FACULTY OF MATHEMATICS, PHYSICS AND INFORMATICS
Comenius University
Bratislava

FACULTY OF ELECTRICAL ENGINEERING
CTU IN PRAGUE

Robust Self-calibration of Focal Lengths from the Fundamental Matrix

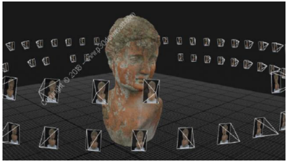
Viktor Kocur, Daniel Kyselica, Zuzana Kúkelová

TERAIS

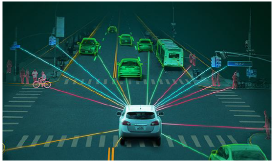
Funded by
the European Union

FUNDED BY
GAČR
CZECH SCIENCE FOUNDATION

3D Vision Applications



3D Reconstruction



Autonomous Driving

Augmented Reality

SLAM and Localization

- calibrate cameras
- estimate camera poses/motion
- triangulate points

Camera Calibration

Offline Calibration

Self-calibration
(autocalibration)

Camera Calibration

Offline Calibration

☺ Can provide complex camera model

☹ Not always possible

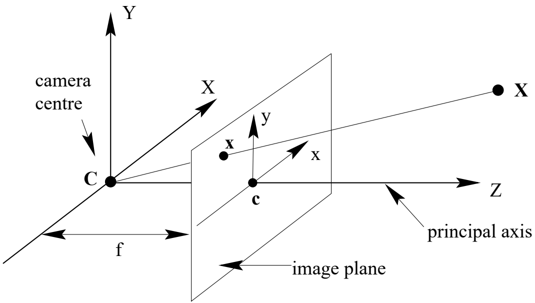
Self-calibration

(autocalibration)

☺ Can be used "in the wild"

☹ Usually simpler camera models

Image adopted from: Zuzana Kukelova et al. “Efficient solution to the epipolar geometry for radially distorted cameras.” *CVPR*. 2015



We want camera intrinsics: focal length, principal point, skew, distortion coeffs...

We can often use additional assumptions: no skew, no distortion, principal point in the image center, square pixels $\rightarrow$ only f remains unknown.

Image adopted from: R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. 2003

Camera Focal Length Estimation

Possible from a single image in some cases:

- Known or learned objects/scene
- Regular geometry observed, e.g.:
  - Calibration patterns
  - Repeating patterns
  - Orthogonal sets of parallel lines

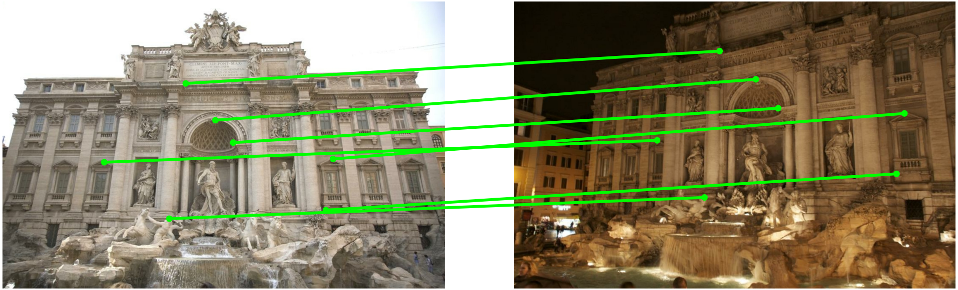
In general we need at least two-views!

Fundamental Matrix

$$\mathbf{F} = \mathbf{K}_2^{-\top}[\boldsymbol{t}]_\times \mathbf{R}\mathbf{K}_1^{-1}$$

$\mathbf{F}$ contains information on camera intrinsics $\mathbf{K}_i = \begin{bmatrix} f_i & 0 & c_{x,i} \\ 0 & f_i & c_{y,i} \\ 0 & 0 & 1 \end{bmatrix}$ and pose $(\mathbf{R}, \boldsymbol{t})$.

