# Supplementary Material for Camera Pose Estimation using Implicit Distortion Models

Linfei Pan
ETH Zurich
linpan@student.ethz.ch

Marc Pollefeys
ETH Zurich, Microsoft
marc.pollefeys@inf.ethz.ch

Viktor Larsson
Lund University
viktor.larsson@math.lth.se

## Abstract

## 1. Overview of Supplementary Material

The supplementary material contains the following:

- Implementation details (Section 2)
- Detailed results from calibration experiment (Table 4)
- Detailed results for the BA experiment (Section 3)
- Evaluation with GT principal point (Section 4)
- Dataset details (Table 1)
- Additional qualitative results (Figure 4)

## 2. Implementation Details

**Neighborhood selection.** To construct the residuals in the regularizer, we select the $k = 2m$ closest neighbours symmetrically around each point. To avoid the instability from the denominator vanishing, a minimum distance $\delta r_{min}$ to the closest neighbor is enforced. Further, since the locally linear assumption fails when the distance between two neighbors is too large, $\delta r_{max}$ is set to ensure the distance of that falls below a threshold.

**Parameters.** In our experiments, $\delta r_{min}$ is set to be 1 px, while $\delta r_{max}$ is set to be 100 px. The number of neighbors is chosen to be 4 since it provides a good trade-off between redundancy in the fitting and the sparsity pattern. This is evaluated in the ablation study in the main paper.

**Outlier filtering.** Given an initial pose estimate we propose to filter outliers in a similar fashion as in [1], using a sliding median filter on the computed $f_i$. Points where $f_i$ deviate significantly from the median of the local neighbourhood ($n = 4$) are removed. The threshold is set as $k = 3 \times 1.4826$ times the median error, which correspond to the $99.7\%$ confidence interval assuming the deviations are normally distributed.

## 3. Evaluation of Bundle Adjustment

The proposed bundle adjustment method optimizes a surrogate problem in each iteration to enable use of the Schur complement trick. However, this inner optimization problem does not minimize the true cost. Figure 1 and Figure 2 show the cost across iterations for bundle adjustment with and without groundtruth principal point. One can observe spikes and energy steps due to the imperfection of the approximation, but still, the surrogate problem is a good approximation and the cost for outer iterations converges almost monotonically as shown in the plots.

## 4. Impact of Estimated Principal Point

For all the experiments in the paper we estimate the principal point by first assuming it is in the center of the image. The principal point is then refined by minimizing the radial reprojection error (see Section 4 in the main paper). For comparison we here show the results using the ground truth principal point (as used in the reference SfM model). See Table 2 and Table 3. Experiments for optimization with estimated principal point are also listed for reference. Similarly, the proposed method consistently outperforms the the Camposeco et al. [1] when ground truth principal point is used. The improvement for estimation of pose for dataset Kazan and Doge Palace using ground truth principal point is significant, which is due to the poor estimation of principal point. In Figure 1 we show the cost across iterations for bundle adjustment using ground truth principal point, and that with estimated principal point can be found in Figure 2. Cumulative reprojection error plots for bundle adjustment with / without groundtruth principal points can be found in Figure 3. Generally, optimizations with groundtruth principal point give better error, and the difference is more significant when the estimation of principal point is less accurate. Empirically we found that the principal point refinement works slightly better if points close to the center are downweighted in the optimization (as the radial reprojection error is very sensitive to principal point shifts for these correspondences).

| | | |
|---|---|---|
| Checkerboard calib. images | [3] | github.com/ylochman/babelcalib |
| Doge Palace and Kazan | [4] | www.maths.lth.se/matematiklth/personal/calle/dataset/dataset.html |
| Grossmunster and Kirchenge | [2] | github.com/vlarsson/radialsfm |
| Aachen Day-Night | [5, 6] | https://www.visuallocalization.net/datasets/ |
| InLoc | [7] | https://www.visuallocalization.net/datasets/ |

Table 1. Datasets used in the experimental evaluation and where to obtain them.

