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].

\[
R = \text{QQ}\{p11, p12, p13, p21, p22, p23, p31, p32, p33\};
\]

\[
p1 = \text{matrix}\{\{p11, p12, p13\}\};
\]

\[
p2 = \text{matrix}\{\{p21, p22, p23\}\};
\]

\[
p3 = \text{matrix}\{\{p31, p32, p33\}\};
\]

\[
\text{plp3} = \text{matrix}\{\{p12*p33-p13*p32, p13*p31-p11*p33, p11*p32-p12*p31\}\};
\]

\[
\text{p2p3} = \text{matrix}\{\{p22*p33-p23*p32, p23*p31-p21*p33, p21*p32-p22*p31\}\};
\]

\[
eq1 = \text{plp3}^\text{transpose}(p2p3);
\]

\[
eq2 = \text{plp3}^\text{transpose}(\text{plp3}) - \text{p2p3}^\text{transpose}(p2p3);
\]

\[
\text{P} = \text{ideal} (\text{eq1}, \text{eq2});
\]

\[
\text{J} = \text{saturate}(I, \det(P));
\]

\[
\text{gg} = \text{mings J}
\]

The 5 constraints of degree 5 are:

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

\[
\text{p}_{11}^2p_{31}p_{33}^2 - p_{11}^2p_{23}p_{32}p_{33} - p_{12}^2p_{31}p_{32}p_{33} + 2p_{11}p_{12}p_{32}p_{33} - p_{12}p_{13}p_{32}p_{33} + p_{12}^2p_{31}p_{32}p_{33} - p_{13}p_{21}p_{32}p_{33} + p_{21}^2p_{23}p_{31}p_{33} - p_{21}^2p_{22}p_{31}p_{33} = 0
\]

\[
\text{p}_{11}^2p_{31}p_{33}^2 - p_{12}^2p_{31}p_{32}p_{33} - p_{13}^2p_{31}p_{32}p_{33} + p_{21}^2p_{23}p_{31}p_{33} - p_{23}^2p_{31}p_{32}p_{33} = 0
\]

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

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.

![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.](image_url)

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Focal length</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean</td>
<td>0.1256</td>
<td>0.0242</td>
<td>0.0234</td>
<td>0.0209</td>
<td>0.0189</td>
<td>0.0142</td>
<td>0.0243</td>
<td>0.1012</td>
<td>0.0211</td>
<td>0.0037</td>
</tr>
<tr>
<td>median</td>
<td>0.0458</td>
<td>0.0137</td>
<td>0.0138</td>
<td>0.0142</td>
<td>0.0138</td>
<td>0.0096</td>
<td>0.0159</td>
<td>0.0264</td>
<td>0.0165</td>
<td>0.0024</td>
</tr>
<tr>
<td>max</td>
<td>1.3042</td>
<td>0.1342</td>
<td>0.1123</td>
<td>0.1449</td>
<td>0.0962</td>
<td><strong>0.0535</strong></td>
<td>0.1176</td>
<td>0.9307</td>
<td>0.0807</td>
<td>0.0186</td>
</tr>
<tr>
<td>Rotation</td>
<td>mean</td>
<td>2.9495</td>
<td>1.2067</td>
<td>1.2559</td>
<td>1.2544</td>
<td>1.2564</td>
<td>1.1228</td>
<td>1.2322</td>
<td>3.0043</td>
<td>0.9005</td>
</tr>
<tr>
<td>Translation</td>
<td>mean</td>
<td>0.6196</td>
<td>0.1163</td>
<td>0.1249</td>
<td>0.0894</td>
<td>0.0849</td>
<td>0.0714</td>
<td>0.1200</td>
<td>0.4941</td>
<td>0.1016</td>
</tr>
<tr>
<td></td>
<td>median</td>
<td>0.2052</td>
<td>0.0522</td>
<td>0.0569</td>
<td>0.0586</td>
<td>0.0629</td>
<td><strong>0.0414</strong></td>
<td>0.0640</td>
<td>0.1315</td>
<td>0.0629</td>
</tr>
<tr>
<td></td>
<td>max</td>
<td>5.4442</td>
<td>0.6711</td>
<td>1.6183</td>
<td>0.5100</td>
<td>0.5809</td>
<td>0.7289</td>
<td>0.7050</td>
<td>4.1039</td>
<td>0.5929</td>
</tr>
<tr>
<td>Principal point</td>
<td>mean</td>
<td>-</td>
<td>0.0319</td>
<td>0.0317</td>
<td>0.0335</td>
<td>0.0335</td>
<td>0.0296</td>
<td>0.0324</td>
<td>-</td>
<td><strong>0.0227</strong></td>
</tr>
<tr>
<td></td>
<td>median</td>
<td>-</td>
<td>0.0226</td>
<td>0.0223</td>
<td>0.0239</td>
<td>0.0236</td>
<td>0.0207</td>
<td>0.0224</td>
<td>-</td>
<td><strong>0.0168</strong></td>
</tr>
<tr>
<td></td>
<td>max</td>
<td>-</td>
<td>0.2263</td>
<td>0.2829</td>
<td>0.4484</td>
<td>0.4411</td>
<td>0.3877</td>
<td>0.3491</td>
<td>-</td>
<td><strong>0.1735</strong></td>
</tr>
<tr>
<td>Distortion</td>
<td>mean</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.0868</td>
<td>-</td>
<td><strong>0.0474</strong></td>
</tr>
<tr>
<td></td>
<td>median</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.0290</td>
<td>-</td>
<td><strong>0.0173</strong></td>
</tr>
<tr>
<td></td>
<td>max</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1.0509</td>
<td>-</td>
<td><strong>1.9341</strong></td>
</tr>
<tr>
<td>Inlier (%)</td>
<td>mean</td>
<td>74.9315</td>
<td>88.0265</td>
<td>86.8770</td>
<td>87.8971</td>
<td>88.6420</td>
<td>89.4039</td>
<td>87.5297</td>
<td>82.1422</td>
<td><strong>93.2453</strong></td>
</tr>
<tr>
<td></td>
<td>median</td>
<td>79.6975</td>
<td>90.6261</td>
<td>90.3534</td>
<td>91.7344</td>
<td>91.8065</td>
<td>92.6474</td>
<td>90.3015</td>
<td>89.3571</td>
<td><strong>94.6934</strong></td>
</tr>
<tr>
<td></td>
<td>max</td>
<td><strong>99.6805</strong></td>
<td><strong>99.2391</strong></td>
<td><strong>99.0881</strong></td>
<td><strong>99.3407</strong></td>
<td><strong>99.4610</strong></td>
<td><strong>99.6711</strong></td>
<td><strong>99.3856</strong></td>
<td><strong>99.6805</strong></td>
<td><strong>99.3610</strong></td>
</tr>
</tbody>
</table>


