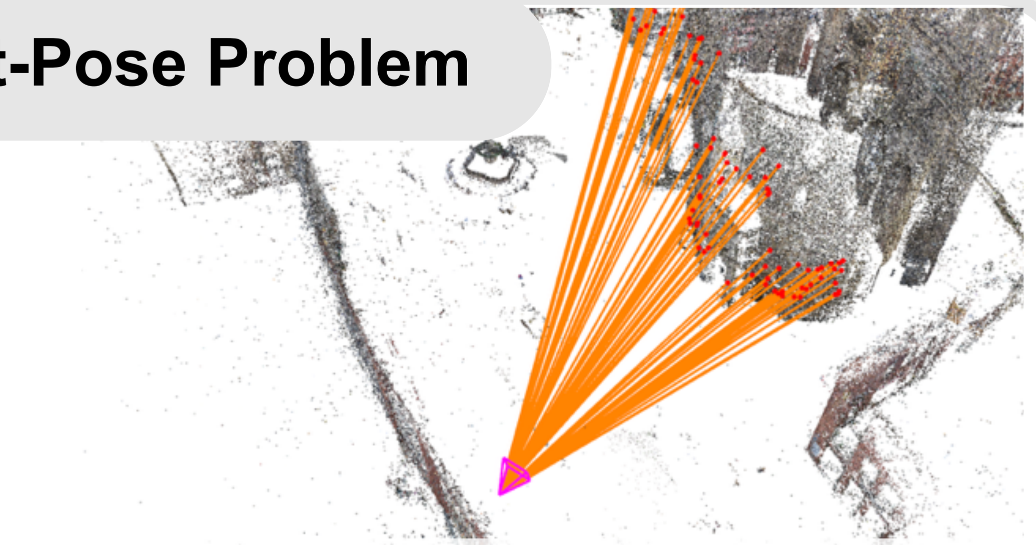
On Sampling Focal Length Values to Solve the Absolute Pose Problem

[1] ETH Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Torsten Sattler[1], Chris Sweeney[2], Marc Pollefeys[1]

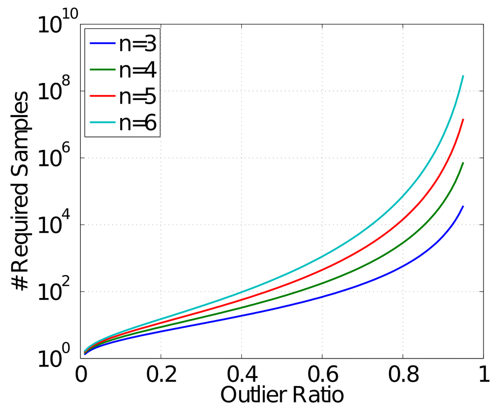
[2] UCSB

1. The *n*-Point-Pose Problem



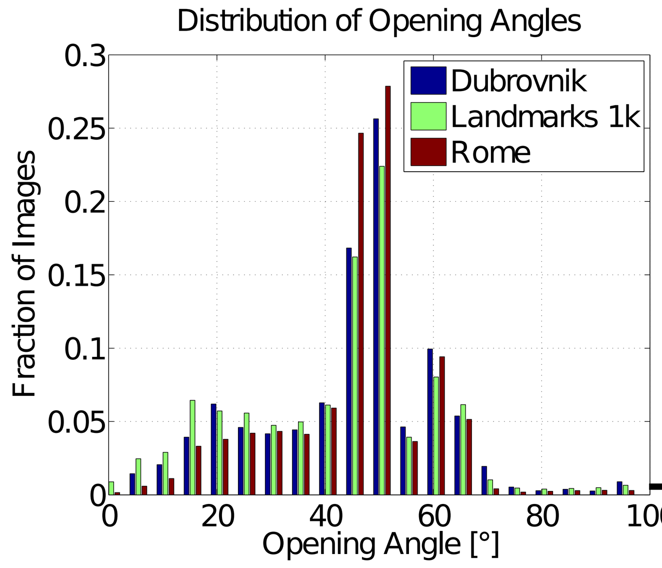
- **Compute camera pose** $[\mathtt{R}|\mathtt{t}]$ from 2D-3D correspondences
- **Estimate focal length** $\mathtt{f}$
- **Applications**: Structure-from-Motion (SfM), image-based localization, …
- **Standard solution**: Minimal solver inside RANSAC-loop