| | | | Kirchenge [2] (369) | | Grossmunster [2] (373) | | Kazan [4] (282) | | Doge Palace [4] (241) | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | $\epsilon_{rot}$ (deg.) | $\epsilon_{pos}$ (cm.) | $\epsilon_{rot}$ (deg.) | $\epsilon_{pos}$ (cm.) | $\epsilon_{rot}$ (deg.) | $\epsilon_{pos}$ (cm.) | $\epsilon_{rot}$ (deg.) | $\epsilon_{pos}$ (cm.) |
| Single image | Proposed | w/ | **0.014** | **0.4** | **0.014** | **0.4** | **0.005** | **1.9** | **0.010** | **2.4** |
| | | w/o | 0.023 | 0.6 | 0.038 | 0.9 | 0.386 | 8.7 | 0.338 | 11.7 |
| | Camposeco et al. [1] | w/ | 0.024 | 0.8 | 0.031 | 3.7 | 0.007 | 2.7 | 0.016 | 91.1 |
| | | w/o | 0.027 | 0.8 | 0.044 | 3.8 | 0.403 | 19.8 | 0.364 | 121.2 |
| Multiple images | Proposed | w/ | **0.012** | **0.3** | **0.016** | **0.3** | **0.004** | **0.4** | **0.006** | **0.4** |
| | | w/o | 0.020 | 0.4 | 0.036 | 0.5 | 0.379 | 3.4 | 0.328 | 3.4 |
| | Camposeco et al. [1] | w/ | 0.019 | 0.4 | 0.027 | 0.7 | 0.007 | 0.6 | 0.013 | 0.7 |
| | | w/o | 0.027 | 0.8 | 0.044 | 3.8 | 0.403 | 19.8 | 0.364 | 121.2 |

Table 2. Average rotation error (in degree) and camera position error (in centimeters) with COLMAP reconstruction result as pseudo groundtruth and optimized with / without groundtruth principal point. The number of images for each dataset is shown in the bracket. Single image optimization and multiple images optimization are presented separately and both compared with method proposed in [1]. Error for principal point estimation for Kirchenge, Grossmunster, Kazan and Doge Palace in px are 0.830, 1.361, 16.456 and 15.377 respectively.

## References


[1] Federico Camposeco, Torsten Sattler, and Marc Pollefeys. Non-parametric structure-based calibration of radially symmetric cameras. In *International Conference on Computer Vision (ICCV)*, 2015. 1, 2, 6

[2] Viktor Larsson, Nicolas Zobernig, Kasim Taskin, and Marc Pollefeys. Calibration-free structure-from-motion with calibrated radial trifocal tensors. In *European Conference on Computer Vision (ECCV)*, 2020. 2, 3, 4

[3] Yaroslava Lochman, Kostiantyn Liepieshov, Jianhui Chen, Michal Perdoch, Christopher Zach, and James Pritts. Babelcalib: A universal approach to calibrating central cameras. In *International Conference on Computer Vision (ICCV)*, 2021. 2, 6

[4] Carl Olsson and Olof Enqvist. Stable structure from motion for unordered image collections. In *Scandinavian Conference on Image Analysis (SCIA)*, 2011. 2, 3, 4

[5] Torsten Sattler, Will Maddern, Carl Toft, Akihiko Torii, Lars Hammarstrand, Erik Stenborg, Daniel Safari, Masatoshi Okutomi, Marc Pollefeys, Josef Sivic, Fredrik Kahl, and Tomas Pajdla. Benchmarking 6dof outdoor visual localization in changing conditions, 2018. 2

[6] Torsten Sattler, Tobias Weyand, Bastian Leibe, and Leif Kobbelt. Image retrieval for image-based localization revisited. In *British Machine Vision Conference (BMVC)*, 2012. 2

[7] Hajime Taira, Masatoshi Okutomi, Torsten Sattler, Mircea Cimpoi, Marc Pollefeys, Josef Sivic, Tomas Pajdla, and Akihiko Torii. Inloc: Indoor visual localization with dense matching and view synthesis, 2018. 2

