

Robust Large-Scale Localization

Fabian Tschopp, Marco Zorzi

Supervisor: Torsten Sattler

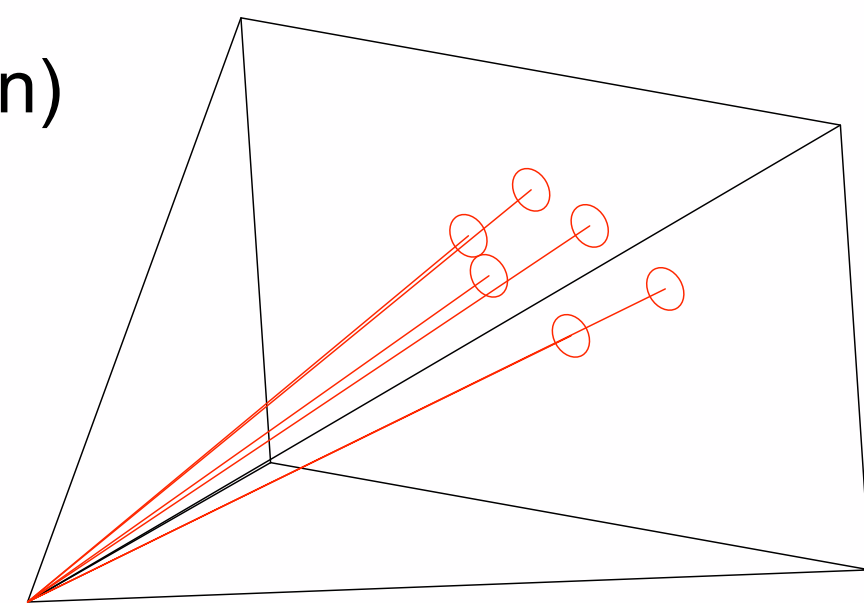
1 Overview

We tackle the problem of getting a full 6-DOF pose estimation of a query image inside a given point cloud. This project re-evaluates the algorithms proposed by Y. Li et al. "Worldwide Pose Estimation using 3D Point Cloud" [3]. Our code computes poses from 3 or 4 points, with both known and unknown focal length. The results can easily be displayed and analyzed with Meshlab. We found both advantages and shortcomings of the methods proposed. Furthermore, additional priors and parameters for point selection, RANSAC and pose quality estimate (inlier test) are proposed and applied. Our new benchmarking analysis gives insights on the problems of completely off poses and proposes algorithmic improvements that could be made.

2 Implementation

Program structure:

1. Load Dubrovnik datasets (golden + information)
2. Choice of basic or advanced processor
3. SIFT averaging and kd-tree population
4. Query an image
5. Start the benchmarking phase
6. Export model for Meshlab visualization
7. Export camera project and projection



- G, F : Sets of good and fit matches.
- w, h : Image width and height.
- $c = \frac{w}{40}$: Half of the area size covered by a match.
- m_x, m_y : Match x and y coordinate on the picture.
- $I_{i,j}^g$: Image map of covered areas by good matches.
- $I_{i,j}^f$: Image map of covered areas by fitted matches.
- A^g : Image area covered by good matches.
- A^f : Image area covered by fitted matches.