Solver	Estimates	Time	n
P3P [4]	$\mathtt{R}, \mathtt{t}$	2µs	3
P4Pf [1,2]	$\mathtt{R}, \mathtt{t}, \mathtt{f}$	~100µs / 46µs	4
P5Pfr [5]	$\mathtt{R}, \mathtt{t}, \mathtt{f}, \mathtt{r}$	2µs	5

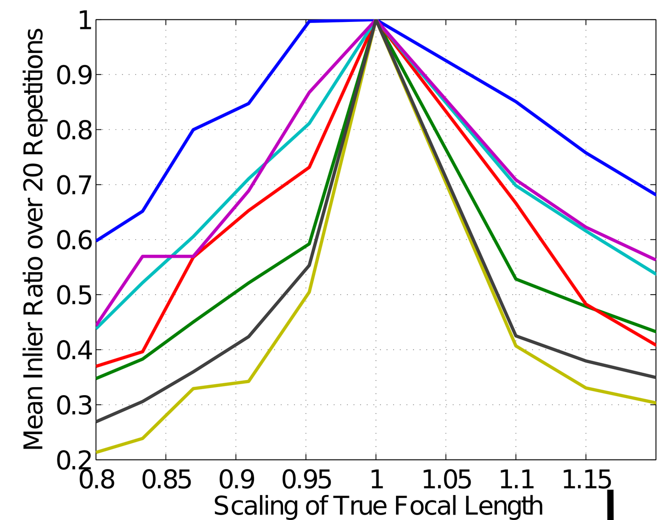


- **Is RANSAC + Minimal Solver the optimal strategy?**
- **Compute the focal length? Or sample it?**
- **Can we do better than brute force search through all focal length values?**

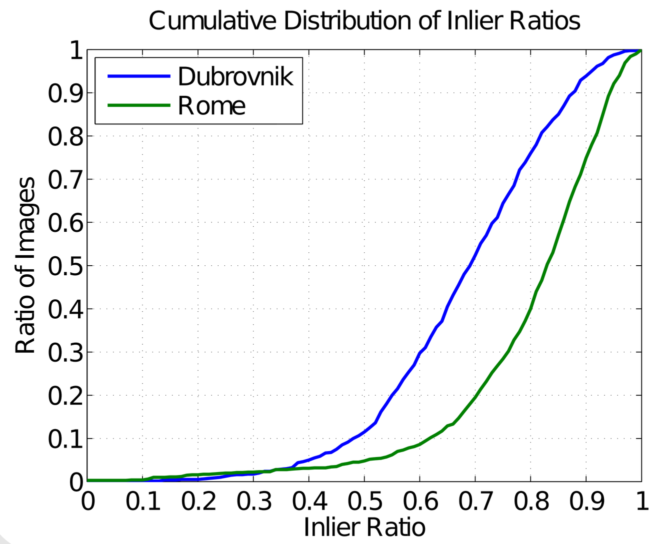
2. Key Observations



- Not all focal length values equally likely
- Prior probabilities from large SfM models
  - Approximation to popularity of cameras
  - Dubrovnik [6]: 6k images
  - Rome [6]: 15k images
  - Landmarks 1k [7]: 205k images