| | | Upgraded (w/o) | | Upgraded (w/) | | Bundle Adjustment (w/o) | | Bundle Adjustment (w/) | |
|---|---|---|---|---|---|---|---|---|---|
| | | mean | median | mean | median | mean | median | mean | median |
| Kirchenge | $\epsilon_{rot}$ (deg.) | 1.720 | 0.398 | 1.449 | 0.169 | 1.726 | 0.403 | **1.361** | **0.084** |
| | $\epsilon_{pos}$ (m.) | 0.414 | 0.028 | 0.424 | 0.028 | **0.401** | 0.026 | 0.414 | **0.025** |
| | $\epsilon_{proj}^{GT}$ (px) | 2.199 | 1.630 | 1.999 | 1.728 | 1.469 | 1.015 | **1.088** | **0.857** |
| | $\epsilon_{proj}^{est}$ (px) | 1.633 | 0.955 | 1.435 | 1.071 | 1.294 | 0.728 | **0.974** | **0.712** |
| Grossmunster | $\epsilon_{rot}$ (deg.) | 2.322 | 0.643 | 1.993 | 0.284 | 2.291 | 0.716 | **1.782** | **0.111** |
| | $\epsilon_{pos}$ (m.) | 0.990 | 0.129 | 1.008 | 0.129 | 0.939 | 0.093 | **0.929** | **0.082** |
| | $\epsilon_{proj}^{GT}$ (px) | 2.866 | 2.110 | 3.223 | 2.563 | 1.631 | 1.325 | **1.227** | **0.935** |
| | $\epsilon_{proj}^{est}$ (px) | 2.866 | 2.110 | 3.015 | 2.356 | 1.516 | 1.075 | **1.191** | **0.869** |
| Kazan | $\epsilon_{rot}$ (deg.) | 0.463 | 0.463 | 0.047 | 0.047 | 0.403 | 0.396 | **0.041** | **0.042** |
| | $\epsilon_{pos}$ (m.) | 0.053 | 0.048 | 0.021 | 0.018 | 0.152 | 0.112 | **0.015** | **0.014** |
| | $\epsilon_{proj}^{GT}$ (px) | 1.359 | 1.084 | 0.923 | 0.750 | 0.834 | **0.682** | **0.827** | 0.699 |
| | $\epsilon_{proj}^{est}$ (px) | 0.957 | 0.720 | 0.708 | 0.487 | 0.658 | 0.488 | **0.569** | **0.401** |
| Doge Palace | $\epsilon_{rot}$ (deg.) | 0.401 | 0.393 | 0.165 | 0.154 | 0.324 | 0.360 | **0.021** | **0.019** |
| | $\epsilon_{pos}$ (m.) | 0.040 | 0.033 | 0.022 | 0.013 | 0.107 | 0.067 | **0.007** | **0.005** |
| | $\epsilon_{proj}^{GT}$ (px) | 0.914 | 0.710 | 0.776 | 0.589 | 1.091 | 0.981 | **0.681** | **0.535** |
| | $\epsilon_{proj}^{est}$ (px) | 0.893 | 0.690 | 0.647 | 0.436 | 0.619 | 0.453 | **0.558** | **0.389** |

Table 3. BA with implicit distortion model with and without the GT principal point. Table shows the rotation (deg.) and position (m) errors. We also report reprojection errors, both with the GT calib. and estimated. The estimation error for the principal point for Kirchenge, Grossmunster, Kazan and Doge Palace are 8.504, 17.268, 21.522 and 17.487 respectively. Note that here the principal point estimate is taken from the framework of Larsson et al. [2] which does not downweight residuals close to the image center.