Fundamental Matrix



$$\mathbf{F} = \mathbf{K}_2^{-\top} [\boldsymbol{t}]_\times \mathbf{R} \mathbf{K}_1^{-1}$$

We estimate **F** using point correspondences and RANSAC

Bougnoux Formula

We can usually assume that the principal points ($\boldsymbol{c}_1, \boldsymbol{c}_2$) are in the image centers and estimate the focal length of the first camera f_1 from $\mathbf{F}$ using:[3]

$$f_1^2 = -\frac{\boldsymbol{c}_2^\top [\boldsymbol{e}_2]_\times \hat{\mathbf{I}} \mathbf{F} \boldsymbol{c}_1 \boldsymbol{c}_1^\top \mathbf{F}^\top \boldsymbol{c}_2}{\boldsymbol{c}_2 [\boldsymbol{e}_2]_\times \hat{\mathbf{I}} \mathbf{F} \hat{\mathbf{I}} \mathbf{F}^\top \boldsymbol{c}_2} \tag{1}$$

Equivalent formula exists for f_2^2.

Both can also be calculated directly from elements of $\mathbf{F}$.[4]

[3] Sylvain Bougnoux. “From projective to euclidean space under any practical situation, a criticism of self-calibration.” *ICCV*. 1998

[4] Oleh Rybkin. “Robust Focal Length Estimation.” Supervised by: Tomáš Pajdla. Bachelor’s Thesis. Czech Technical University in Prague, 2017

Problems

Singularity: The denominator vanishes when the principal axes of the two cameras are coplanar (they intersect). This makes the formula unstable in many practical situations!

Problems

Singularity: The denominator vanishes when the principal axes of the two cameras are coplanar (they intersect). This makes the formula unstable in many practical situations!

Problems

Singularity: The denominator vanishes when the principal axes of the two cameras are coplanar (they intersect). This makes the formula unstable in many practical situations!

Imaginary f: The formula provides f_i in squared form, thus sometimes yielding imaginary f_i. This can be caused by instability or errors in $\mathbf{F}, \boldsymbol{c}$.

Alternative - Iterative Optimization

Hatley and Silpa-Anan[5] propose to optimize $\mathbf{F}, \boldsymbol{c}_1, \boldsymbol{c}_2$ with LM w.r.t cost:

$$C = w_{\mathbf{F}} C_{\mathbf{F}} + \sum_{i \in \{1,2\}} w_{f_{i,p}} (f_{i,p}^2 - f_i^2)^2 + w_{\boldsymbol{c}_{i,p}} ||\boldsymbol{c}_{i,p} - \boldsymbol{c}_i||^2 + w_p C_p(f_i) \tag{2}$$

- $w_{\mathbf{F}} C_{\mathbf{F}}$: Sampson Error w.r.t. inlier matches
- $w_{f_{i,p}} (f_{i,p}^2 - f_i^2)^2$: Prior for f_i
- $w_{\boldsymbol{c}_{i,p}} ||\boldsymbol{c}_{i,p} - \boldsymbol{c}_i||^2$: Prior for $\boldsymbol{c}_i$
- $w_p C_p(f_i)$: Penalty for low f_i

[5] Richard Hartley and Chanop Silpa-Anan. “Reconstruction from two views using approximate calibration.” *ACCV*. 2002

Alternative - Iterative Optimization

Hatley and Silpa-Anan[5] propose to optimize $\mathbf{F}, \boldsymbol{c}_1, \boldsymbol{c}_2$ with LM w.r.t cost:

$$C = \underbrace{w_{\mathbf{F}} C_{\mathbf{F}}}_{\text{Sampson Error w.r.t. inlier matches}} + \sum_{i \in \{1,2\}} \underbrace{w_{f_{i,p}} (f_{i,p}^2 - f_i^2)^2}_{\text{Prior for } f_i} + \underbrace{w_{\boldsymbol{c}_{i,p}} ||\boldsymbol{c}_{i,p} - \boldsymbol{c}_i||^2}_{\text{Prior for } \boldsymbol{c}_i} + \underbrace{w_p C_p(f_i)}_{\text{Penalty for low } f_i} \tag{2}$$

- ☺ Able to correct for error in $\boldsymbol{c}_1, \boldsymbol{c}_2, \mathbf{F}$
- ☺ Outputs real f_1, f_2

[5] Richard Hartley and Chanop Silpa-Anan. "Reconstruction from two views using approximate calibration." *ACCV*. 2002