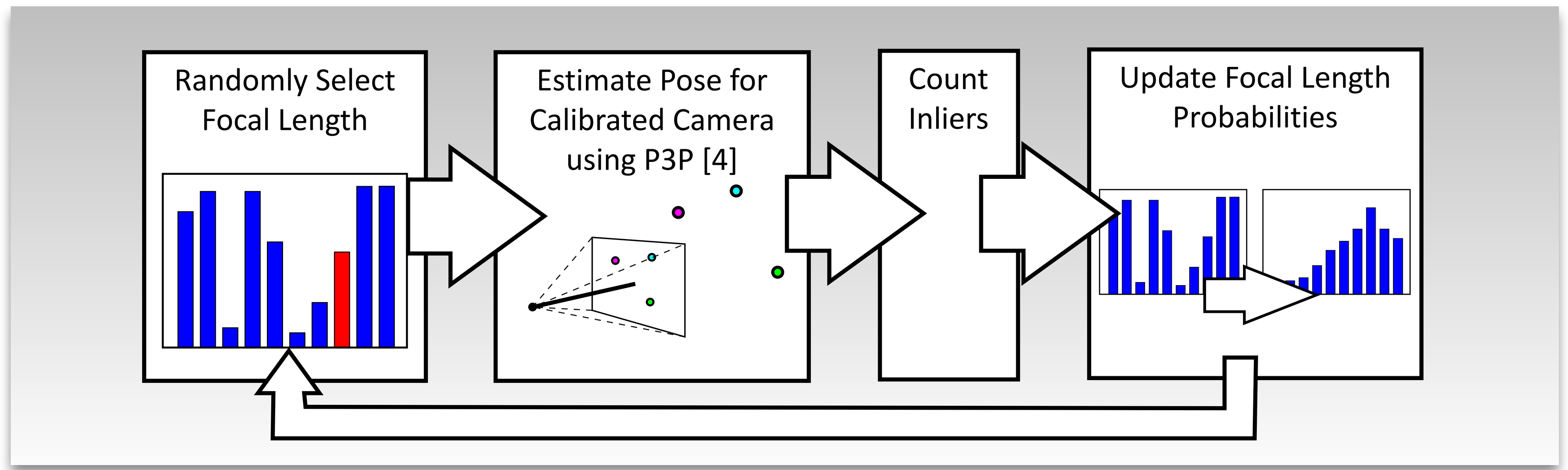


Simplifying assumption:
Inlier ratio decreases monotonically when moving away from correct focal length.



- Inlier ratio only depends on matching algorithm & scene
- Can learn cumulative distribution function (**cdf**) offline

3. P3P(f)-RANSAC



- **Modified RANSAC scheme**: Sample instead of computing focal length in each iteration.
- **Same termination guarantee**: Stop if probability of finding better pose < η
- Probabilistic model:

$$P_{\text{sampling}}(f) = \frac{P_{\text{prior}}(f) \cdot P(\varepsilon(f) > \varepsilon^* \mid f)}{\sum_{f' \in \mathcal{F}} P_{\text{prior}}(f') \cdot P(\varepsilon(f') > \varepsilon^* \mid f')}$$

Probability of finding new best pose for focal length f

- Assumes minimal inlier ratio ε_0 to limit number of RANSAC iterations
- Current best inlier ratio ε^*

- **Update of probability distribution**:
  1. No good model found so far ($\varepsilon^* = \varepsilon_0$): Treat all focal length values independently
  2. Best model found for focal length $\mathtt{f}^*$: Model dependency between focal length values
     - Probability of finding better model for focal length $\mathtt{f}$, used in $\mathtt{k(f)}$ iterations so far:

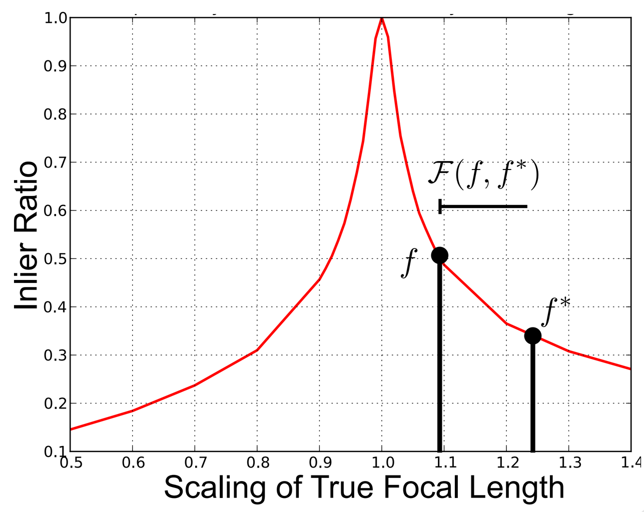
$$P(\varepsilon(f) > \varepsilon_0 \mid f) \leq P(\varepsilon(\mathcal{F}(f, f^*)) > \varepsilon^* \mid f) = \text{cdf}(\max(\varepsilon_{\max}(f), \varepsilon_0)) - \text{cdf}(\varepsilon_0)$$