$$\forall m \in G. i = m_x - c, \dots, m_x + c, \quad (1)$$

$$j = m_y - c, \dots, m_y + c \quad I_{i,j}^g = 1 \quad (2)$$

$$\forall m \in F. i = m_x - c, \dots, m_x + c, \quad (3)$$

$$j = m_y - c, \dots, m_y + c \quad I_{i,j}^f = 1 \quad (4)$$

$$A^g = \sum_{i=1, j=1}^{i=w, j=h} I_{i,j}^g \quad (5)$$

$$A^f = \sum_{i=1, j=1}^{i=w, j=h} I_{i,j}^f \quad (6)$$

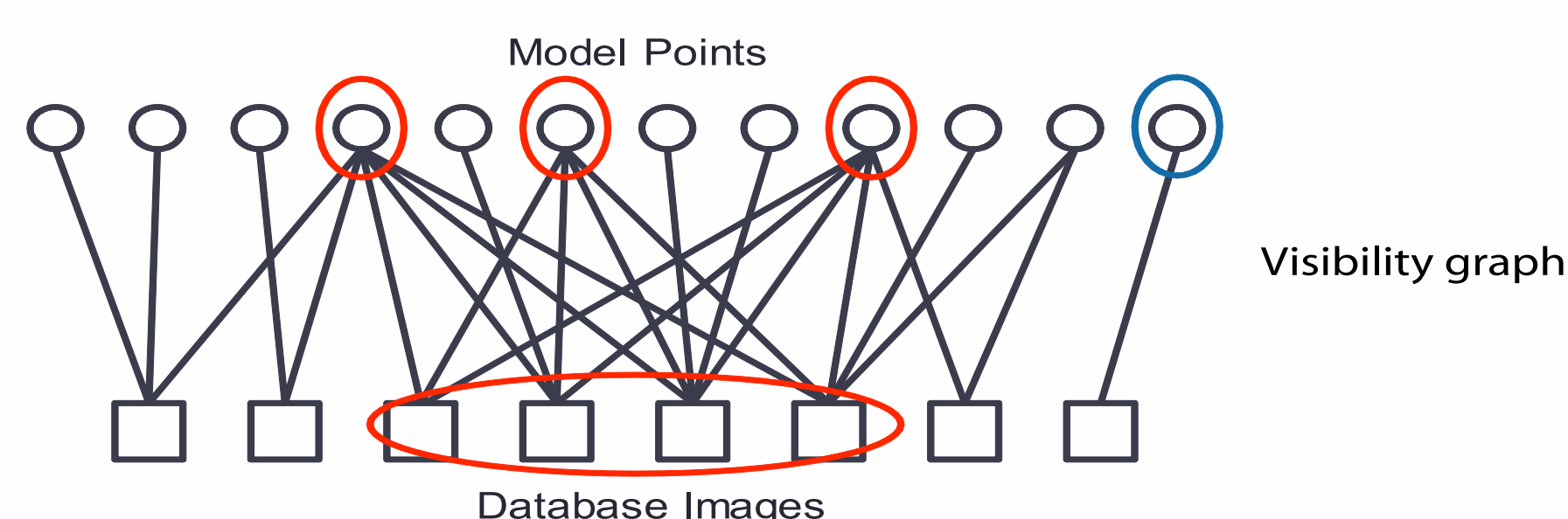
$$q = \frac{A^f}{A^g} \quad (7)$$

Figures: Matches found in a picture, circles filtering technique realized for better features extraction.

3 Advanced approach

Key Steps:

1. Co-occurrence prior sampling
2. Probability of co-occurrence
3. Backmatching

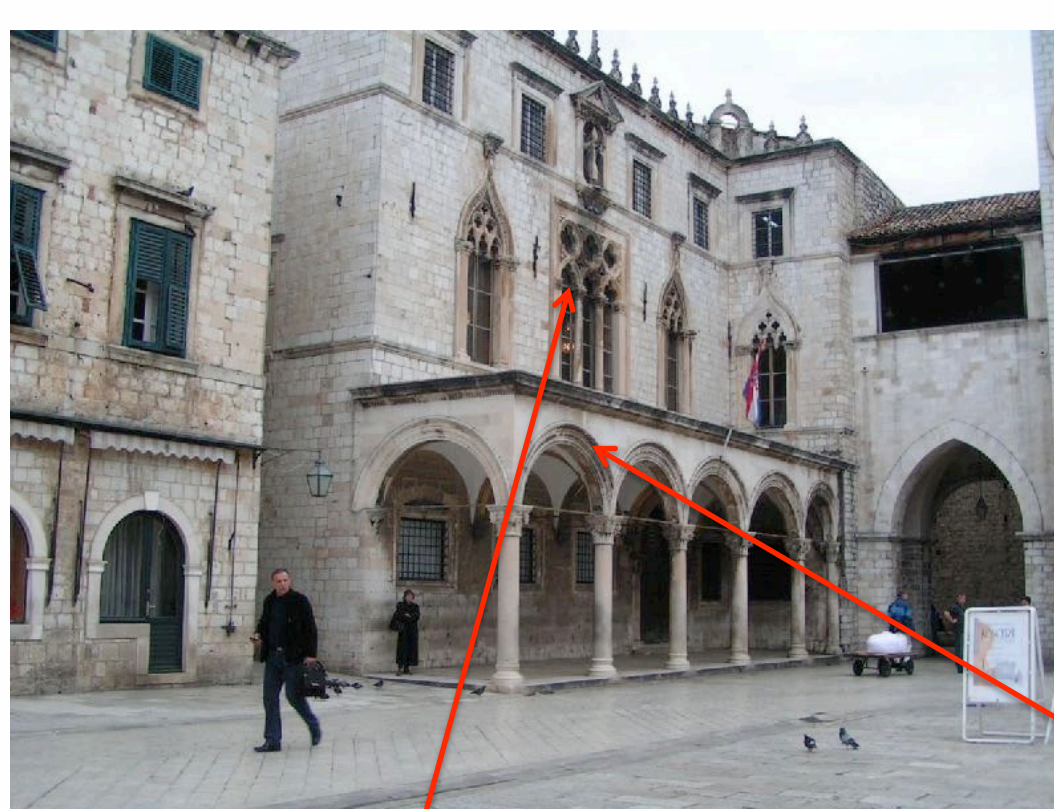


$$\tilde{P}r_{select}(p_i | p_1, \dots, p_{i-1}) \propto |A_{p_1} \cap \dots \cap A_{p_i}|$$

