

# Camera Pose Estimation with Unknown Principal Point

## Supplementary material

Viktor Larsson  
Lund University  
Lund, Sweden

viktorl@maths.lth.se

Zuzana Kukelova  
Czech Technical University  
Prague, Czech Republic

kukelzuz@fel.cvut.cz

Yinqiang Zheng  
National Institute of Informatics  
Tokyo, Japan

yqzheng@nii.ac.jp

### 1. New Camera Matrix Constraints

Below is the code for computing the constraints in Macaulay2 [9].

```
R = QQ[p11,p12,p13,p21,p22,p23,p31,p32,p33];

p1 = matrix({{p11,p12,p13}});
p2 = matrix({{p21,p22,p23}});
p3 = matrix({{p31,p32,p33}});

p1p3 = matrix({{p12*p33-p13*p32,
                p13*p31-p11*p33,
                p11*p32-p12*p31}});
p2p3 = matrix({{p22*p33-p23*p32,
                p23*p31-p21*p33,
                p21*p32-p22*p31}});

eq1 = p1p3*transpose(p2p3);
eq2 = p1p3*transpose(p1p3)-p2p3*transpose(p2p3);

P = p1 || p2 || p3;

I = ideal {eq1,eq2};
J = saturate(I, det(P));

gg = mingens J
```

The 5 constraints of degree 5 are:

$$p_{11}p_{12}p_{32}p_{33} + p_{11}p_{12}p_{33}^3 - p_{11}p_{13}p_{32}^3 - p_{11}p_{13}p_{32}p_{33}^2 - p_{12}^2p_{31}p_{32}p_{33} + p_{12}p_{13}p_{31}p_{32}^2 - p_{12}p_{13}p_{31}p_{33}^2 + p_{13}^2p_{31}p_{32}p_{33} + p_{21}p_{22}p_{32}^2p_{33} + p_{21}p_{22}p_{33}^3 - p_{21}p_{23}p_{32}^2 - p_{21}p_{23}p_{32}p_{33}^2 - p_{22}^2p_{31}p_{32}p_{33} + p_{22}p_{23}p_{31}p_{32}^2 - p_{22}p_{23}p_{31}p_{33}^2 + p_{23}^2p_{31}p_{32}p_{33} = 0$$

$$p_{11}p_{12}p_{31}p_{33}^2 - p_{11}p_{13}p_{31}p_{32}p_{33} + p_{12}^2p_{32}p_{33}^2 - p_{12}p_{13}p_{31}p_{33}^2 - 2p_{12}p_{13}p_{32}^2p_{33} + p_{13}^2p_{31}^2p_{32} + p_{13}^2p_{32}^2 - p_{21}^2p_{32}p_{33}^2 + 2p_{21}p_{22}p_{31}p_{32}^2 + p_{21}p_{22}p_{31}p_{33}^2 + p_{21}p_{23}p_{31}p_{32}p_{33} - p_{22}^2p_{31}^2p_{32} - p_{22}p_{23}p_{31}^2p_{33} = 0$$

$$p_{11}^2p_{32}^2p_{33} + p_{11}^2p_{33}^3 - p_{11}p_{12}p_{31}p_{32}p_{33} - p_{11}p_{13}p_{31}p_{32}^2 - 2p_{11}p_{13}p_{31}p_{33}^2 + p_{12}p_{13}p_{31}^2p_{32} + p_{13}^2p_{31}^2p_{33} + p_{21}p_{22}p_{31}p_{32}p_{33} - p_{21}p_{23}p_{31}p_{32}^2 - p_{22}^2p_{31}^2p_{33} - p_{22}^2p_{33}^2 +$$

$$p_{22}p_{23}p_{31}^2p_{32} + 2p_{22}p_{23}p_{32}p_{33}^2 - p_{23}^2p_{32}^2p_{33} = 0$$

$$p_{11}^2p_{31}p_{33}^2 + p_{11}p_{12}p_{32}p_{33}^2 - 2p_{11}p_{13}p_{31}^2p_{33} - p_{11}p_{13}p_{32}^2p_{33} - p_{12}p_{13}p_{31}p_{32}p_{33} + p_{13}^2p_{31}^3 + p_{13}^2p_{31}p_{32}^2 - p_{21}^2p_{31}p_{32}^2 + 2p_{21}p_{22}p_{31}^2p_{32} + p_{21}p_{22}p_{32}p_{33}^2 - p_{21}p_{23}p_{32}^2p_{33} - p_{22}^2p_{31}^3 - p_{22}^2p_{31}p_{33}^2 + p_{22}p_{23}p_{31}p_{32}p_{33} = 0$$