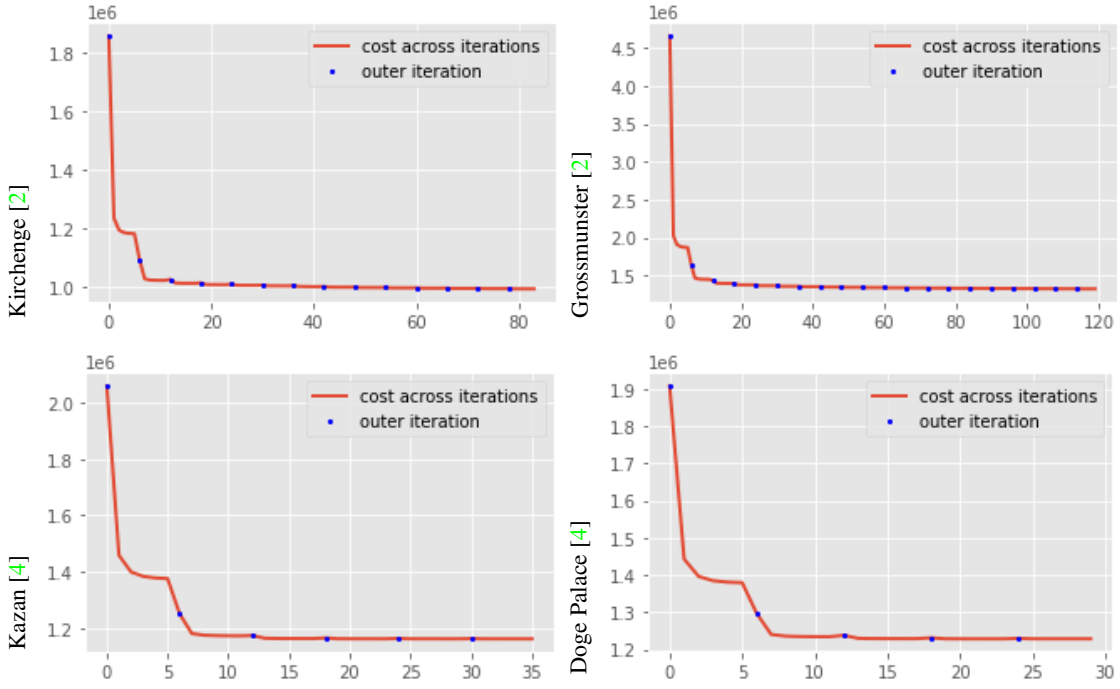


Figure 1. Cost across iterations of optimization for bundle adjustment with groundtruth principal point. Optimization stops when the decrease in cost is less than 0.1%. Start of the outer iterations are marked with blue dots.

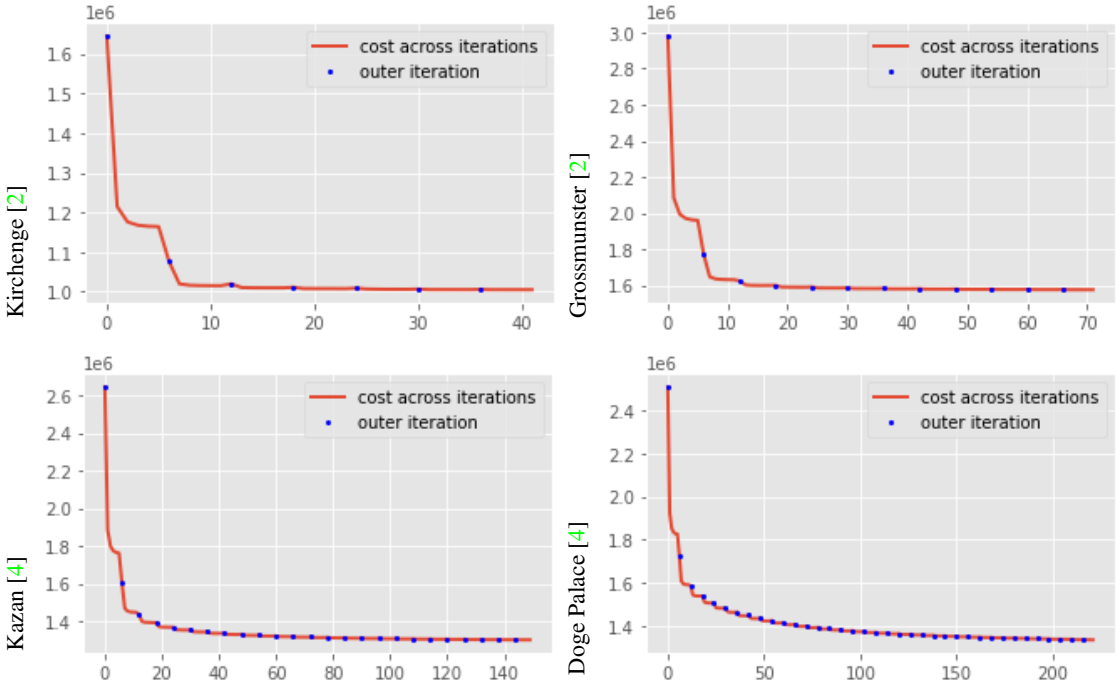


Figure 2. Cost across iterations of optimization for bundle adjustment without groundtruth principal point. Optimization stops when the decrease in cost is less than 0.1%. Start of the outer iterations are marked with blue dots.

