Supplementary Material for *Loosely-Coupled Semi-Direct Monocular SLAM*

Seong Hun Lee and Javier Civera

November 2018

1 Median Errors

Table 1: **[EuRoC MAV]** Median absolute trajectory errors e_{ate} [m] over 10 runs on each sequence. The best and the worst results are in blue and red, respectively. The mark × indicates that more than half of the total runs lost tracking due to challenging scene textures, camera motions or illumination changes.

	ORB VO	ORB SLAM	DSO default	DSO reduced	Ours VO	Ours SLAM
M1L	0.046	0.046	0.053	**0.076**	**0.044**	0.046
M2L	0.038	**0.037**	0.050	**0.071**	0.038	**0.037**
M3L	**0.039**	**0.039**	0.166	**0.249**	0.040	0.040
M4L	0.064	0.065	**0.172**	0.163	**0.060**	0.064
M5L	0.078	**0.051**	0.101	**0.144**	0.061	0.065
V11L	0.096	0.097	0.104	**0.115**	**0.089**	0.090
V12L	0.127	**0.103**	**0.516**	0.136	0.104	0.105
V13L	0.938	0.423	×	0.757	0.608	**0.412**
V21L	**0.080**	**0.080**	0.089	**0.090**	0.089	0.088
V22L	0.185	**0.178**	0.210	**0.233**	0.181	**0.178**
V23L	×	×	1.429	**1.055**	1.174	1.062

	ORB VO	ORB SLAM	DSO default	DSO reduced	Ours VO	Ours SLAM
M1R	0.040	**0.038**	0.052	**0.076**	**0.038**	0.039
M2R	0.038	0.038	0.052	**0.071**	**0.036**	**0.036**
M3R	0.045	**0.043**	0.149	**0.185**	**0.043**	0.045
M4R	0.074	**0.069**	0.129	**0.174**	0.070	0.074
M5R	0.056	**0.055**	0.095	**0.122**	0.065	0.060
V11R	0.103	0.104	0.119	**0.215**	0.100	**0.099**
V12R	0.125	0.112	0.133	**0.161**	**0.111**	**0.111**
V13R	×	0.130	×	**0.614**	1.277	0.825
V21R	**0.103**	0.104	0.104	**0.116**	0.113	0.114
V22R	0.198	**0.183**	**0.255**	0.209	0.201	0.191
V23R	×	×	0.707	0.432	0.345	**0.238**

Table 2: **[TUM monoVO]** Median alignment errors e_{align} [m] over 10 runs on each sequence. The best and the worst results are in blue and red, respectively. The mark × indicates that more than half of the total runs lost tracking due to challenging scene textures, camera motions or illumination changes.

	ORB VO	ORB SLAM	DSO default	DSO reduced	Ours VO	Ours SLAM
1	0.744	**1.063**	**0.526**	0.799	0.885	0.860
2	**19.08**	**56.03**	0.570	**0.487**	1.412	0.908
3	0.731	**0.510**	3.393	**4.747**	3.283	0.529
4	**2.458**	2.234	0.689	0.770	**0.677**	0.764
5	×	×	**1.718**	2.004	1.921	1.880
6	1.328	0.037	0.808	1.328	**1.358**	**0.522**
7	**0.406**	0.454	0.604	**0.969**	0.926	0.931
8	**596.9**	**425.0**	**0.353**	0.483	0.600	0.625
9	0.534	**0.393**	0.617	**1.069**	1.057	0.994
10	**1.632**	1.312	**0.289**	0.427	0.400	0.564
11	**1.086**	**0.152**	0.592	0.847	0.961	0.172
12	**2.533**	1.646	**0.597**	0.937	1.000	0.812
13	2.232	**2.366**	1.388	1.312	**1.130**	1.216
14	**10.05**	**20.63**	0.718	**0.666**	0.863	0.847
15	1.241	**1.498**	**0.832**	1.310	1.340	1.323
16	0.832	**0.086**	0.514	0.739	**0.931**	0.603
17	1.777	**0.417**	2.300	**2.684**	2.681	1.300
18	5.381	×	1.579	2.135	1.986	**0.411**
19	**6.493**	5.621	1.898	3.534	3.673	**1.232**
20	1.266	**0.273**	0.754	1.095	**1.331**	0.295
21	×	×	4.219	2.707	2.504	**0.230**
22	×	×	3.954	4.785	4.732	**1.131**
23	**5.140**	0.161	0.478	2.344	2.538	**0.061**
24	**5.731**	0.263	0.297	0.361	0.275	**0.156**
25	1.657	0.099	0.821	**2.275**	2.070	**0.081**