$$p_{11}^2p_{31}p_{32}p_{33} - p_{11}p_{12}p_{31}^2p_{33} - p_{11}p_{12}p_{33}^3 - p_{11}p_{13}p_{31}^2p_{32} + p_{11}p_{13}p_{32}p_{33}^2 + p_{12}p_{13}p_{31}^3 + p_{12}p_{13}p_{31}p_{33}^2 - p_{13}^2p_{31}p_{32}p_{33} + p_{21}^2p_{31}p_{32}p_{33} - p_{21}p_{22}p_{31}^2p_{33} - p_{21}p_{22}p_{33}^3 - p_{21}p_{23}p_{31}^2p_{32} + p_{21}p_{23}p_{32}p_{33}^2 + p_{22}p_{23}p_{31}^3 + p_{22}p_{23}p_{31}p_{32}^2 - p_{23}^2p_{31}p_{32}p_{33} = 0$$

### 2. Synthetic Experiments

In this section we show some more results from the synthetic experiments. Figure 1 shows the error in the focal length, principal point, rotation and translation for the varying noise experiment.

Figure 2 shows the distribution of the relative focal length error for the 2 px noise for the P7Pfruv solver, with and without bundle adjustment.

### 3. Real Images

Here we show more errors for the experiments with real images, including the results for the additional non-linear refinement for the P7Pfruv solver (marked with (b) in the tables). Table 1 shows the mean, median and maximum errors for the *NotreDame* dataset. Table 2 shows the results for the images where the principal point was shifted more than 6% of the image size and Table 3 shows the errors for the images with less than 6% shift. We also show the results for the cropped *Rotunda* dataset in Table 4.

### References

- [1] RealityCapture. [www.capturingreality.com](http://www.capturingreality.com).
- [2] M. Bujnak, Z. Kukelova, and T. Pajdla. A general solution to the p4p problem for camera with unknown focal length. In *Computer Vision and Pattern Recognition (CVPR)*, 2008.

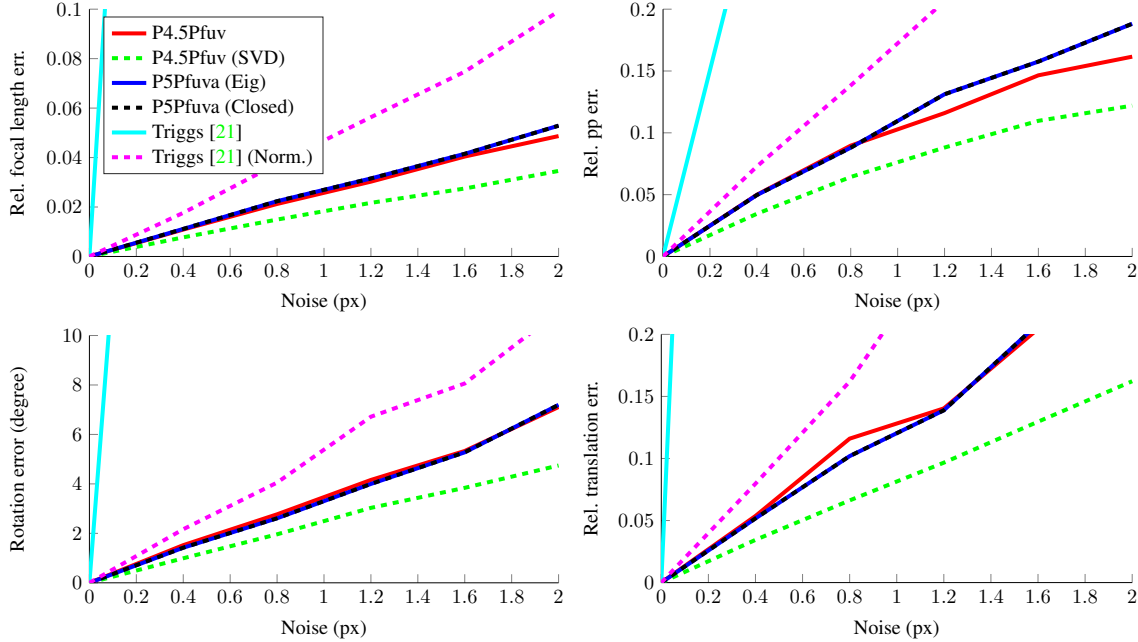


Figure 1. Error for varying noise. *Top left*: Relative focal length error. *Top right*: Relative principal point error. *Bottom left*: Rotation error (in degrees). *Bottom right*: Relative translation error.