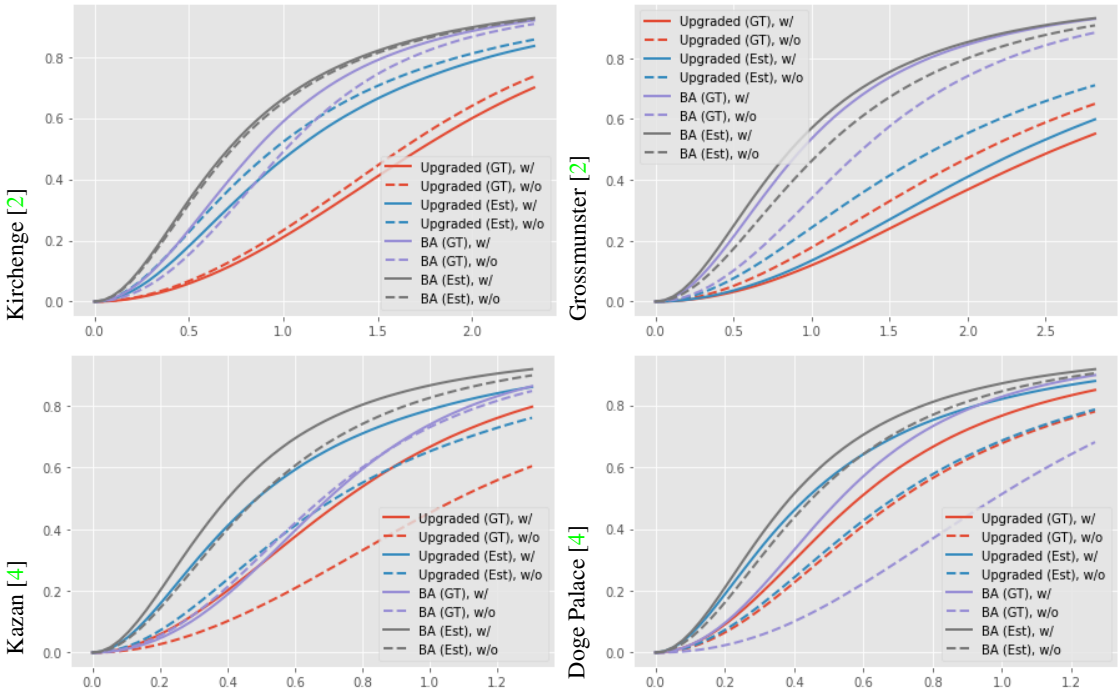


Figure 3. Bundle adjustment with implicit distortion optimized with / without ground truth principal point. Cumulative reprojection errors using GT calib and estimated. Bundle adjustment using groundtruth principal point in general gives better reprojection error. The estimation error for the principal point for Kirchenge, Grossmunster, Kazan and Doge Palace are 8.504, 17.268, 21.522 and 17.487 respectively. Note that here the principal point estimate is taken from the framework of Larsson et al. [2] which does not downweight residuals close to the image center.