Alternative - Iterative Optimization

Hatley and Silpa-Anan[5] propose to optimize $\mathbf{F}, \boldsymbol{c}_1, \boldsymbol{c}_2$ with LM w.r.t cost:

$$C = w_{\mathbf{F}} C_{\mathbf{F}} + \sum_{i \in \{1,2\}} w_{f_{i,p}} (f_{i,p}^2 - f_i^2)^2 + w_{\boldsymbol{c}_{i,p}} ||\boldsymbol{c}_{i,p} - \boldsymbol{c}_i||^2 + w_p C_p(f_i) \quad (2)$$

- $w_{\mathbf{F}} C_{\mathbf{F}}$ → Sampson Error w.r.t. inlier matches
- $w_{f_{i,p}} (f_{i,p}^2 - f_i^2)^2$ → Prior for f_i
- $w_{\boldsymbol{c}_{i,p}} ||\boldsymbol{c}_{i,p} - \boldsymbol{c}_i||^2$ → Prior for $\boldsymbol{c}_i$
- $w_p C_p(f_i)$ → Penalty for low f_i

☺ Able to correct for error in $\boldsymbol{c}_1, \boldsymbol{c}_2, \mathbf{F}$

☺ Outputs real f_1, f_2

☹ f_i still from eq. (1) → singularities

☹ Sampson Error is expensive

☹ Many iterations to converge

☹ Weights need to be tuned

[5] Richard Hartley and Chanop Silpa-Anan. "Reconstruction from two views using approximate calibration." *ACCV*. 2002

Our Contribution

- New method for self-calibration of f_1, f_2 from $\mathbf{F}$
  - Solved as optimization problem using a simpler cost function
  - We propose an efficient iterative algoritm for optimization of the cost function
  - More robust (accurate) than SOTA
  - Significantly faster than other iterative approaches

Our Contribution

- New method for self-calibration of f_1, f_2 from $\mathbf{F}$
  - Solved as optimization problem using a simpler cost function
  - We propose an efficient iterative algoritm for optimization of the cost function
  - More robust (accurate) than SOTA
  - Significantly faster than other iterative approaches
- New RANSAC check $\rightarrow$ further precision improvement and faster RANSAC

Our Contribution

- New method for self-calibration of f_1, f_2 from $\mathbf{F}$
  - Solved as optimization problem using a simpler cost function
  - We propose an efficient iterative algoritm for optimization of the cost function
  - More robust (accurate) than SOTA
  - Significantly faster than other iterative approaches
- New RANSAC check $\rightarrow$ further precision improvement and faster RANSAC
- Both extensively tested on different real-world datasets with different features and feature matchers and RANSAC variants

Estimating f_i - Our Approach

We solve an optimization problem:

$$\min_{f_1,f_2,\mathbf{c}_1,\mathbf{c}_2} \sum_{i=1,2} w_i^f (f_i - f_i^p)^2 + w_i^{\mathbf{c}} \|\mathbf{c}_i - \mathbf{c}_i^p\|^2 \tag{3}$$

[6] Manolis IA Lourakis and Rachid Deriche. “Camera self-calibration using the Kruppa equations and the SVD of the fundamental matrix: The case of varying intrinsic parameters.” PhD thesis. INRIA, 2000

We solve an optimization problem:

$$\min_{f_1, f_2, \mathbf{c}_1, \mathbf{c}_2} \sum_{i=1,2} w_i^f (f_i - f_i^p)^2 + w_i^{\mathbf{c}} \|\mathbf{c}_i - \mathbf{c}_i^p\|^2 \tag{3}$$

$$\text{s.t.} \quad \kappa_1 = \sigma_1(\mathbf{v}_1^\top \omega_1^* \mathbf{v}_1)(\mathbf{u}_1^\top \omega_2^* \mathbf{u}_2) + \sigma_2(\mathbf{v}_1^\top \omega_1^* \mathbf{v}_2)(\mathbf{u}_2^\top \omega_2^* \mathbf{u}_2) = 0 \tag{4}$$

$$\kappa_2 = \sigma_1(\mathbf{v}_1^\top \omega_1^* \mathbf{v}_2)(\mathbf{u}_1^\top \omega_2^* \mathbf{u}_1) + \sigma_2(\mathbf{v}_2^\top \omega_1^* \mathbf{v}_2)(\mathbf{u}_1^\top \omega_2^* \mathbf{u}_2) = 0, \tag{5}$$

where κ_1, κ_2 are two Kruppa equations[6] derived from $\omega_i^* = \mathbf{K}_i \mathbf{K}_i^\top$ and SVD of $\mathbf{F}$.