		P3.5Pf [18]	P6P [11]	P5Pfvu [21]	P4.5Pfvu	P4.5Pfvu (SVD)	P4.5Pfvu (6pt)	P5Pfvu	P4Pfr [18]	P7Pfruv	P7Pfruv (b)
Focal length	mean	0.1256	0.0242	0.0234	0.0209	0.0189	<b>0.0142</b>	0.0243	0.1012	0.0211	0.0037
	median	0.0458	0.0137	0.0138	0.0142	0.0138	<b>0.0096</b>	0.0159	0.0264	0.0165	0.0024
	max	1.3042	0.1342	0.1123	0.1449	0.0962	<b>0.0535</b>	0.1176	0.9307	0.0807	0.0186
Rotation	mean	2.9495	1.2067	1.2559	1.2544	1.2564	<b>1.1228</b>	1.2322	3.0043	0.9005	0.3166
	median	1.6891	0.8814	0.9778	1.0582	1.1260	0.9255	<b>0.8546</b>	1.6816	0.7009	0.2085
	max	19.5789	5.1510	4.1978	3.9726	3.8889	4.0754	4.5883	21.7933	<b>2.7497</b>	1.9464
Translation	mean	0.6196	0.1163	0.1249	0.0894	0.0849	<b>0.0714</b>	0.1200	0.4941	0.1016	0.0188
	median	0.2052	0.0522	0.0569	0.0586	0.0629	<b>0.0414</b>	0.0640	0.1315	0.0629	0.0097
	max	5.4442	0.6711	1.6183	<b>0.5100</b>	0.5809	0.7289	0.7050	4.1039	0.5929	0.1665
Principal point	mean	-	0.0319	0.0317	0.0335	0.0335	0.0296	0.0324	-	<b>0.0227</b>	0.0089
	median	-	0.0226	0.0223	0.0239	0.0236	0.0207	0.0224	-	<b>0.0168</b>	0.0054
	max	-	0.2263	0.2829	0.4484	0.4411	0.3877	0.3491	-	<b>0.1735</b>	0.0808
Distortion	mean	-	-	-	-	-	-	-	0.0868	<b>0.0474</b>	0.0165
	median	-	-	-	-	-	-	-	0.0290	<b>0.0173</b>	0.0071
	max	-	-	-	-	-	-	-	<b>1.0509</b>	1.9341	0.3840
Inlier (%)	mean	74.9315	88.0265	86.8770	87.8971	88.6420	89.4039	87.5297	82.1422	<b>93.2453</b>	98.3413
	median	79.6975	90.6261	90.3534	91.7344	91.8065	92.6474	90.3015	89.3571	<b>94.6934</b>	98.3413
	max	<b>99.6805</b>	99.2391	99.0581	99.3407	99.4610	99.6711	99.3856	99.6805	99.3610	100.0000

Table 1. Full results for the *NotreDame* dataset: Comparison of different solvers on 81 images downloaded from the Internet. The table shows the relative errors except for the rotation errors which are in degrees. For the principal point the error is relative to the image size. The best results (excluding P7Pfruv with bundle adjustment) are marked bold.

[3] M. Bujnak, Z. Kukelova, and T. Pajdla. New efficient solution to the absolute pose problem for camera with unknown focal length and radial distortion. In *Asian Conference on Computer Vision (ACCV)*, 2010.

[4] D. Cox, J. Little, and D. O’shea. *Ideals, varieties, and algorithms*, volume 3. Springer, 1992.

[5] O. Faugeras. *Three-dimensional computer vision: a geometric viewpoint*. MIT press, 1993.

[6] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image anal-

ysis and automated cartography. *Comm. ACM*, 24(6):381–395, 1981.

[7] A. W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Computer Vision and Pattern Recognition (CVPR)*, 2001.

