Camera Pose Estimation using Implicit Distortion Models

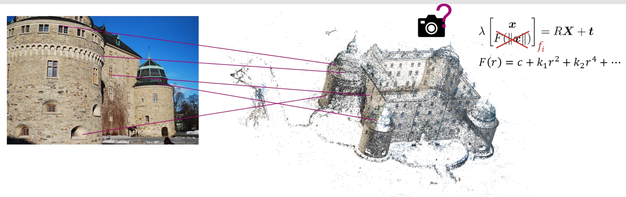
Linfei Pan[1] Marc Pollefeys[1,2] Viktor Larsson[3]

[1] ETH Zurich [2] Microsoft [3] Lund University

CVPR JUNE 19-24 2022 NEW ORLEANS • LOUISIANA

Task

- Estimating full 6 degree-of-freedom (DoF) without knowledge of intrinsic calibration
- Full bundle adjustment with implicit camera calibration model

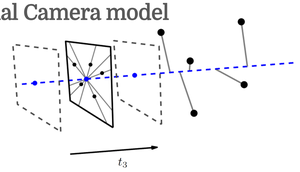


$$\lambda \begin{bmatrix} x \\ F(r)\cdot(y) \end{bmatrix}_{f_i} = RX + t$$

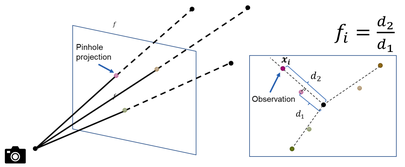
$$F(r) = c + k_1 r^2 + k_2 r^4 + \cdots$$

Terminology

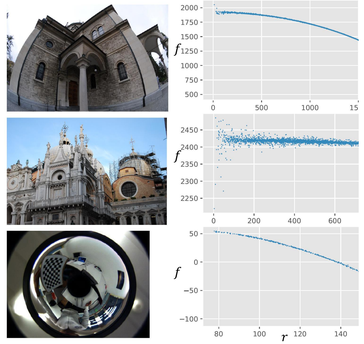
Radial Camera model



Pointwise Focal length



$$f_i = \frac{d_2}{d_1}$$

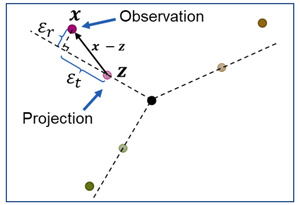


Methodology

Objective Function

- Decomposition of reprojection error

$$\|\boldsymbol{x}-\boldsymbol{z}\|^2 = \underbrace{\|(I-\frac{\boldsymbol{z}\boldsymbol{z}^\top}{\boldsymbol{z}^\top\boldsymbol{z}})\boldsymbol{x}\|^2}_{\varepsilon_r} + \underbrace{\|\frac{\boldsymbol{z}\boldsymbol{z}^\top}{\boldsymbol{z}^\top\boldsymbol{z}}(\boldsymbol{x}-\boldsymbol{z})\|^2}_{\varepsilon_t}$$



- ε_r can be calculated without knowing the focal length or radial distortion
- Replacement of ε_t with $\mathcal{R}(\boldsymbol{f})$
- Proposed objective function

$$\min_{R,\boldsymbol{t}} \sum_{i=1}^{N} \varepsilon_r(R,\boldsymbol{t},\boldsymbol{x}_i,\boldsymbol{X}_i) + \mathcal{R}\left(\{f_i\}_{i=1}^N\right)$$

References

[1] Y. Lochman, et. al. Babelcalib: A universal approach to calibrating central cameras. ICCV 2021

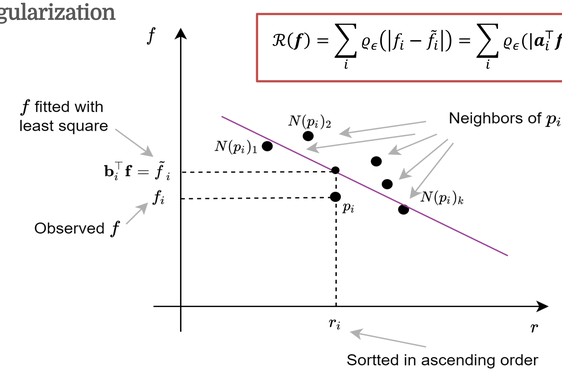
[2] V. Larsson et. al. Calibration-free structure-from-motion with calibrated radial trifocal tensors. ECCV, 2020

[3] C. Olsson et. al. Stable structure from motion for unordered image collections. SCIA, 2011.