[6] Manolis IA Lourakis and Rachid Deriche. "Camera self-calibration using the Kruppa equations and the SVD of the fundamental matrix: The case of varying intrinsic parameters." PhD thesis. INRIA, 2000

To solve the problem we use the Lagrangian.

$$L = \sum_{i=1,2} w_i^f (f_i - f_i^p)^2 + w_i^{\mathbf{c}} \|\mathbf{c}_i - \mathbf{c}_i^p\|^2 - 2\lambda_1\kappa_1 - 2\lambda_2\kappa_2 \quad (6)$$

Solving $\nabla L = \mathbf{0}$ using algebraic methods is not feasible - system too complex

[7] Peter Lindstrom. “Triangulation made easy.” *CVPR*. 2010

[8] Viktor Larsson, Kalle Astrom, and Magnus Oskarsson. “Efficient solvers for minimal problems by syzygy-based reduction.” *CVPR*. 2017

Estimating f_i - Our Approach

To solve the problem we use the Lagrangian.

$$L = \sum_{i=1,2} w_i^f (f_i - f_i^p)^2 + w_i^{\mathbf{c}} \|\mathbf{c}_i - \mathbf{c}_i^p\|^2 - 2\lambda_1\kappa_1 - 2\lambda_2\kappa_2 \tag{6}$$

Solving $\nabla L = \mathbf{0}$ using algebraic methods is not feasible - system too complex

We propose a hybrid (iterative + algebraic) approach:[7]

[7] Peter Lindstrom. "Triangulation made easy." *CVPR*. 2010

[8] Viktor Larsson, Kalle Astrom, and Magnus Oskarsson. "Efficient solvers for minimal problems by syzygy-based reduction." *CVPR*. 2017

Estimating f_i - Our Approach

To solve the problem we use the Lagrangian.

$$L = \sum_{i=1,2} w_i^f (f_i - f_i^p)^2 + w_i^{\mathbf{c}} \|\mathbf{c}_i - \mathbf{c}_i^p\|^2 - 2\lambda_1\kappa_1 - 2\lambda_2\kappa_2 \quad (6)$$

Solving $\nabla L = \mathbf{0}$ using algebraic methods is not feasible - system too complex

We propose a hybrid (iterative + algebraic) approach:[7]

- We fix unknowns $(f_i, c_{x,i}, c_{y,i})$ in higher order monomials to values estimated in previous step.

[7] Peter Lindstrom. "Triangulation made easy." *CVPR*. 2010

[8] Viktor Larsson, Kalle Astrom, and Magnus Oskarsson. "Efficient solvers for minimal problems by syzygy-based reduction." *CVPR*. 2017

Estimating f_i - Our Approach

To solve the problem we use the Lagrangian.

$$L = \sum_{i=1,2} w_i^f (f_i - f_i^p)^2 + w_i^{\mathbf{c}} \|\mathbf{c}_i - \mathbf{c}_i^p\|^2 - 2\lambda_1\kappa_1 - 2\lambda_2\kappa_2 \quad (6)$$

Solving $\nabla L = \mathbf{0}$ using algebraic methods is not feasible - system too complex

We propose a hybrid (iterative + algebraic) approach:[7]

- We fix unknowns $(f_i, c_{x,i}, c_{y,i})$ in higher order monomials to values estimated in previous step.
- In each iteration we solve 2 polynomial equations of degree 4 using GB solver.[8]

[7] Peter Lindstrom. “Triangulation made easy.” *CVPR*. 2010

[8] Viktor Larsson, Kalle Astrom, and Magnus Oskarsson. “Efficient solvers for minimal problems by syzygy-based reduction.” *CVPR*. 2017

To solve the problem we use the Lagrangian.

$$L = \sum_{i=1,2} w_i^f (f_i - f_i^p)^2 + w_i^{\mathbf{c}} \|\mathbf{c}_i - \mathbf{c}_i^p\|^2 - 2\lambda_1\kappa_1 - 2\lambda_2\kappa_2 \quad (6)$$

Solving $\nabla L = \mathbf{0}$ using algebraic methods is not feasible - system too complex

We propose a hybrid (iterative + algebraic) approach:[7]

- We fix unknowns $(f_i, c_{x,i}, c_{y,i})$ in higher order monomials to values estimated in previous step.
- In each iteration we solve 2 polynomial equations of degree 4 using GB solver.[8]
- We iterate until the relative change is small or we hit 50 iterations.