[8] X. Gao, X. Hou, J. Tang, and H. Cheng. Complete solution classification for the perspective-three-point problem. *Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 25(8):930–943, 2003.

[9] D. R. Grayson and M. E. Stillman. Macaulay 2, a software

		P3.5Pf [18]	P6P [11]	P5Pfov [21]	P4.5Pfov	P4.5Pfov (SVD)	P4.5Pfov (6pt)	P5Pfuva	P4Pfr [18]	P7Pfruv	P7Pfruv (b)
Focal length	mean	0.2615	0.0295	0.0228	0.0185	0.0186	<b>0.0123</b>	0.0286	0.2177	0.0192	0.0043
	median	0.1409	0.0150	0.0138	0.0139	<b>0.0089</b>	0.0093	0.0224	0.1363	0.0163	0.0031
	max	1.3042	0.1342	0.0809	0.0572	0.0762	<b>0.0437</b>	0.1176	0.9307	0.0807	0.0186
Rotation	mean	5.6991	1.6485	1.5334	1.3566	1.3602	1.3361	1.5225	5.9204	<b>0.8759</b>	0.3942
	median	4.6145	1.1001	1.0628	1.0924	1.1260	1.0800	1.1426	4.6211	<b>0.6414</b>	0.2657
	max	19.5789	5.1510	4.1978	3.9726	3.8889	4.0754	4.5883	21.7933	<b>2.7497</b>	1.9127
Translation	mean	1.2483	0.1679	0.1528	0.1004	0.0978	<b>0.0699</b>	0.1646	1.0641	0.1140	0.0283
	median	0.7351	0.0941	0.0735	0.0745	0.0759	<b>0.0601</b>	0.1105	0.7020	0.0797	0.0135
	max	5.4442	0.6711	1.6183	0.3328	0.3557	<b>0.2896</b>	0.6612	4.1039	0.4998	0.1665
Principal point	mean	-	0.0413	0.0373	0.0333	0.0334	0.0319	0.0396	-	<b>0.0214</b>	0.0108
	median	-	0.0291	0.0277	0.0300	0.0270	0.0258	0.0286	-	<b>0.0163</b>	0.0071
	max	-	0.1559	0.0971	0.0965	0.1114	0.0907	0.1375	-	<b>0.0669</b>	0.0572
Distortion	mean	-	-	-	-	-	-	-	0.1647	<b>0.0240</b>	0.0121
	median	-	-	-	-	-	-	-	0.1102	<b>0.0176</b>	0.0066
	max	-	-	-	-	-	-	-	1.0509	<b>0.1082</b>	0.1137
Inlier (%)	mean	61.7820	89.4821	87.9500	89.3505	90.0521	90.8913	88.6629	66.8490	<b>93.0248</b>	98.2320
	median	68.0180	91.7491	90.3534	92.2825	92.1621	92.7361	90.9337	70.3435	<b>94.9796</b>	98.2320
	max	94.9309	99.1837	99.0581	99.3407	99.3563	<b>99.6711</b>	99.3856	96.7742	99.1031	100.0000

Table 2. Full results for the *NotreDame* dataset: Comparison of different solvers on images with principal point shift  $> 6\%$ . The table shows the relative errors except for the rotation errors which are in degrees. For the principal point the error is relative to the image size. The best results (excluding P7Pfruv with bundle adjustment) are marked bold.