[4] T. Sattler et. al. Benchmarking 6dof outdoor visual localization in changing conditions, 2018

[5] H. Taira et. al. Inloc: Indoor visual localization with dense matching and view synthesis, 2018

Regularization



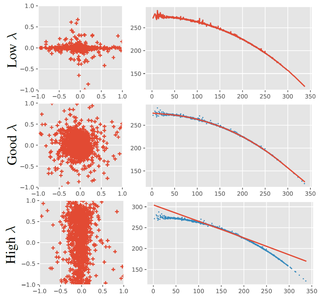
$$\mathcal{R}(\boldsymbol{f}) = \sum_i \varrho_\epsilon(|f_i - \tilde{f}_i|) = \sum_i \varrho_\epsilon(|\boldsymbol{a}_i^\top \boldsymbol{f}|)$$

Non-Parametric Intrinsic Calibration

Algorithm 1: Automatic selection of λ

$\lambda \leftarrow \lambda_{init}$, $best\_res \leftarrow \infty$
$\epsilon_{rad} \leftarrow$ radial reprojection error
for $i \leftarrow 0$ **to** *max-iters* **do**
  Solve (21) to recover distortion mapping
  $\epsilon_{tan} \leftarrow$ tangential reprojection error
  $res \leftarrow |\epsilon_{rad} - \epsilon_{rms}|$
  if $\epsilon_{tan} \leq \epsilon_{rad}$ **then**
    $\lambda \leftarrow 10\lambda$ `// under-regularized`
  else
    $\lambda \leftarrow \lambda/2$ `// over-regularized`
  end
  if $res < best\_res$ **then**
    $best\_res \leftarrow res$
    Save best calibration found so far
  end
end

$$\min_{\boldsymbol{f}} \sum_i \varrho\left(|f_i - \hat{f}_i|\right) + \lambda\mathcal{R}(\boldsymbol{f}) \qquad (21)$$



Bundle Adjustment

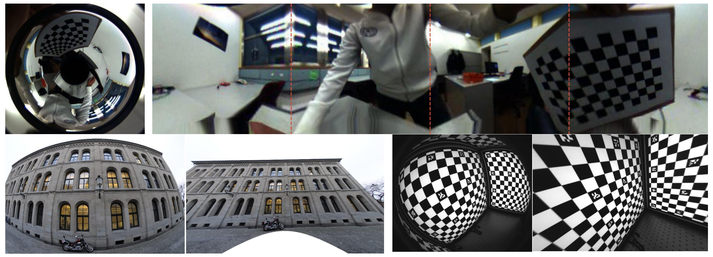
- The regularization relies on the dependence between structure
- To break the dependence of structure, solve a surrogate problem instead

$$\tilde{\boldsymbol{f}}^i = \left[f_{i-m}(\boldsymbol{X}_{i-m}^{t-1}) \cdots f_i(\boldsymbol{X}_i) \cdots f_{i+m}(\boldsymbol{X}_{i+m}^{t-1})\right]^\top$$

- Fix the structure of neighboring points, but still optimize R, t
- Schur complement trick becomes applicable again