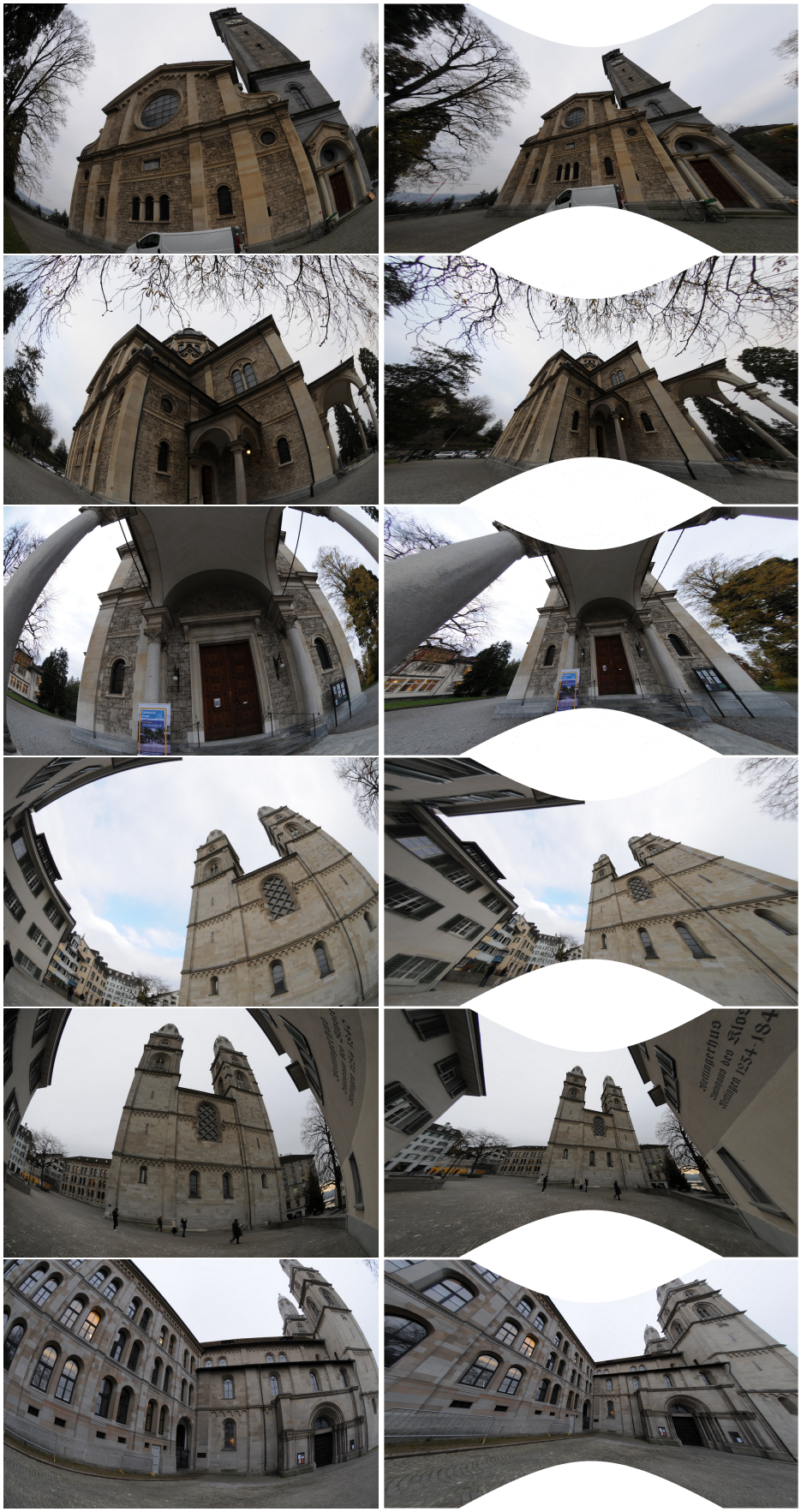

Figure 4. Qualitative results of implicit self-calibration. The images were not seen during calibration showing that our calibration does not overfit to the calibration data.