		P3.5Pf [18]	P6P [11]	P5Pfov [21]	P4.5Pfov	P4.5Pfov (SVD)	P4.5Pfov (6pt)	P5Pfuva	P4Pfr [18]	P7Pfruv	P7Pfruv (b)
Focal length	mean	0.0351	0.0206	0.0238	0.0225	0.0191	<b>0.0155</b>	0.0214	0.0236	0.0223	0.0034
	median	0.0240	0.0124	0.0136	0.0142	0.0157	<b>0.0096</b>	0.0150	0.0159	0.0166	0.0015
	max	0.2235	0.0920	0.1123	0.1449	0.0962	<b>0.0535</b>	0.1137	0.0870	0.0757	0.0173
Rotation	mean	1.1164	<b>0.9121</b>	1.0709	1.1863	1.1872	0.9806	1.0387	1.0603	0.9169	0.2649
	median	0.9137	0.8007	0.9329	0.9518	1.1145	0.8911	<b>0.7650</b>	0.9135	0.7830	0.1866
	max	3.2030	2.5408	3.0633	3.4579	3.6908	2.6247	4.2505	2.7917	<b>2.3552</b>	1.9464
Translation	mean	0.2005	0.0818	0.1064	0.0820	0.0763	<b>0.0724</b>	0.0902	0.1142	0.0934	0.0126
	median	0.0861	0.0361	0.0478	0.0533	0.0542	<b>0.0356</b>	0.0481	0.0568	0.0553	0.0072
	max	2.2931	<b>0.4828</b>	0.5347	0.5100	0.5809	0.7289	0.7050	1.2441	0.5929	0.0772
Principal point	mean	-	0.0257	0.0280	0.0336	0.0335	0.0280	0.0275	-	<b>0.0236</b>	0.0077
	median	-	0.0191	0.0204	0.0213	0.0231	0.0188	<b>0.0152</b>	-	0.0187	0.0043
	max	-	0.2263	0.2829	0.4484	0.4411	0.3877	0.3491	-	<b>0.1735</b>	0.0808
Distortion	mean	-	-	-	-	-	-	-	<b>0.0349</b>	0.0631	0.0195
	median	-	-	-	-	-	-	-	<b>0.0136</b>	0.0161	0.0091
	max	-	-	-	-	-	-	-	<b>0.6777</b>	1.9341	0.3840
Inlier (%)	mean	83.6978	87.0561	86.1617	86.9281	87.7019	88.4123	86.7742	92.3377	<b>93.3923</b>	98.4141
	median	85.4583	89.7185	90.3243	90.7809	90.5397	92.1720	90.1257	93.8954	<b>94.5743</b>	98.4141
	max	<b>99.6805</b>	99.2391	99.0415	99.2063	99.4610	99.3976	98.8903	<b>99.6805</b>	99.3610	100.0000

Table 3. Full results for the *NotreDame* dataset: Comparison of different solvers on images with principal point shift  $< 6\%$ . The table shows the relative errors except for the rotation errors which are in degrees. For the principal point the error is relative to the image size. The best results (excluding P7Pfruv with bundle adjustment) are marked bold.

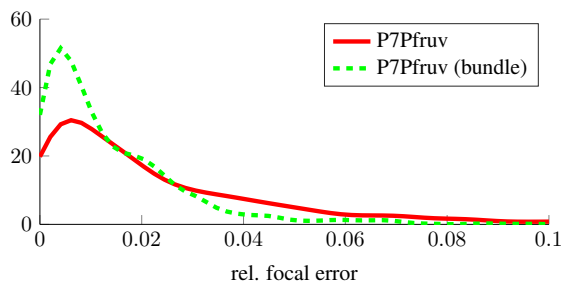


Figure 2. Distribution of relative focal length error for 2 px noise with radial distortion.

system for research in algebraic geometry, 2002. 1

- [10] B. Haralick, C. Lee, K. Ottenberg, and M. Ise. Review and analysis of solutions of the three point perspective pose estimation problem. *International Journal of Computer Vision (IJCV)*, 13(3):331–356, 1994.
- [11] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge Univ. Press, 2nd edition, 2003. 2, 3, 4
- [12] A. Heyden. *Geometry and Algebra of Multiple Projective Transforms*. PhD thesis, PhD thesis, Lund University, Lund Institute of Technology, Department of Mathematics, 1995.
- [13] K. Josephson and M. Byröd. Pose estimation with radial distortion and unknown focal length. In *Computer Vision and Pattern Recognition (CVPR)*, 2009.

