Learning Features for visual SLAM & SfM

Jonas Scheer, Mario Fritz, Oliver Grau

Abstract

Reducing the amount of image features as in [1] speeds-up SfM computation time, but still leaves a lot of redundant information. Our approach aims a higher feature reduction, by applying a classifier (random forest), optimized for SLAM/SfM. The main advantages are:

- Computational speed-up, due to fewer features
- Higher accuracy, due to lower amount of outliers
- Less redundant data in a SfM/SLAM pipeline

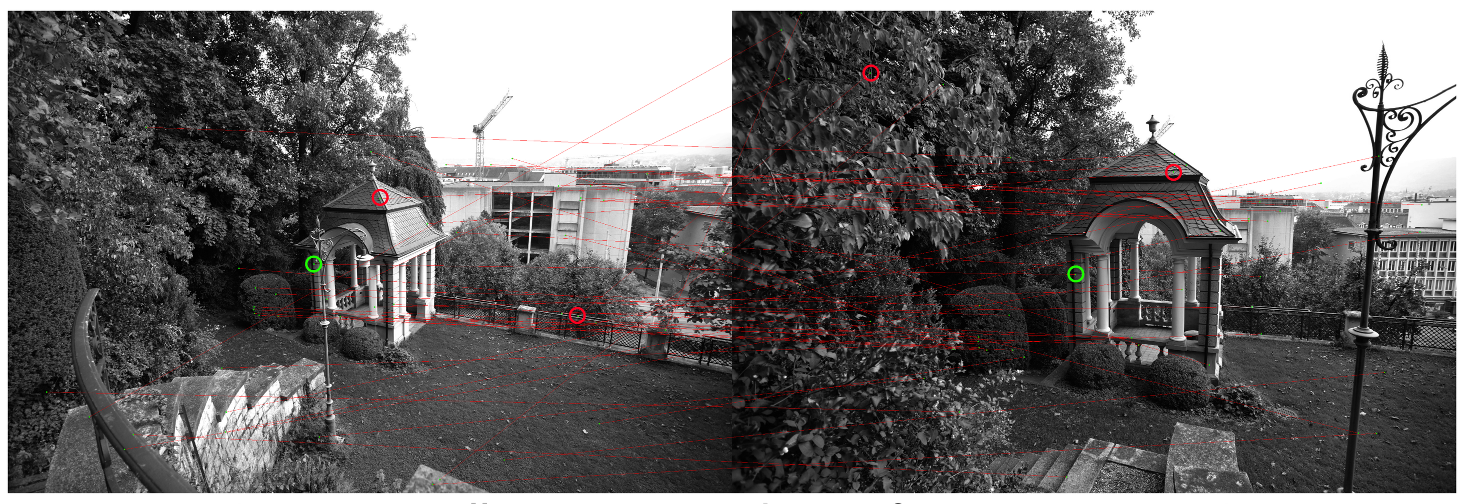
Predicting Matchability

Hartmann et al. [1] deploy a classifier, optimized for feature matching:

- Predicting SIFT [5] feature matches
- Speed-up subsequent matching procedure

Drawbacks:

- Predicting a single sift match does not consider the whole scene.
- wrong sift matches, according to an underlying 3D model



(geometrically wrong matching of 3D scene points)

(geometrically wrong matches across scenes)

Training with Long-Track Features

Solution: Training with more appropriate features as in [1]

$$All \supseteq Matchable \supseteq Geo = Track_1 \supseteq Track_{n>1}$$

All:	All available sift features
Matchable:	Features with valid sift-match as in [1]
Geo:	Features with valid 3D point
$Track_n$:	Features viewable from n cameras

—> Training Random-Forest [3] with Long-Track Features

Feature Reduction

Comparison with [1] and hDog (reducing features by increasing the sift threshold)

- 48% feature reduction
- Lower reproduction error —> more accurate 3D model

Fountain scene				
	feature amount	avg features	3D points	mean repr. error
all features	2545581	101823	221311	0.735637
high DoG	50686	2027	4596	0.746021
Hartmann et. al	39903	1596	2454	0.842059
Our Classifier	20379	815	1691	0.574793

Accuracy & Speed-up

We Create a reconstruction with all features and align the reconstructions with reduced features and compute the rotation & position error:

- Position error: 15% lower
- Rotation error: up to 30% lower

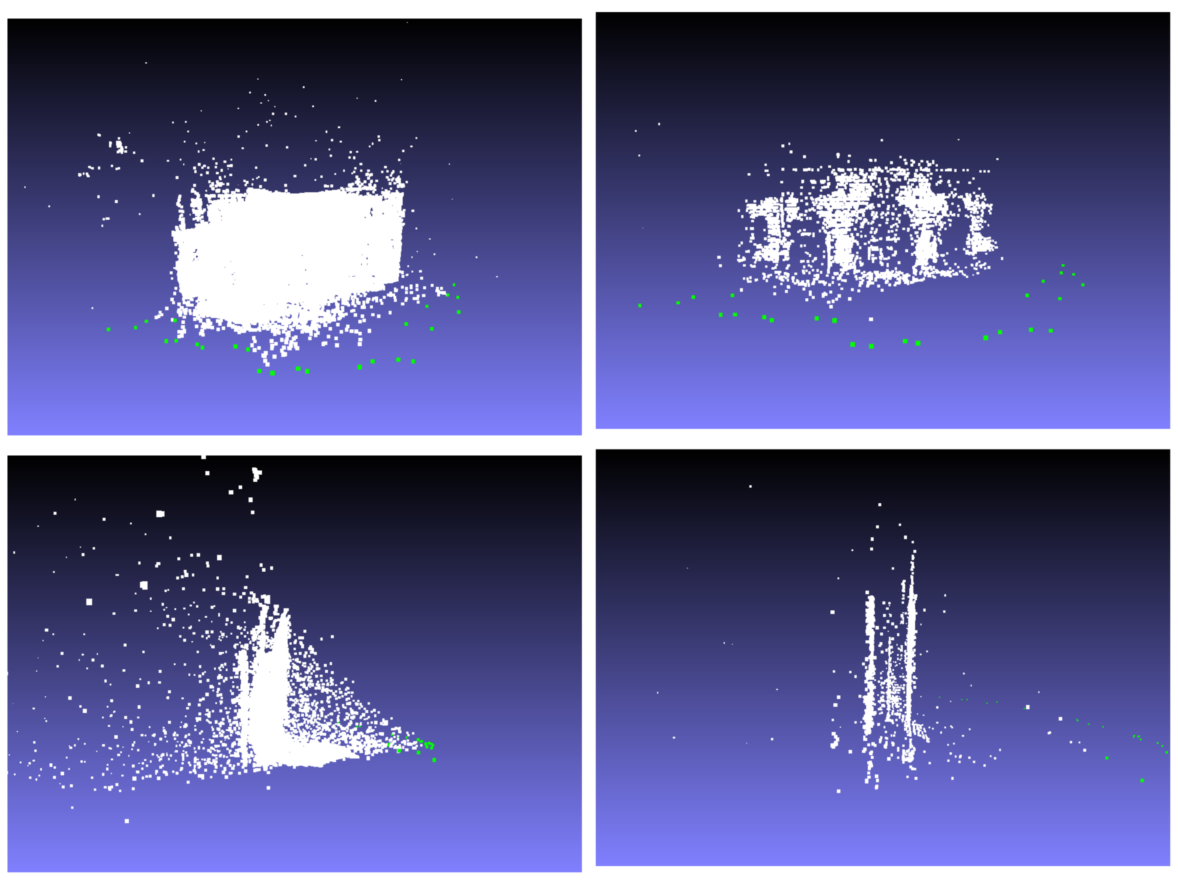
Fountain scene			
	3D points	mean position error	mean rotation error
high DoG	4624	0.005991	0.004881
Hartmann et. al	9129	0.006487	0.002644
Our Classifier	3586	0.005087	0.001826

- Computational speed-up: 60% (BA); 34% (Matching)

	features	feature matching	bundle adjustment
all features	85459	3.477	0.146
Hartmann et. al	5263	0.0695	0.0162
Our Classifier	2107	0.0239	0.011

Point clouds & Outlier

- Dense model with many outliers (left)
- 3D model by using long-track features (right)



References

[1] Predicting Matchability - W. Hartmann, M. Havlena, K. Schindler - CVPR 2014
[2] D. G. Lowe, "Distinctive image features from scale-invariant keypoints,"
[3] W. Stefan, "random-forests." https://github.com/stefan-w

Contact: jonas@jscheer.de