	ORB VO	ORB SLAM	DSO default	DSO reduced	Ours VO	Ours SLAM
26	**5.386**	0.323	3.336	4.756	3.705	**0.241**
27	**3.810**	**0.198**	0.965	1.496	1.530	0.321
28	**9.517**	0.433	2.185	2.309	2.186	**0.121**
29	**8.255**	0.325	0.430	1.232	1.311	**0.039**
30	1.278	0.046	0.663	**1.919**	0.821	**0.037**
31	×	×	0.593	0.770	0.738	**0.078**
32	**3.094**	0.077	0.320	0.816	0.876	**0.070**
33	**2.726**	**0.074**	1.449	1.730	1.632	0.078
34	2.687	**0.105**	0.884	**9.434**	**22.54**	0.355
35	**7.046**	0.705	0.578	2.620	2.758	**0.290**
36	1.465	0.590	**5.851**	2.898	3.661	**0.528**
37	0.442	**0.103**	0.379	0.685	**0.709**	0.180
38	×	×	0.768	1.662	1.966	**0.171**
39	**32.58**	0.314	1.293	2.221	2.896	**0.210**
40	×	×	1.805	1.185	1.007	**0.164**
41	×	×	0.897	0.542	0.449	**0.195**
42	×	0.547	0.889	1.293	1.158	**0.084**
43	1.388	0.181	0.458	**1.963**	1.902	**0.102**
44	1.388	**0.065**	0.552	**1.780**	1.713	0.120
45	**4.085**	**0.204**	1.258	2.442	2.693	0.227
46	**12.09**	1.345	0.608	1.415	1.605	**0.403**
47	**12.68**	0.252	1.519	2.019	2.234	**0.122**
48	**3.753**	0.208	1.089	3.314	2.621	**0.060**
49	**11.40**	**0.164**	×	1.023	1.172	0.195
50	**137.2**	4.557	**0.771**	1.830	1.999	1.586

2 Derivation of Equation (7) and (8)

Let subscript w, i and j denote the world reference frame, the previous and the current keyframe, respectively. Also, let subscript D and F denote the direct and the feature-based module, respectively. Suppose that two modules have the same camera axes convention. Now, let the relative scale between the two modules be s, such that for an arbitrary point $\mathbf{p}$,

$$s\mathbf{p}_{i|\mathrm{D}} = \mathbf{p}_{i|\mathrm{F}}, \tag{1}$$
$$s\mathbf{p}_{j|\mathrm{D}} = \mathbf{p}_{j|\mathrm{F}} = \mathbf{R}_{ji|\mathrm{F}}\mathbf{p}_{i|\mathrm{F}} + \mathbf{t}_{ji|\mathrm{F}}. \tag{2}$$

Then, we have

$$s\mathbf{p}_{j|\mathrm{D}} = s(\mathbf{R}_{ji|\mathrm{D}}\mathbf{p}_{i|\mathrm{D}} + \mathbf{t}_{ji|\mathrm{D}}) \tag{3}$$
$$= \mathbf{R}_{ji|\mathrm{D}}(s\mathbf{p}_{i|\mathrm{D}}) + s\mathbf{t}_{ji|\mathrm{D}} \tag{4}$$
$$\overset{(1)}{=} \mathbf{R}_{ji|\mathrm{D}}\mathbf{p}_{i|\mathrm{F}} + s\mathbf{t}_{ji|\mathrm{D}} \tag{5}$$

