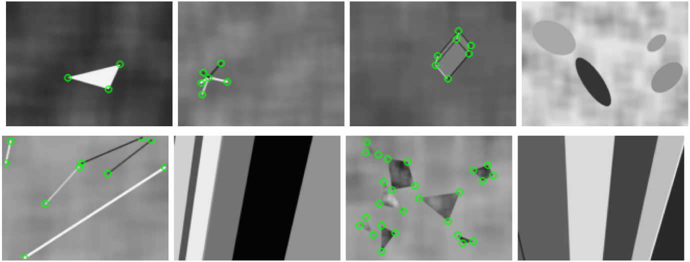
Base Detector [4]



[4] image taken from "SuperPoint: Self-Supervised Interest Point Detection and Description"

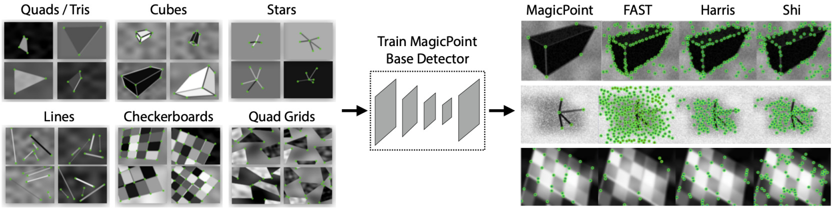
Synthetic Shapes [5]

- Synthetic Shapes are simple geometric shapes with no ambiguity in the interest point locations.



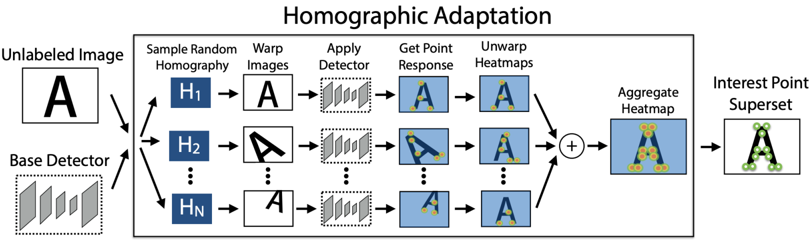
[5] image taken from SuperPoint's open source implementation of Paul-Édouard Sarlin and Rémi Pautrat

Interest Point Pre-Training [6]



[6] image taken from "SuperPoint: Self-Supervised Interest Point Detection and Description"

Homographic Adaptation [7] and Interest Point Self-Labeling



[7] image taken from SuperPoint

Joint Training

- The result of self-labeling phase is a pseudo-ground truth of interest points for the unlabeled input images.
- This pseudo-ground truth is used to train the network and obtain SuperPoint's interest point detector and descriptor.

RD-SuperPoint

- To make SuperPoint invariant to radial distortion we apply two modifications:
    1. **In pre-training and joined training phases:** in addition to random homographies, warp the images with radial distortion as well.
    2. **In self-labeling phase:** combine homographic adaptation with RD-adaptation to obtain homographic-RD-adaptation.

Distortion

- We use the "division model" to distort the images:

$$x_u = \frac{x_d - x_c}{1 + \lambda r^2} + x_c, \tag{1}$$

$$y_u = \frac{y_d - y_c}{1 + \lambda r^2} + y_c. \tag{2}$$

- Where
  - (x_u, y_u) is the pixel coordinate in the undistorted (original) image.
  - (x_d, y_d) is the pixel coordinate in the distorted image.
  - (x_c, y_c) is the coordinate of the distortion center.
  - λ is the distortion factor.
  - and

$$r = \sqrt{(x_d - x_c)^2 + (y_d - y_c)^2} \tag{3}$$

- Given the pixel coordinate of the distorted image, these formulas give the coordinate of its corresponding pixel in the original image.

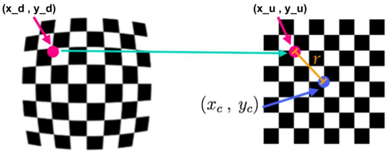
Undistortion

- We use the inverse of equations 1 and 2 to undistort the distorted image:

$$x_d = \frac{1 - \sqrt{1 - 4\lambda||x_u - x_c||^2}}{2\lambda||x_u - x_c||^2}(x_u - x_c) + x_c, \tag{4}$$

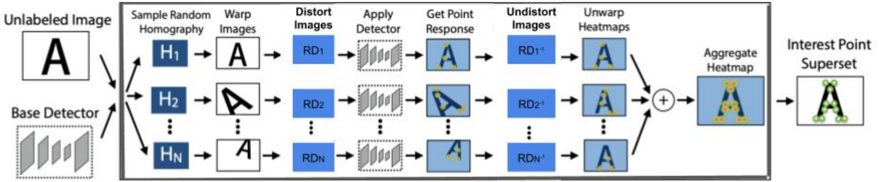
$$y_d = \frac{1 - \sqrt{1 - 4\lambda||y_u - y_c||^2}}{2\lambda||y_u - y_c||^2}(y_u - y_c) + y_c. \tag{5}$$

- Given the pixel coordinate of the undistorted image, these formulas give the coordinate of its corresponding pixel in the distorted image.