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.

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.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Focal length</strong></td>
<td>mean 0.3696</td>
<td>0.2199</td>
<td>0.1676</td>
<td>0.1699</td>
<td>0.1666</td>
<td>0.1696</td>
<td>0.1928</td>
<td>0.0845</td>
<td>0.0012</td>
</tr>
<tr>
<td></td>
<td>median 0.2629</td>
<td>0.1564</td>
<td>0.1386</td>
<td>0.1319</td>
<td>0.1290</td>
<td>0.1368</td>
<td>0.1494</td>
<td>0.0540</td>
<td>0.0010</td>
</tr>
<tr>
<td></td>
<td>max 4.1083</td>
<td>0.5135</td>
<td>0.5066</td>
<td>0.4187</td>
<td>0.4461</td>
<td>0.4217</td>
<td>0.5450</td>
<td>0.5896</td>
<td>0.0040</td>
</tr>
<tr>
<td><strong>Rotation</strong></td>
<td>mean 14.6800</td>
<td>10.7499</td>
<td>8.4318</td>
<td>8.4168</td>
<td>8.3194</td>
<td>8.2880</td>
<td>9.7119</td>
<td>9.4980</td>
<td>0.1719</td>
</tr>
<tr>
<td></td>
<td>median 13.3526</td>
<td>10.0873</td>
<td>7.8776</td>
<td>7.8686</td>
<td>7.6738</td>
<td>7.6900</td>
<td>9.1371</td>
<td>21.1274</td>
<td>0.7516</td>
</tr>
<tr>
<td></td>
<td>max 176.3515</td>
<td>24.4273</td>
<td>22.8580</td>
<td>22.7519</td>
<td>22.4078</td>
<td>27.0481</td>
<td>21.1274</td>
<td>0.7516</td>
<td>0.7629</td>
</tr>
<tr>
<td><strong>Translation</strong></td>
<td>mean 0.4538</td>
<td>0.1687</td>
<td>0.1446</td>
<td>0.1434</td>
<td>0.1450</td>
<td>0.1535</td>
<td>0.2181</td>
<td>0.0043</td>
<td>0.0031</td>
</tr>
<tr>
<td></td>
<td>median 0.2673</td>
<td>0.1218</td>
<td>0.1170</td>
<td>0.1136</td>
<td>0.1129</td>
<td>0.1211</td>
<td>0.2062</td>
<td>0.0038</td>
<td>0.0026</td>
</tr>
<tr>
<td></td>
<td>max 5.9461</td>
<td>0.8867</td>
<td>0.6083</td>
<td>0.6217</td>
<td>0.6804</td>
<td>0.7358</td>
<td>0.7070</td>
<td>0.0138</td>
<td>0.0138</td>
</tr>
<tr>
<td><strong>Principal point</strong></td>
<td>mean -</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.1271</td>
</tr>
<tr>
<td></td>
<td>median -</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.1271</td>
</tr>
<tr>
<td></td>
<td>max -</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2.9758</td>
</tr>
<tr>
<td><strong>Distortion</strong></td>
<td>mean 21.26</td>
<td>40.46</td>
<td>34.73</td>
<td>36.04</td>
<td>36.81</td>
<td>36.93</td>
<td>37.45</td>
<td>33.18</td>
<td>98.78</td>
</tr>
<tr>
<td></td>
<td>median 18.67</td>
<td>38.55</td>
<td>31.97</td>
<td>32.73</td>
<td>33.58</td>
<td>33.98</td>
<td>35.22</td>
<td>29.93</td>
<td>97.06</td>
</tr>
<tr>
<td></td>
<td>max 55.03</td>
<td>67.12</td>
<td>63.42</td>
<td>62.06</td>
<td>64.12</td>
<td>65.65</td>
<td>65.76</td>
<td>64.50</td>
<td>98.93</td>
</tr>
</tbody>
</table>

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 P7Pfuv with bundle adjustment) are marked bold.