$$f_{scaling} = \frac{1}{1 + e^{-\frac{|A_{p_1} \cap \dots \cap A_{p_i}|}{k}}}$$

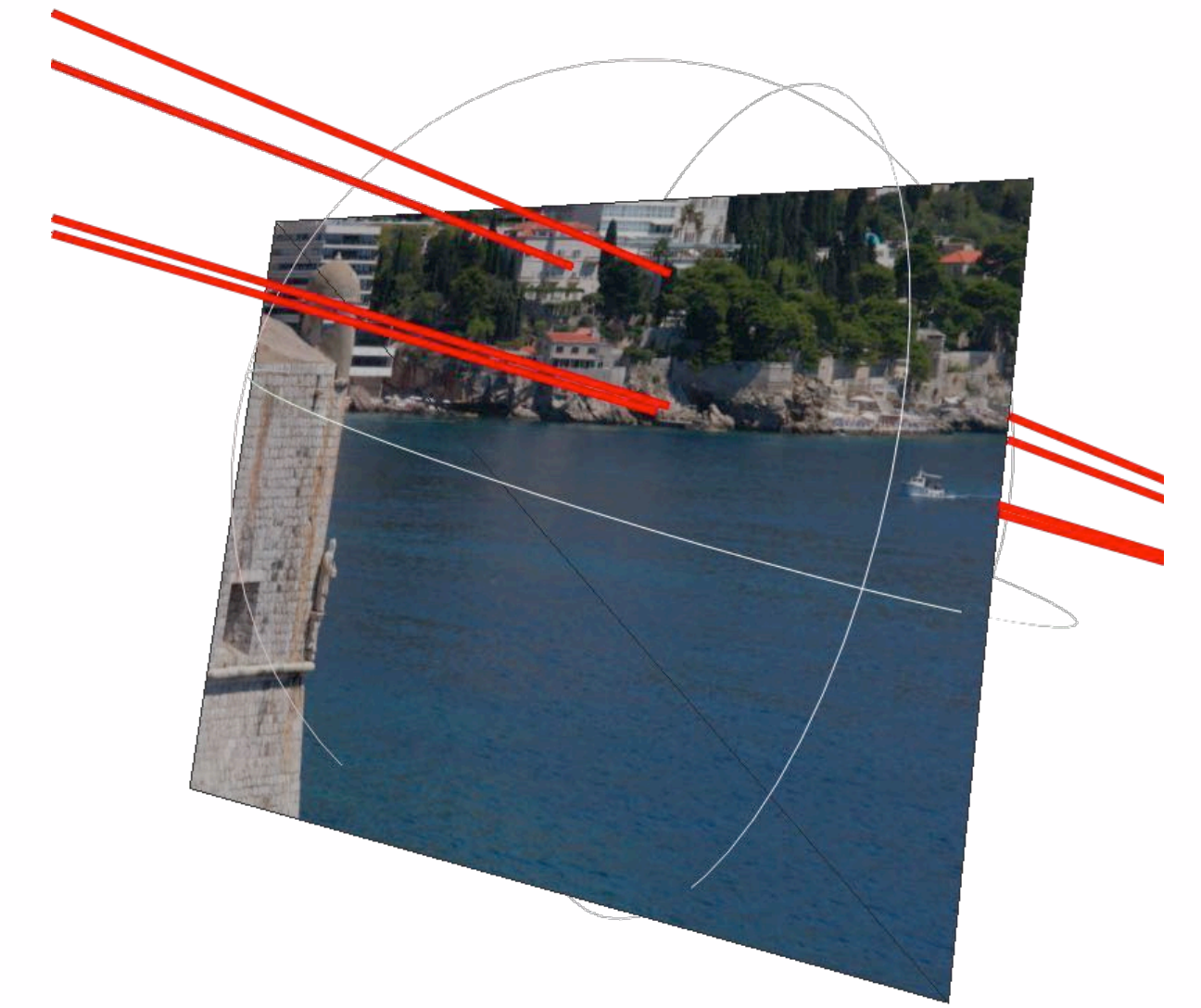
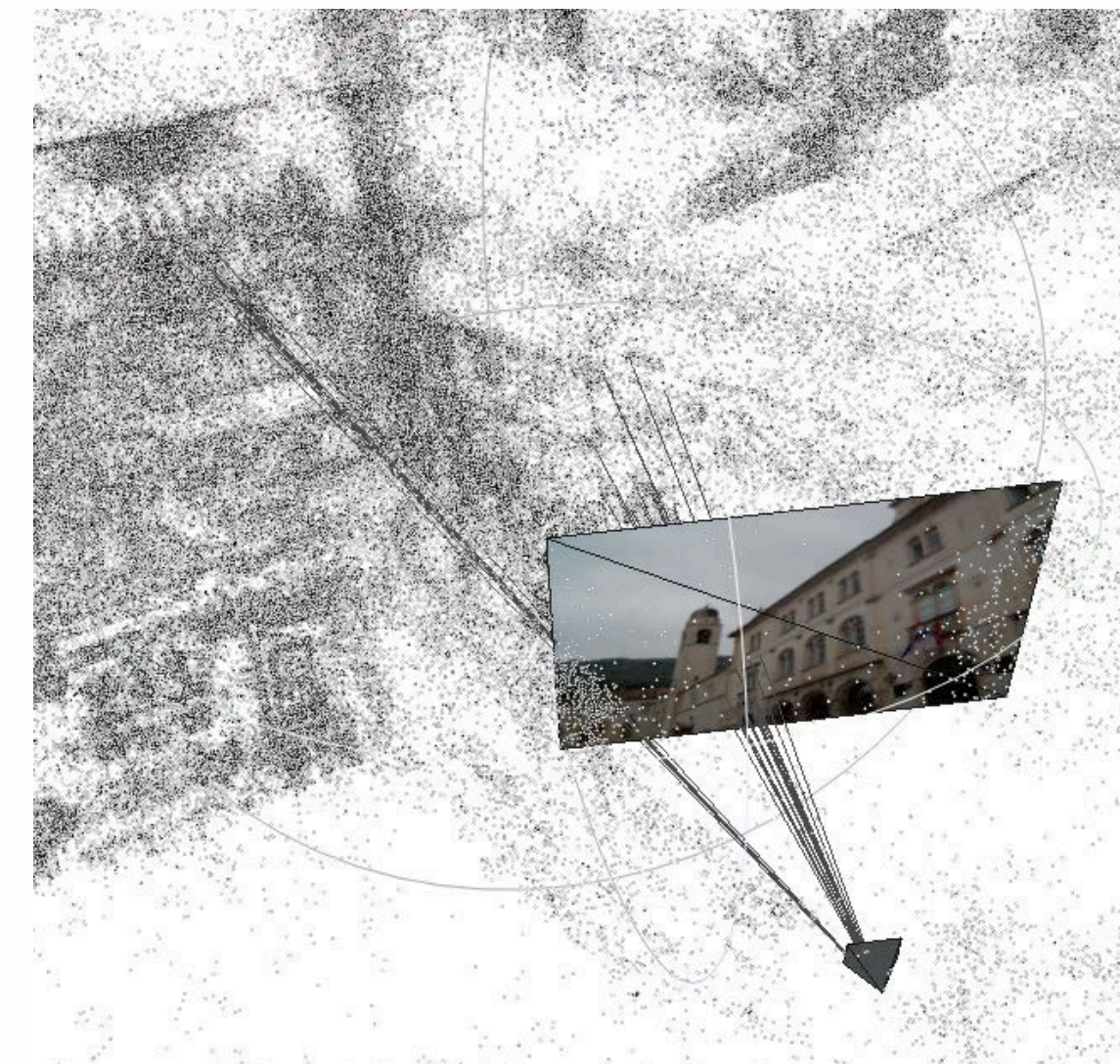
$$f_{ratio} = \frac{|A_{p_1} \cap \dots \cap A_{p_i}|}{\min(|A_{p_1} \cap \dots \cap A_{p_{i-1}}|, |A_{p_i}|)}$$