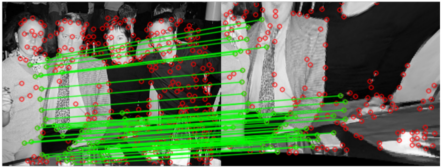
Distortion & Undistortion

Homographic-RD-adaptation

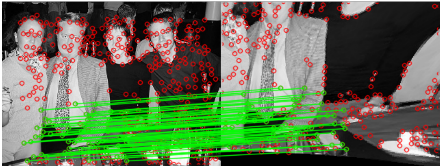


Visual Comparison

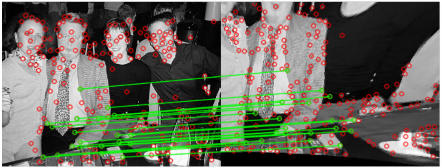
RD-SuperPoint



SuperPoint



SIFT

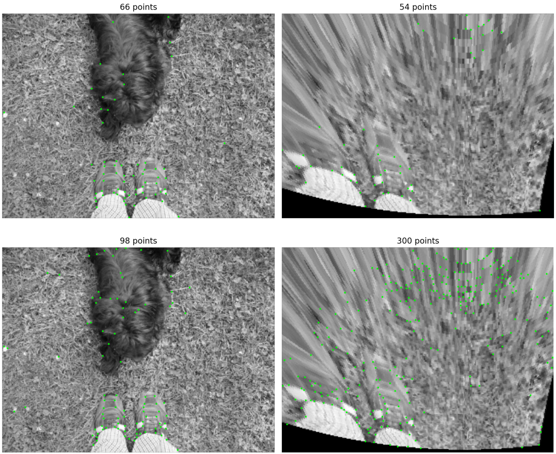


What to Evaluate?

Detector	Descriptor
Repeatability	Mean Average Precision
	Matching Score

Detector Evaluation

- Intuitively, repeatability rate evaluates the detector's geometric stability under different transformations.



RD-SuperPoint

SuperPoint

Detector Evaluation

- We have computed the repeatability rate for different detectors between pairs of images from MS-COCO dataset with at most 300 shared keypoints per image:

	Without RD	With RD
RD-SuperPoint	0.748	**0.461**
SuperPoint	0.681	0.437
Harris	**0.819**	0.457
Fast	0.692	0.414
Shi	0.684	0.376

- The results show that in the presence of radial distortion, detectors' repeatability rate considerably decrease.
- RD-SuperPoint has better repeatability rate than SuperPoint both in presence and absence of radial distortion.
- And it has the highest repeatability rate in the presence of radial distortion.

Descriptor Evaluation

- **Mean-Average precision (mAP)** evaluates how discriminating the descriptor is. It is computed as the Area Under Curve (AUC) of the Precision-Recall curve for multiple correctness thresholds.
- **Matching score** evaluates the performance of the detector and the descriptor. In simple words, it is the ratio of number of true positives over the number of features proposed by the pipeline in the shared viewpoint region.

Without Radial Distortion

- We computed the homography estimation using a correct distance equal 3, the mAP using correctness thresholds from 1 to 30, and the matching score using a correctness threshold of 3 for RD-SuperPoint, SuperPoint and SIFT:

	mAP	Matching Score
RD-SuperPoint	0.27	0.29
SuperPoint	0.23	0.25
SIFT	**0.35**	**0.31**

- The results show that SIFT has the best performance in the absence of radial distortion.

With Radial Distortion

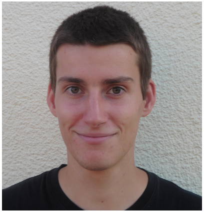
	mAP	Matching Score
RD-SuperPoint	0.26	**0.26**
SuperPoint	0.24	0.22
SIFT	**0.28**	**0.25**

- The results show that RD-SuperPoint is almost as good as SIFT in the presence of radial distortion.

Conclusion

- We showed that by applying two modifications we can improve SuperPoint's performance both in the presence and absence of radial distortion.
- These modifications provide a real-time interest point detector and descriptor even in the presence of radial distortion in the images without adding considerable complication to the original method and while preserving all the original invariance of SuperPoint.

Acknowledgement



Questions?


thank you