Experiments

Qualitative Results



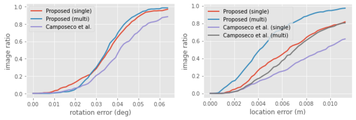
Checkerboard Calibration

- on dataset from BabelCalib [1] , Lochman et al., 2021

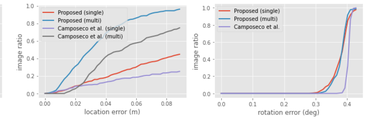
	Proposed			Camposeco et al. [6]		
	ϵ_{rot}	ϵ_{pos}	$< 1°, 1\%$	ϵ_{rot}	ϵ_{pos}	$< 1°, 1\%$
OV corner	**1.07**	**0.58**	**122 / 280**	1.20	0.59	81 / 280
OV cube	**0.07**	**0.03**	**105 / 105**	0.04	0.11	**105 / 105**
OV plane	1.23	6.78	**35 / 92**	**1.06**	**1.78**	32 / 92
Kalibr	**0.17**	**0.18**	**277 / 280**	0.31	0.86	231 / 280
OCamCalib	0.62	**0.26**	**61 / 79**	**0.58**	0.59	55 / 79
UZH DAVIS	**0.74**	**1.91**	**110 / 140**	2.14	8.28	62 / 140
UZH Snapdragon	**0.16**	**0.25**	**137 / 140**	0.43	0.89	122 / 140

	[25]		Proposed		Camposeco et al. [6]	
	ϵ_{rms}^{BC}	ϵ_{pp}	ϵ_{rms}	$<$ 1px	ϵ_{rms}	$<$ 1px
OV corner	1.52	16.28	**2.09**	**16/120**	2.96	0/120
OV cube	0.29	0.40	**0.31**	**49/49**	0.40	49/49
OV plane	0.60	0.89	**0.82**	**33/41**	2.84	9/41
Kalibr	0.21	0.88	**0.30**	**118/120**	0.61	113/120
OCamCalib	0.68	2.17	**0.97**	**31/40**	2.62	17/40
UZH DAVIS	0.41	0.37	**0.42**	**58/60**	0.72	49/60
UZH Snapdragon	0.26	0.56	**0.28**	**60/60**	0.46	59/60

Structure from Motion



Grossmunster [2]



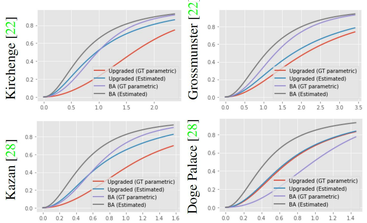
Kazan [3]

Self Calibration in Visual Localization

Aachen Day-Night [37]		day	night
Single image	Proposed (w/ filter)	**58.3 / 76.5 / 94.2**	**61.2 / 77.6 / 99.0**
	Proposed (w/o filter)	51.3 / 67.4 / 92.8	50.0 / 68.4 / 94.9
	Camposeco et al. [6]	46.0 / 61.9 / 83.1	45.9 / 69.4 / 85.7
Multiple images	Proposed (w/ filter)	**82.6 / 92.4 / 98.3**	**73.5 / 88.8 / 100.0**
	Proposed (w/o filter)	77.8 / 90.8 / **98.3**	**73.5 / 88.8 / 100.0**
	Camposeco et al. [6]	18.6 / 34.3 / 83.5	37.8 / 63.3 / 99.0
Parametric model	hloc [33] + GT calib.	**89.6 / 95.4 / 98.8**	**86.7 / 93.9 / 100.0**
	hloc [33] + [21]	60.6 / 82.8 / 98.2	64.3 / 82.7 / 100.0
InLoc [40]		duc1	duc2
Single image	Proposed (w/ filter)	28.3 / 46.0 / 63.6	**26.7 / 48.1 / 61.8**
	Proposed (w/o filter)	**29.8 / 46.5 / 64.6**	26.0 / 42.7 / 59.5
	Camposeco et al. [6]	23.2 / 40.4 / 55.1	18.3 / 31.3 / 42.7
Multiple images	Proposed (w/ filter)	34.8 / 52.5 / **69.7**	**38.9** / 57.3 / **74.0**
	Proposed (w/o filter)	**35.4 / 53.0 / 69.7**	35.9 / **58.0 / 74.0**
	Camposeco et al. [6]	34.8 / 51.0 / 69.2	35.1 / **58.0 / 74.0**
Parametric model	hloc [33] + GT calib.	**46.5 / 66.2 / 78.3**	**51.9 / 74.8 / 78.6**
	hloc [33] + [21]	25.8 / 47.5 / 62.6	27.5 / 55.0 / 66.4

Aachen Day-Night [4] and InLoc [5] Challenge

Implicit Distortion in Bundle Adjustment



Kirchenge and Grossmunster [2]; Kazan and Doge Palace [3]