		P3.5Pf [18]	P6P [11]	P5Pfuv [21]	P4.5Pfuv	P4.5Pfuv (SVD)	P4.5Pfuv (6pt)	P5Pfuva	P4Pfr [18]	P7Pfruv	P7Pfruv (b)
Focal length	mean	0.3696	0.2199	0.1667	0.1699	0.1666	0.1696	0.1928	0.0845	<b>0.0012</b>	0.0008
	median	0.2629	0.1564	0.1386	0.1319	0.1290	0.1368	0.1494	0.0540	<b>0.0010</b>	0.0007
	max	4.1083	0.5135	0.5066	0.4187	0.4461	0.4217	0.5450	0.5896	<b>0.0040</b>	0.0053
Rotation	mean	14.6800	10.7499	8.4318	8.4168	8.3194	8.2880	9.7119	13.6371	<b>0.2558</b>	0.1837
	median	13.3526	10.0873	7.8776	8.1739	7.8868	7.6738	9.4980	13.3771	<b>0.2191</b>	0.1588
	max	176.5515	24.4273	22.8580	20.9626	22.7519	22.4078	27.0481	21.1274	<b>0.7516</b>	0.7629
Translation	mean	0.4538	0.1687	0.1446	0.1436	0.1434	0.1450	0.1535	0.2181	<b>0.0043</b>	0.0031
	median	0.2673	0.1218	0.1175	0.1170	0.1136	0.1129	0.1211	0.2062	<b>0.0038</b>	0.0026
	max	5.9461	0.8867	0.6083	0.6217	0.6804	0.7358	0.6327	0.7070	<b>0.0138</b>	0.0138
Principal point	mean	-	0.0635	0.0619	0.0616	0.0614	0.0608	0.0655	-	<b>0.0022</b>	0.0016
	median	-	0.0633	0.0536	0.0556	0.0539	0.0535	0.0624	-	<b>0.0018</b>	0.0012
	max	-	0.1710	0.2004	0.1857	0.2008	0.1764	0.1798	-	<b>0.0080</b>	0.0080
Distortion	mean	-	-	-	-	-	-	-	0.2042	<b>0.0067</b>	0.0045
	median	-	-	-	-	-	-	-	0.1271	<b>0.0054</b>	0.0034
	max	-	-	-	-	-	-	-	2.9758	<b>0.0334</b>	0.0232
Inlier (%)	mean	21.26	40.46	34.73	36.04	36.81	36.93	37.45	33.18	<b>96.78</b>	97.49
	median	18.67	38.55	31.97	32.73	33.58	33.98	35.22	29.93	<b>97.06</b>	97.80
	max	55.03	67.12	63.42	62.06	64.12	65.65	65.76	64.50	<b>98.93</b>	99.19

Table 4. Full results for the cropped *Rotunda* dataset: Comparison of different solvers on 248 images with radial distortion and shifted principal point. The table shows the relative errors except for the rotation errors which are in degrees. For the principal point the error is relative to the image size. The best results (excluding P7Pfruv with bundle adjustment) are marked bold.

- [14] 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. In *Computer Vision and Pattern Recognition (CVPR)*, 2011.
- [15] Z. Kukulova, M. Bujnak, and T. Pajdla. Real-time solution to the absolute pose problem with unknown radial distortion and focal length. In *International Conference on Computer Vision (ICCV)*, 2013.
- [16] Z. Kukulova, J. Heller, M. Bujnak, A. Fitzgibbon, and T. Pajdla. Efficient solution to the epipolar geometry for radially distorted cameras. In *International Conference on Computer Vision (ICCV)*, 2015.
- [17] V. Larsson and K. Åström. Uncovering symmetries in polynomial systems. In *European Conference on Computer Vision (ECCV)*, 2016.
- [18] V. Larsson, Z. Kukulova, and Y. Zheng. Making minimal solvers for absolute pose estimation compact and robust. In *International Conference on Computer Vision (ICCV)*, 2017. [2](#), [3](#), [4](#)
- [19] V. Larsson, K. Åström, and M. Oskarsson. Efficient solvers for minimal problems by syzygy-based reduction. In *Computer Vision and Pattern Recognition (CVPR)*, 2017.
- [20] V. Larsson, K. Åström, and M. Oskarsson. Polynomial solvers for saturated ideals. In *International Conference on Computer Vision (ICCV)*, 2017.
- [21] B. Triggs. Camera pose and calibration from 4 or 5 known 3d points. In *International Conference on Computer Vision (ICCV)*, 1999. [2](#), [3](#), [4](#)
- [22] R. Tsai. A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses. *IEEE Journal on Robotics and Automation*, 3(4):323–344, 1987.
- [23] C. Wu. P3.5p: Pose estimation with unknown focal length. In *Computer Vision and Pattern Recognition (CVPR)*, 2015.
- [24] Y. Zheng, S. Sugimoto, I. Sato, and M. Okutomi. A general and simple method for camera pose and focal length determination. In *Computer Vision and Pattern Recognition (CVPR)*, 2014.