$$\tilde{P}r_{select}(p_i | p_1, \dots, p_{i-1}) = f_{scaling} \cdot f_{ratio}$$

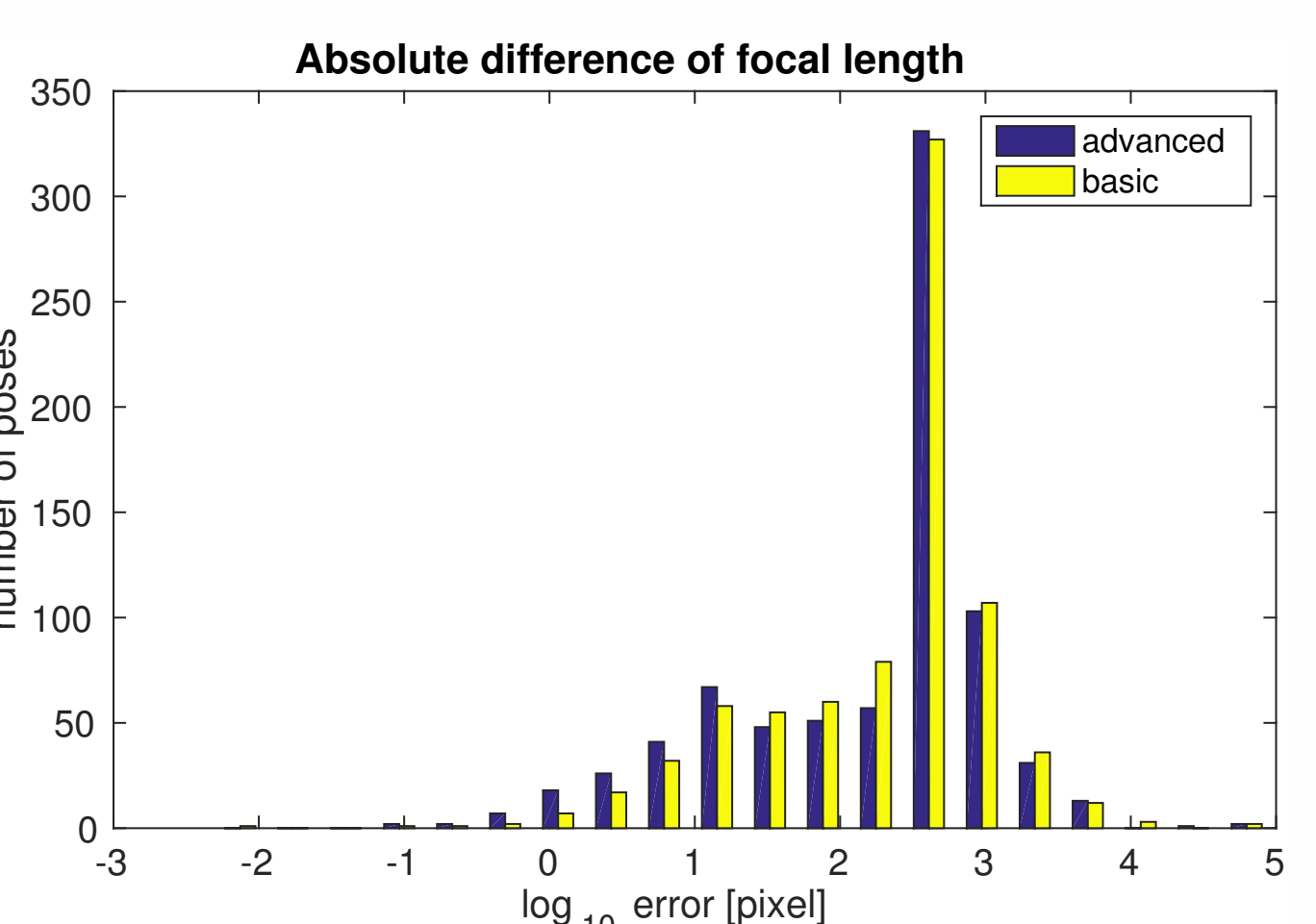
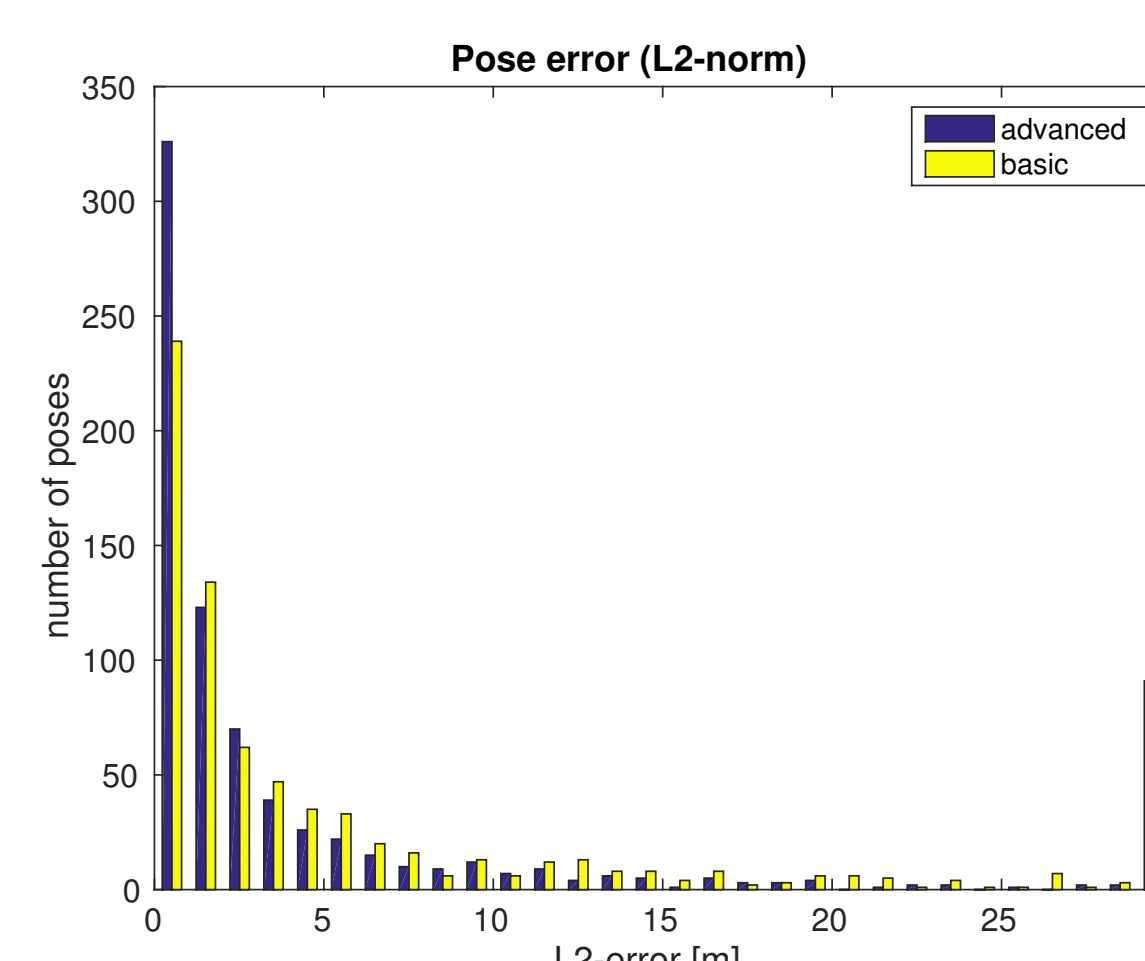