- Maximal inlier ratio that can be found with probability ≥ η

$$\left(1 - \varepsilon^3\right)^k = \eta \Longrightarrow \varepsilon_{\max}(\mathcal{F}(f, f^*)) = \sqrt[3]{1 - \sqrt[k(\mathcal{F}(f,f^*))]{\eta}}$$

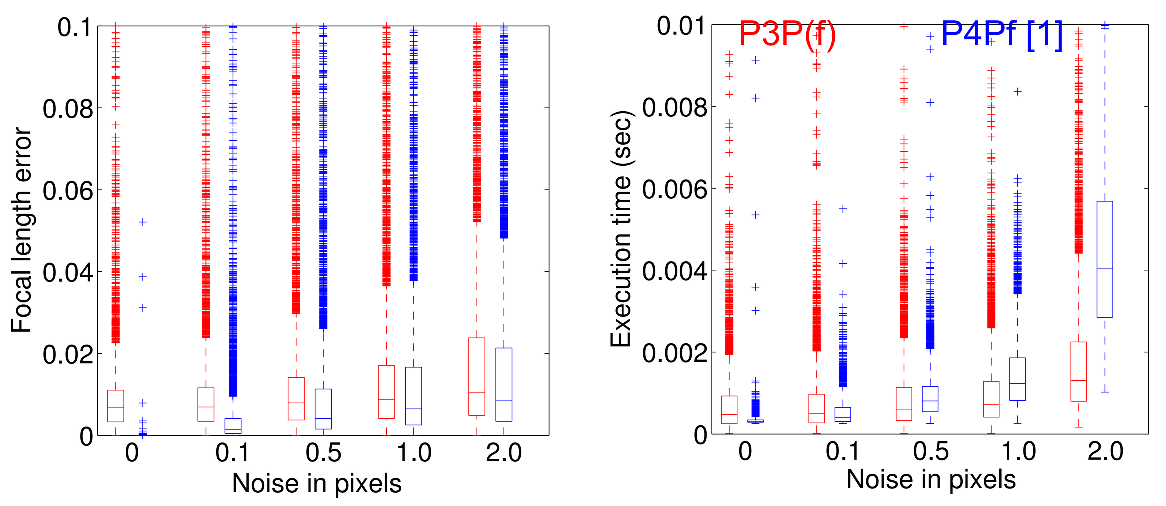
- Number of iterations in which focal length from $\mathcal{F}(f, f^*)$ was used: $k(\mathcal{F}(f, f^*))$

- **Early pose rejection:**
  - Most models will be generated for wrong focal length
  - All-inlier sample + wrong focal length ⇒ low inlier ratio
  - $\mathrm{T}_{1,1}$ test [3] extremely effective:
    - Evaluate pose on all matches only if random match inlier