| | | Training images | | | | | | Test images | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Proposed | | | Camposeco et al. [1] | | | [3] | | Proposed | | Camposeco et al. [1] | |
| Name | $N_{img}$ | $\epsilon_{rot}$ | $\epsilon_{pos}$ | $< 1°, 1\%$ | $\epsilon_{rot}$ | $\epsilon_{pos}$ | $< 1°, 1\%$ | $\epsilon_{rms}^{BC}$ | $\epsilon_{pp}$ | $\epsilon_{rms}$ | $<$ 1px | $\epsilon_{rms}$ | $<$ 1px |
| **Kalibr** | | | | | | | | | | | | | |
| BF2M2020S23 | 35+15 | 0.28 | 0.28 | 33/35 | 0.21 | 0.86 | 28/35 | 0.18 | 1.57 | 0.24 | 15/15 | 0.51 | 14/15 |
| BF5M13720 | 35+15 | 0.12 | 0.09 | 35/35 | 0.25 | 1.92 | 31/35 | 0.18 | 0.45 | 0.20 | 15/15 | 0.66 | 14/15 |
| BM2820 | 35+15 | 0.15 | 0.11 | 35/35 | 0.18 | 0.55 | 33/35 | 0.18 | 1.27 | 0.21 | 15/15 | 0.41 | 15/15 |
| BM4018S118 | 35+15 | 0.08 | 0.14 | 35/35 | 0.33 | 0.93 | 24/35 | 0.24 | 0.08 | 0.27 | 15/15 | 0.62 | 15/15 |
| ENTANIYA | 35+15 | 0.19 | 0.13 | 35/35 | 0.40 | 0.40 | 30/35 | 0.42 | 0.99 | 0.52 | 15/15 | 1.00 | 12/15 |
| EUROC | 35+15 | 0.25 | 0.47 | 34/35 | 0.63 | 1.39 | 20/35 | 0.13 | 1.51 | 0.21 | 15/15 | 0.48 | 15/15 |
| GOPRO | 35+15 | 0.11 | 0.11 | 35/35 | 0.17 | 0.39 | 34/35 | 0.16 | 0.53 | 0.54 | 13/15 | 0.93 | 13/15 |
| TUMVI | 35+15 | 0.20 | 0.12 | 35/35 | 0.29 | 0.48 | 31/35 | 0.16 | 0.68 | 0.19 | 15/15 | 0.30 | 15/15 |
| **OCamCalib** | | | | | | | | | | | | | |
| Fisheye1 | 10+5 | 0.21 | 0.13 | 10/10 | 0.70 | 0.43 | 8/10 | 0.59 | 0.97 | 0.60 | 5/5 | 1.12 | 1/5 |
| Fisheye190deg | 5+3 | 0.14 | 0.16 | 5/5 | 0.98 | 0.96 | 2/5 | 0.63 | 0.41 | 0.63 | 3/3 | 0.93 | 2/3 |
| Fisheye2 | 11+5 | 0.19 | 0.20 | 11/11 | 0.35 | 0.38 | 9/11 | 0.46 | 0.04 | 0.46 | 5/5 | 0.57 | 5/5 |
| GOPR | 7+4 | 0.14 | 0.15 | 7/7 | 0.16 | 0.20 | 7/7 | 1.23 | 1.34 | 2.51 | 0/4 | 6.82 | 0/4 |
| KaidanOmni | 11+6 | 1.21 | 0.00 | 2/11 | 0.78 | 0.00 | 11/11 | 0.56 | 2.43 | 0.68 | 6/6 | 0.81 | 6/6 |
| Ladybug | 9+4 | 1.03 | 0.57 | 4/9 | 0.92 | 1.89 | 0/9 | 0.67 | 8.56 | 1.18 | 3/4 | 1.70 | 0/4 |
| MiniOmni | 10+5 | 0.89 | 0.00 | 7/10 | 0.38 | 0.03 | 10/10 | 0.71 | 1.90 | 1.13 | 4/5 | 7.60 | 0/5 |
| Omni | 9+5 | 0.76 | 0.67 | 8/9 | 0.46 | 1.05 | 3/9 | 0.78 | 1.16 | 1.03 | 2/5 | 2.93 | 0/5 |
| VMRImage | 7+3 | 0.67 | 0.64 | 7/7 | 0.61 | 0.70 | 5/7 | 0.56 | 2.76 | 0.60 | 3/3 | 0.67 | 3/3 |
| **OV** | | | | | | | | | | | | | |
| corner ov00 | 35+15 | 1.28 | 0.49 | 0/35 | 1.46 | 0.64 | 0/35 | 0.70 | 20.05 | 1.34 | 0/15 | 2.62 | 0/15 |
| corner ov01 | 35+15 | 0.61 | 0.25 | 35/35 | 0.71 | 0.31 | 35/35 | 0.71 | 9.32 | 0.96 | 5/15 | 1.56 | 0/15 |
| corner ov02 | 35+15 | 1.08 | 0.80 | 9/35 | 1.23 | 0.58 | 1/35 | 2.11 | 16.16 | 2.59 | 1/15 | 3.48 | 0/15 |