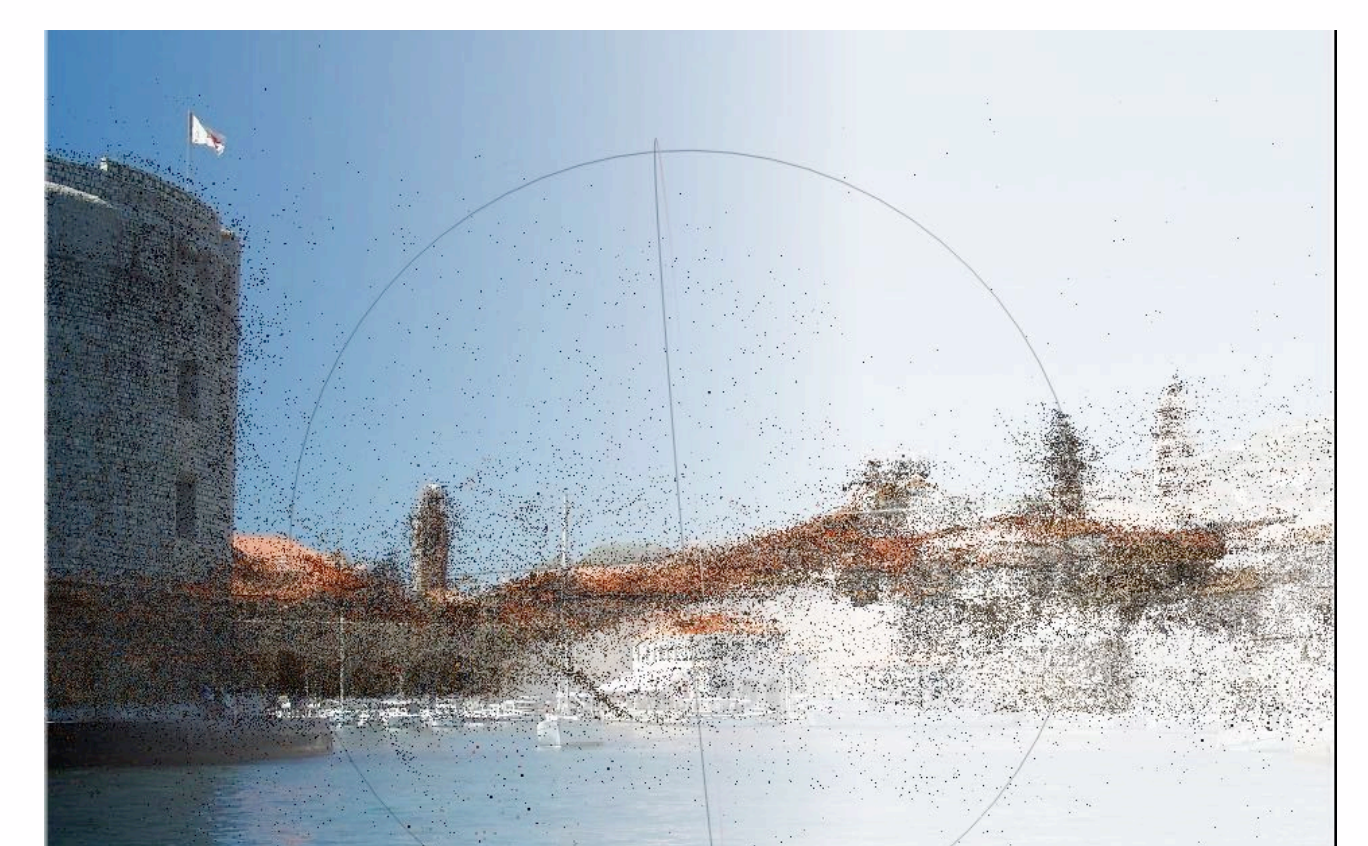
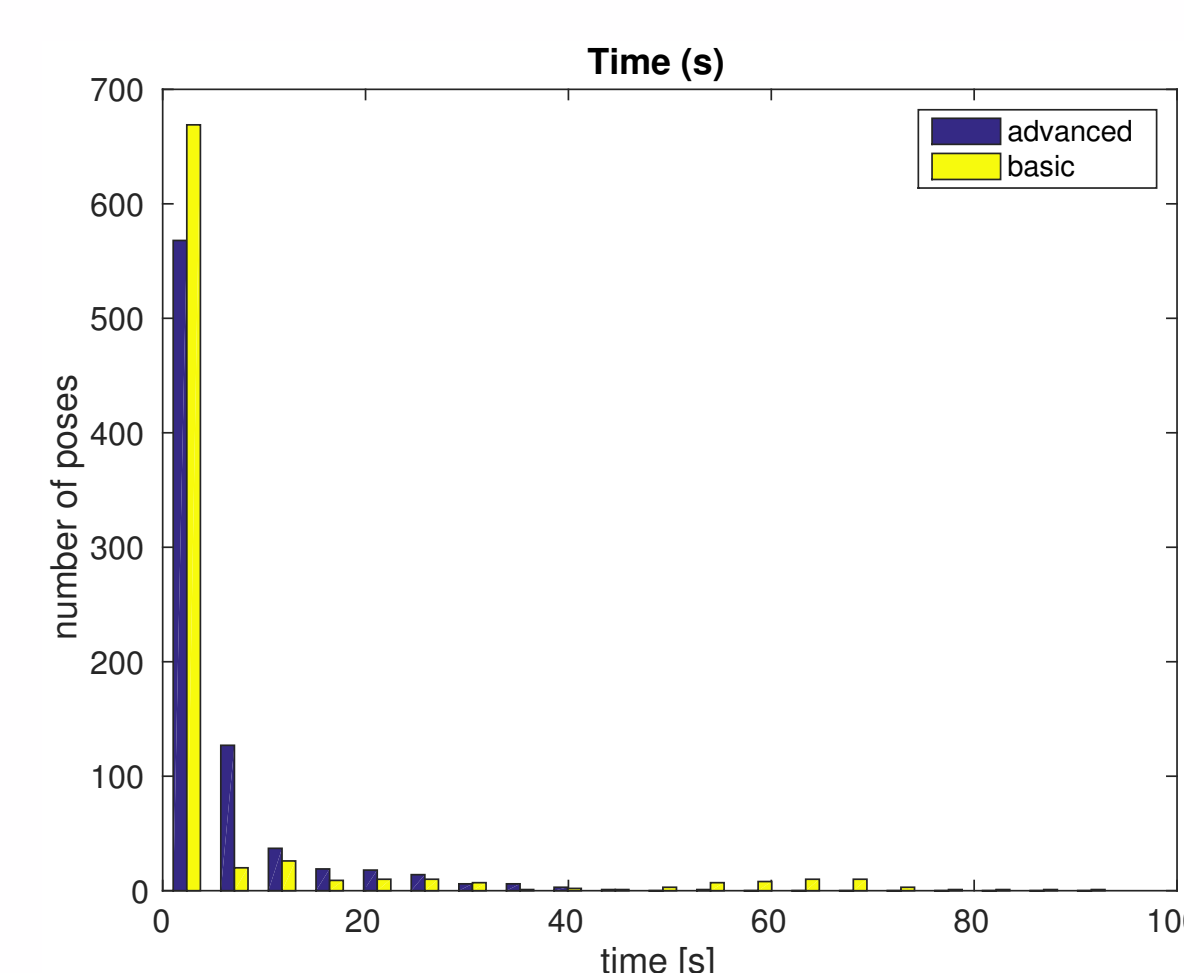