Subtracting (5) from (2) leads to

$$(\mathbf{R}_{ji|\mathrm{F}} - \mathbf{R}_{ji|\mathrm{D}})\mathbf{p}_{i|\mathrm{F}} + \mathbf{t}_{ji|\mathrm{F}} - s\mathbf{t}_{ji|\mathrm{D}} = \mathbf{0}. \tag{6}$$

For (6) to hold for $\mathbf{p}_{i|\mathrm{F}} = \mathbf{0}$:

$$\mathbf{t}_{ji|\mathrm{F}} = s\mathbf{t}_{ji|\mathrm{D}}. \tag{7}$$

Likewise, for (6) to hold for $\mathbf{p}_{i|\mathrm{F}} \neq \mathbf{0}$,

$$\mathbf{R}_{ji|\mathrm{F}} = \mathbf{R}_{ji|\mathrm{D}}. \tag{8}$$

On the other hand, we know that

$$\mathbf{p}_{i|\mathrm{D}} = \mathbf{R}_{iw|\mathrm{D}}\mathbf{p}_{w|\mathrm{D}} + \mathbf{t}_{iw|\mathrm{D}}, \tag{9}$$
$$\mathbf{p}_{j|\mathrm{D}} = \mathbf{R}_{jw|\mathrm{D}}\mathbf{p}_{w|\mathrm{D}} + \mathbf{t}_{jw|\mathrm{D}}. \tag{10}$$

Hence,

$$\mathbf{p}_{w|\mathrm{D}} \overset{(9)}{=} (\mathbf{R}_{iw|\mathrm{D}})^{-1}(\mathbf{p}_{i|\mathrm{D}} - \mathbf{t}_{iw|\mathrm{D}}) = (\mathbf{R}_{iw|\mathrm{D}})^{\mathrm{T}}(\mathbf{p}_{i|\mathrm{D}} - \mathbf{t}_{iw|\mathrm{D}}) = (\mathbf{R}_{iw|\mathrm{D}})^{\mathrm{T}}\mathbf{p}_{i|\mathrm{D}} - (\mathbf{R}_{iw|\mathrm{D}})^{\mathrm{T}}\mathbf{t}_{iw|\mathrm{D}}, \tag{11}$$
$$\mathbf{p}_{w|\mathrm{D}} \overset{(10)}{=} (\mathbf{R}_{jw|\mathrm{D}})^{-1}(\mathbf{p}_{j|\mathrm{D}} - \mathbf{t}_{jw|\mathrm{D}}). \tag{12}$$

Equating the right-hand sides of (11) and (12), we get

$$\mathbf{p}_{j|\mathrm{D}} = \mathbf{R}_{jw|\mathrm{D}}(\mathbf{R}_{iw|\mathrm{D}})^{\mathrm{T}}\mathbf{p}_{i|\mathrm{D}} - \mathbf{R}_{jw|\mathrm{D}}(\mathbf{R}_{iw|\mathrm{D}})^{\mathrm{T}}\mathbf{t}_{iw|\mathrm{D}} + \mathbf{t}_{jw|\mathrm{D}}. \tag{13}$$

This shows that

$$\mathbf{R}_{ij|\mathrm{D}} = \mathbf{R}_{jw|\mathrm{D}}(\mathbf{R}_{iw|\mathrm{D}})^{\mathrm{T}} \tag{14}$$

and

$$\mathbf{t}_{ij|\mathrm{D}} = -\mathbf{R}_{jw|\mathrm{D}}(\mathbf{R}_{iw|\mathrm{D}})^{\mathrm{T}}\mathbf{t}_{iw|\mathrm{D}} + \mathbf{t}_{jw|\mathrm{D}} \tag{15}$$
$$\overset{(14)}{=} -\mathbf{R}_{ij|\mathrm{D}}\mathbf{t}_{iw|\mathrm{D}} + \mathbf{t}_{jw|\mathrm{D}}. \tag{16}$$

Finally, substituting (14) and (16) into (7) and (8) leads to the two equations we wanted to prove in the main paper.