[7] Peter Lindstrom. "Triangulation made easy." *CVPR*. 2010

[8] Viktor Larsson, Kalle Astrom, and Magnus Oskarsson. "Efficient solvers for minimal problems by syzygy-based reduction." *CVPR*. 2017

Estimating f_i - Our Approach

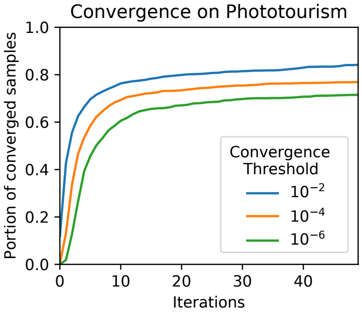
To solve the problem we use the Lagrangian.

$$L = \sum_{i=1,2} w_i^f (f_i - f_i^p)^2 + w_i^{\mathbf{c}} \|\mathbf{c}_i - \mathbf{c}_i^p\|^2 - 2\lambda_1\kappa_1 - 2\lambda_2\kappa_2 \quad (6)$$

Solving $\nabla L = \mathbf{0}$ using algebraic methods is not feasible - system too complex

We propose a hybrid (iterative + algebraic) approach:[7]

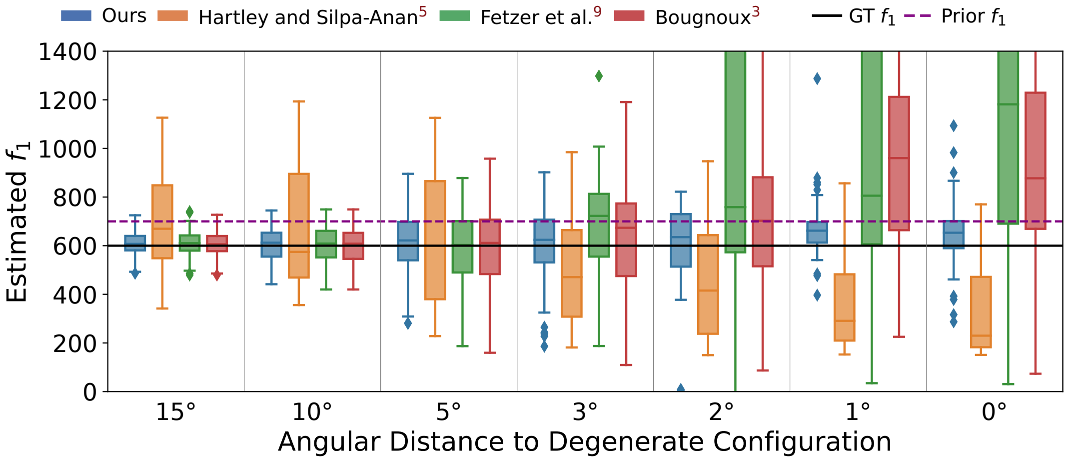
- We fix unknowns $(f_i, c_{x,i}, c_{y,i})$ in higher order monomials to values estimated in previous step.
- In each iteration we solve 2 polynomial equations of degree 4 using GB solver.[8]
- We iterate until the relative change is small or we hit 50 iterations.