| corner ov03 | 35+15 | 0.60 | 0.33 | 35/35 | 0.65 | 0.33 | 34/35 | 5.67 | 9.07 | 5.82 | 4/15 | 6.45 | 0/15 |
| corner ov04 | 35+15 | 1.08 | 0.78 | 10/35 | 1.23 | 1.01 | 0/35 | 0.71 | 16.77 | 1.36 | 2/15 | 2.30 | 0/15 |
| corner ov05 | 35+15 | 0.93 | 0.43 | 27/35 | 1.07 | 0.58 | 8/35 | 0.79 | 14.60 | 1.19 | 3/15 | 1.90 | 0/15 |
| corner ov06 | 35+15 | 1.95 | 0.94 | 0/35 | 2.06 | 0.77 | 0/35 | 0.71 | 27.96 | 2.00 | 0/15 | 3.16 | 0/15 |
| corner ov07 | 35+15 | 1.05 | 0.61 | 6/35 | 1.16 | 0.47 | 3/35 | 0.74 | 16.33 | 1.44 | 1/15 | 2.20 | 0/15 |
| cube ov00 | 31+14 | 0.11 | 0.03 | 31/31 | 0.04 | 0.15 | 31/31 | 0.30 | 0.70 | 0.34 | 14/14 | 0.49 | 14/14 |
| cube ov01 | 25+12 | 0.04 | 0.02 | 25/25 | 0.05 | 0.07 | 25/25 | 0.27 | 0.23 | 0.29 | 12/12 | 0.37 | 12/12 |
| cube ov02 | 24+11 | 0.06 | 0.01 | 24/24 | 0.04 | 0.06 | 24/24 | 0.28 | 0.25 | 0.29 | 11/11 | 0.34 | 11/11 |
| cube ov03 | 25+12 | 0.08 | 0.04 | 25/25 | 0.04 | 0.14 | 25/25 | 0.30 | 0.44 | 0.31 | 12/12 | 0.38 | 12/12 |
| plane 130108MP | 35+15 | 0.09 | 0.04 | 35/35 | 0.32 | 0.31 | 32/35 | 0.47 | 0.49 | 0.49 | 15/15 | 1.01 | 9/15 |
| plane 2012-A0 | 30+13 | 0.65 | 3.16 | 0/30 | 1.43 | 2.69 | 0/30 | 0.56 | 0.92 | 0.61 | 13/13 | 3.94 | 0/13 |
| plane 3136-H0 | 19+9 | 4.17 | 25.90 | 0/19 | 1.50 | 2.73 | 0/19 | 0.95 | 2.00 | 1.72 | 2/9 | 3.18 | 0/9 |
| plane 5501-C4 | 8+4 | 1.45 | 4.42 | 0/8 | 1.90 | 2.52 | 0/8 | 0.45 | 0.15 | 0.68 | 3/4 | 5.37 | 0/4 |
| **UZH** | | | | | | | | | | | | | |
| DAVIS in45 | 35+15 | 0.12 | 0.18 | 35/35 | 0.27 | 0.56 | 33/35 | 0.36 | 0.14 | 0.37 | 15/15 | 0.43 | 15/15 |
| DAVIS inf | 35+15 | 2.08 | 5.59 | 18/35 | 5.03 | 19.81 | 0/35 | 0.50 | 0.45 | 0.52 | 14/15 | 1.01 | 10/15 |
| DAVIS o45 | 35+15 | 0.19 | 0.46 | 33/35 | 0.82 | 2.29 | 27/35 | 0.30 | 0.09 | 0.30 | 15/15 | 0.50 | 15/15 |
| DAVIS of | 35+15 | 0.57 | 1.42 | 24/35 | 2.42 | 10.48 | 2/35 | 0.49 | 0.79 | 0.49 | 14/15 | 0.94 | 9/15 |
| Snapdragon in45 | 35+15 | 0.12 | 0.09 | 35/35 | 0.33 | 0.68 | 31/35 | 0.26 | 0.47 | 0.30 | 15/15 | 0.68 | 14/15 |
| Snapdragon inf | 35+15 | 0.23 | 0.32 | 33/35 | 0.69 | 1.79 | 26/35 | 0.24 | 0.82 | 0.24 | 15/15 | 0.42 | 15/15 |
| Snapdragon o45 | 35+15 | 0.21 | 0.49 | 34/35 | 0.45 | 0.55 | 33/35 | 0.32 | 0.53 | 0.33 | 15/15 | 0.41 | 15/15 |
| Snapdragon of | 35+15 | 0.08 | 0.09 | 35/35 | 0.25 | 0.52 | 32/35 | 0.21 | 0.43 | 0.23 | 15/15 | 0.31 | 15/15 |

Table 4. Complete results for the checkerboard calibration experiment from Section 6.1 in the main paper. The table shows the per-camera results. See the main paper for details on the experimental setup.