Priority queue

4 Results



	Basic	Advanced
Time [s]	6.4	5.3
Queries with less than 0.5m error [%]	16.5	27.12
Median error [m]	2.42	1.6
Mean error [m]	33.0	32.5



Figures: Correct matches and benchmarking results

5 Conclusion

This method allows to robustly scale the point cloud, including multiple cities and still get good pose estimations, so it represents a really scalable solution.

The advanced method cannot really show good improvements on the Dubrovnik dataset alone, as there is not much to improve on it. The better quality found does not justify the much more complex algorithm, which also has a lot more parameters to tune. The algorithm proposed would be most useful when multiple datasets with similar features (such as Rome and Dubrovnik) are fused together. However, testing this would require to fuse datasets and re-index all matches and cameras. For a world-wide pose estimation, multiple FLANN kd-trees and distribution across clusters seems to be necessary, but not hard to achieve.

6 References

1. M. Bujnak, Z. Kukelova, and T. Pajdla. A general solution to the p4p problem for camera with unknown focal length. CVPR, 2008.
2. L. Kneip, D. Scaramuzza, and R. Siegwart. A novel parametrization of the perspective-three-point problem for a direct computation of absolute camera position and orientation. CVPR, 2011.
3. Y. Li, N. Snavely, D. Huttenlocher, and P. Fua. Worldwide pose estimation using 3d point clouds. ECCV, 2012.
4. T. Sattler, B. Leibe, and L. Kobbelt. Improving image-based localization by active correspondence search. ECCV, 2012.