[7] Peter Lindstrom. "Triangulation made easy." *CVPR*. 2010

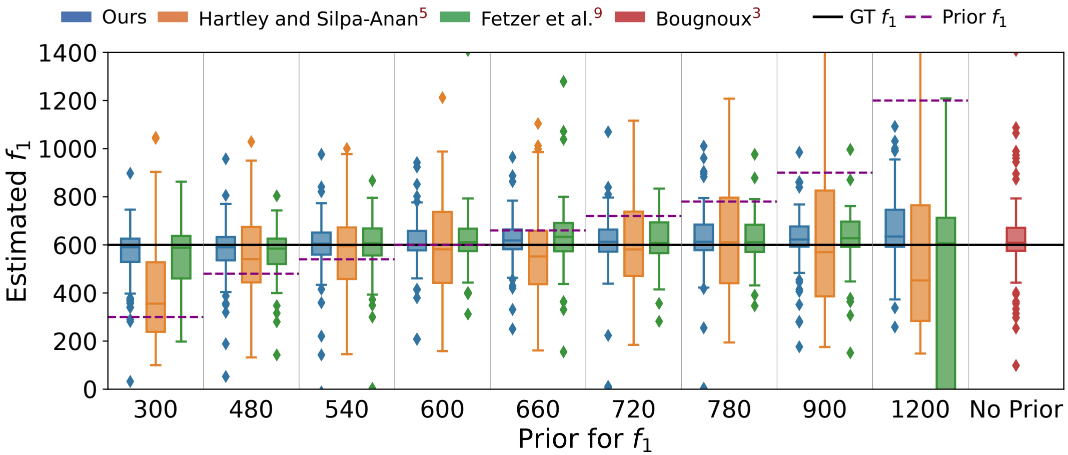
[8] Viktor Larsson, Kalle Astrom, and Magnus Oskarsson. "Efficient solvers for minimal problems by syzygy-based reduction." *CVPR*. 2017

Synth Data - Performance in Degenerate Configurations



[9] Torben Fetzer, Gerd Reis, and Didier Stricker. "Stable intrinsic auto-calibration from fundamental matrices of devices with uncorrelated camera parameters." *CVPR*. 2020

Synth Data - Priors



Our method works with inaccurate priors!

Real Focal Length Checking (RFC)

F with imaginary f_i from the Bougnoux formula usually lead to fewer inliers than **F** with real f_i.

[4] Oleh Rybkin. “Robust Focal Length Estimation.” Supervised by: Tomáš Pajdla. Bachelor’s Thesis. Czech Technical University in Prague, 2017

[10] Yuhe Jin et al. “Image matching across wide baselines: From paper to practice.” *IJCV* (2021)

[11] Jiaming Sun et al. “LoFTR: Detector-free local feature matching with transformers.” *CVPR*. 2021

Real Focal Length Checking (RFC)

F with imaginary f_i from the Bougnoux formula usually lead to fewer inliers than **F** with real f_i.

We use this with an efficient formula[4] to reject models in RANSAC:

- faster - not scoring bad models
- better accuracy

[4] Oleh Rybkin. "Robust Focal Length Estimation." Supervised by: Tomáš Pajdla. Bachelor's Thesis. Czech Technical University in Prague, 2017

[10] Yuhe Jin et al. "Image matching across wide baselines: From paper to practice." *IJCV* (2021)

[11] Jiaming Sun et al. "LoFTR: Detector-free local feature matching with transformers." *CVPR*. 2021

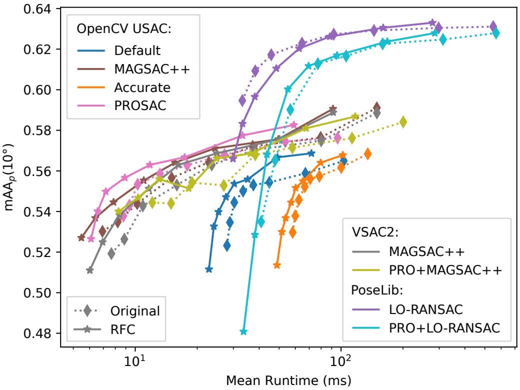
Real Focal Length Checking (RFC)

F with imaginary f_i from the Bougnoux formula usually lead to fewer inliers than **F** with real f_i.

We use this with an efficient formula[4] to reject models in RANSAC:

- faster - not scoring bad models
- better accuracy

Results on real data[10] with LoFTR[11] →



[4] Oleh Rybkin. “Robust Focal Length Estimation.” Supervised by: Tomáš Pajdla. Bachelor’s Thesis. Czech Technical University in Prague, 2017

[10] Yuhe Jin et al. “Image matching across wide baselines: From paper to practice.” *IJCV* (2021)

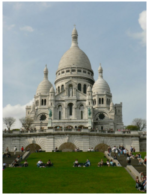
[11] Jiaming Sun et al. “LoFTR: Detector-free local feature matching with transformers.” *CVPR*. 2021

Real Data Evaluation

Evaluated on two large-scale datasets with COLMAP GT:

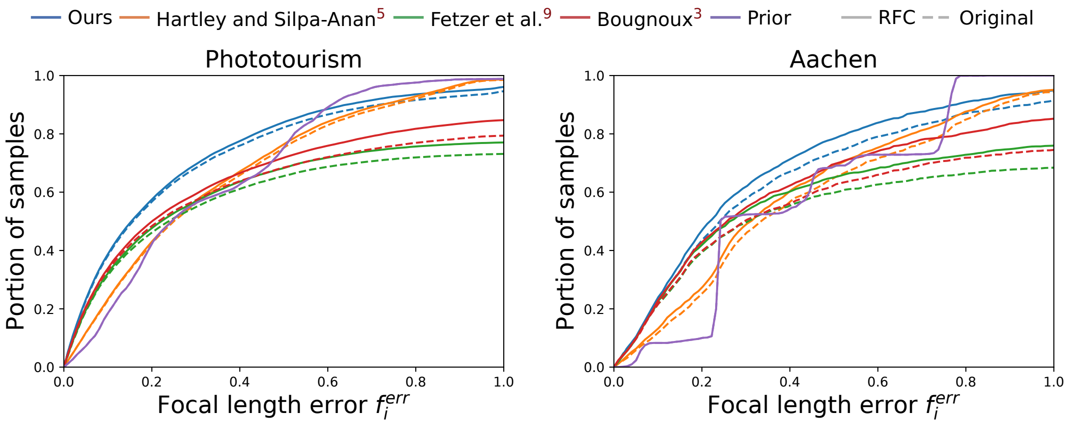
- Phototourism[10]
  - 12 evaluation scenes of historical landmarks
  - 1000 image pairs per scene
- Aachen Day-Night v1.1[12]
  - 5000 image pairs

We use COLMAP priors: $f_i^p = 1.2 \max(width_i, height_i)$



[12] Zichao Zhang, Torsten Sattler, and Davide Scaramuzza. "Reference pose generation for long-term visual localization via learned features and view synthesis." *IJCV* (2021)

f_i Estimation Accuracy



Our iterative method and RFC beat SOTA in terms of $f_i^{err} = \frac{|f_i^{est} - f_i^{gt}|}{f_i^{gt}}$.

F estimated using Magsac++[13] + LoFTR[11] matches - more experiments in paper.

[13] Daniel Barath et al. “MAGSAC++, a fast, reliable and accurate robust estimator.” *CVPR*. 2020

Pose Estimation Accuracy

		Phototourism[10]			Aachen Day-Night v1.1 [12]		
Method	RFC	Median p_{err}	mAA_p 10°	mAA_p 20°	Median p_{err}	mAA_p 10°	mAA_p 20°
Ours		6.43°	40.48	56.01	9.52°	27.62	46.28
Ours	✓	**6.29°**	**41.03**	**56.78**	**8.78°**	**29.29**	**48.13**
Hartley[5]		9.19°	30.15	46.94	12.19°	21.10	38.74
Hartley[5]	✓	9.00°	30.61	47.69	11.37°	22.12	40.61
Fetzer[9]		9.40°	33.00	47.25	12.33°	24.34	39.69
Fetzer[9]	✓	8.94°	33.64	48.51	10.66°	25.62	42.16
Bougnoux[3]		7.55°	37.17	52.70	10.25°	26.73	45.07
Bougnoux[3]	✓	7.39°	37.57	53.21	9.69°	27.25	45.61
Prior		11.17°	22.63	41.98	13.00°	17.06	37.27
Prior	✓	11.05°	22.73	42.31	12.73°	17.68	38.05

Pose error $p_{err} = \max\left(\angle\left(\mathbf{R}^{est}, \mathbf{R}^{gt}\right), \angle\left(\boldsymbol{t}^{est}, \boldsymbol{t}^{gt}\right)\right)$ and mean average accuracy (mAA_p) evaluation shows that our method leads to better poses.

Summary

- Our proposed method for robust self-calibration of f from **F** beats SOTA in terms of accuracy and is faster than previous iterative approaches.
- If you plan to use images with unknown intrinsics (crowdsourced, changing f of a single camera) consider using our method.
- If you need to estimate **F** with RANSAC consider using RFC.
- More info (method details, experiments - case of $f_1 = f_2$) in the paper.

Code available on Github.

Thank you for your attention!