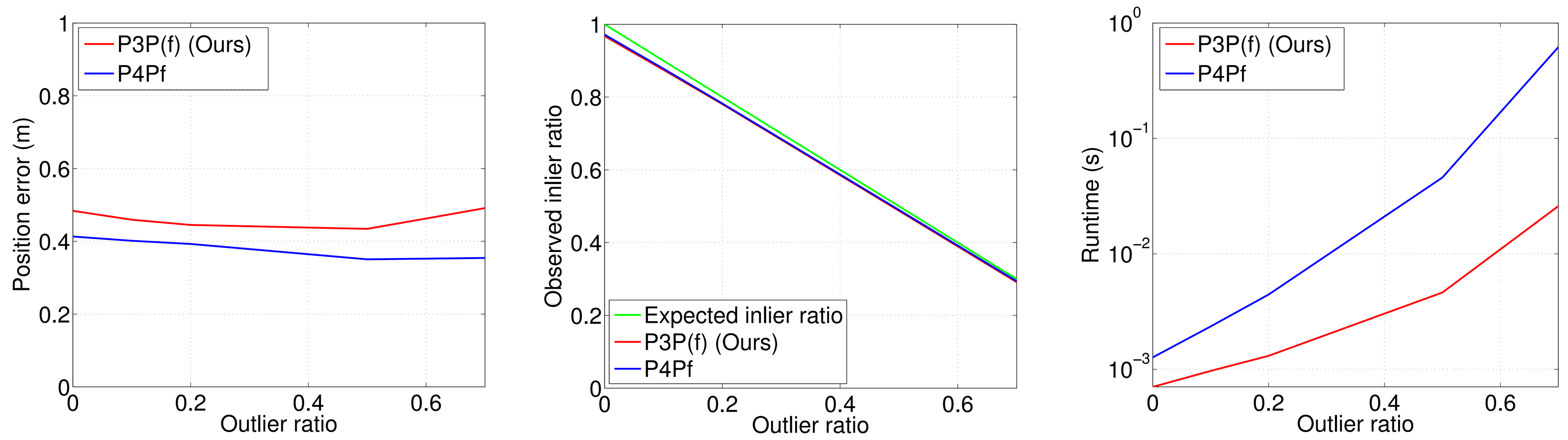


4. Experimental Evaluation

Synthetic data: Reprojected points into images (Dubrovnik), added image noise and outliers



Robustness to image noise



Robustness to outliers: Simplifying assumption holds well in practice!

Real data: Priors from Landmarks 1k, results on Dubrovnik (800 query images) using correspondences from [8]

Solver	# loc. images	Localization Accuracy [m]					Localization Times [ms]			
			Quantiles					Quantiles		
		Mean	25%	50%	75%	90%	Mean	50%	75%	90%
P3P [4] (exact focal)	792	40.3	1.0	7.6	26.4	111.8	**1.21**	**0.20**	**1.00**	3.01
P4Pf [1]	795	38.7	**0.4**	**1.3**	**4.7**	**20.1**	32.09	4.84	10.78	28.73
P5Pfr [5]	**796**	227.2	0.5	2.0	31.3	200.9	6.02	0.54	3.07	16.44
P3P(f) (Ours)	795	**20.8**	**0.4**	1.6	5.4	27.6	1.68	0.68	1.27	**2.72**
P3P(f) uniform prior	795	28.1	0.5	1.7	5.9	24.3	1.89	0.85	1.46	3.08

- Ground truth position given by SfM
- P3P(f) run-times similar to P3P!
- Accuracy similar to P4Pf, but significantly faster
- Improvements come from sampling strategy, not priors

References

[1] M. Bujnak, Z. Kukelova, T. Pajdla. A general solution to the P4P problem for camera with unknown focal length. CVPR 2008.
[2] M. Bujnak, Z. Kukelova, T. Pajdla. Making Minimal Solvers Fast. CVPR 2012.
[3] O. Chum, J. Matas. Randomized RANSAC with T(d,d) test. BMVC 2002.
[4] L. Kneip, D. Scaramuzza, R. Siegwart. A Novel Parameterization of the Perspective-Three-Point Problem for a Direct Computation of Absolute Camera Position and Orientation. CVPR 2011.
[5] Z. Kukelova, M. Bujnak, T. Pajdla. Real-Time Solution to the Absolute Pose Problem with Unknown Radial Distortion and Focal Length. ICCV 2013.
[6] Y. Li, N. Snavely, D. Huttenlocher. Location Recognition using Prioritized Feature Matching. ECCV 2010.
[7] Y. Li, N. Snavely, D. Huttenlocher, P. Fua. Worldwide Pose Estimation Using 3D Point Clouds. ECCV 2012.
[8] T. Sattler, B. Leibe, L. Kobbelt. Improving Image-Based Localization by Active Correspondence Search. ECCV 2012.

Source code will be available at http://people.inf.ethz.ch/sattlert

This work was supported in part by NSF Grant IIS-1219261, NSF Graduate Research Fellowship Grant DGE-1144085, the CTI Switzerland grant #13086.1 PFES-ES 4DSites, and the European Union's Seventh Framework Programme (FP6/2007-2013) under grant #269916 (V-Charge)

V-CHARGE