Generalized LOAM: LiDAR Odometry Estimation with Trainable Local Geometric Features – Supplementary Material –

Kohei Honda[1], Kenji Koide[2], Masashi Yokozuka[2], Shuji Oishi[2], and Atsuhiko Banno[2]

I. Compalison with Baseline Methods

We compared the proposed method with four lidar odometry estimation algorithms; LOAM [1], LeGO-LOAM [2], (plane-to-plane) GICP [3], and SuMa [4]. LOAM and GICP are implemented by ourselves, and share common processes (e.g., preprocessing and pose optimization) and parameters for a fair comparison. LeGO-LOAM and SuMa show the results with parameters tuned by Optuna [5] in [6]. Note that all algorithms are compared using only the front-end part (loop closure is disabled).

Table I and II show the translational and rotational RTEs compared with all baseline methods. Table III shows the results of the comparison with different scan matching schemes. Two types of scan matching methods were implemented: Frame-to-Frame (F-F) and Frame-to-Model (F-M) matching. F-F matches consecutive frames frame-by-frame and estimates the sensor trajectory by accumulating the matching results. F-M matches the current scan and a model that is created by accumulating past frames. Our proposed two MLPs were trained for F-F scan matching.

References

[1] J. Zhang and S. Singh, "Loam: Lidar odometry and mapping in real-time," in Robotics: Science and Systems, vol. 2, no. 9, 2014, pp. 1–9.

[2] T. Shan and B. Englot, "Lego-loam: Lightweight and ground-optimized lidar odometry and mapping on variable terrain," in International Conference on Intelligent Robots and Systems, 2018, pp. 4758–4765.

[3] A. Segal, D. Haehnel, and S. Thrun, "Generalized-icp." in Robotics: science and systems, vol. 2, no. 4, 2009, p. 435.

[4] J. Behley and C. Stachniss, "Efficient surfel-based slam using 3d laser range data in urban environments," in Robotics: Science and Systems, vol. 2018, 2018, p. 59.

[5] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, "Optuna: A next-generation hyperparameter optimization framework," in International Conference on Knowledge Discovery and Data Mining, 2019.

[6] K. Koide, M. Yokozuka, S. Oishi, and A. Banno, "Adaptive hyperparameter tuning for black-box lidar odometry," in International Conference on Intelligent Robots and Systems. IEEE, 2021, pp. 7708–7714.

[7] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in Conference on Computer Vision and Pattern Recognition, 2012, pp. 3354–3361.

[1]Kohei Honda is with the Department of Mechanical Systems Engineering, Graduate School of Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi, Japan, honda.kohei.b0@s.mail.nagoya-u.ac.jp

[2]Kenji Koide, Masashi Yokozuka, Shuji Oishi, and Atsuhiko Banno are with the Department of Information Technology and Human Factors, the National Institute of Advanced Industrial Science and Technology, Umezono 1-1-1, Tsukuba, 3050061, Ibaraki, Japan , {k.koide, yokotsuka-masashi, shuji.oishi, atsuhiko.banno }@aist.go.jp

TABLE I

AVERAGE TRANSLATIONAL RTES [%] FOR KITTI DATASET [7]

Seq. ID	Training dataset 01	03	05	07	09	Avg.	Test dataset 00	02	04	06	08	10	Avg.
LOAM	2.33	4.18	5.34	2.02	7.35	4.24	3.51	10.8	3.40	1.19	4.66	4.08	4.61
LeGO-LOAM	3.07	1.63	1.02	1.02	1.29	1.61	1.87	1.83	1.33	1.09	1.76	1.83	1.62
GICP	3.18	1.06	0.90	0.57	1.26	1.40	1.11	1.54	1.07	0.72	1.15	1.40	1.17
SuMa	5.59	1.14	0.87	0.66	0.86	1.82	0.84	1.27	0.56	0.60	1.43	1.73	1.07
Generalized LOAM (PN++)	3.16	1.04	0.82	0.53	1.15	1.34	1.04	1.45	1.01	0.68	1.10	1.24	1.09
Generalized LOAM (RN)	2.91	0.95	0.76	0.53	1.00	1.23	0.94	1.33	0.87	0.64	1.06	0.99	0.97

[1] Red and blue respectively indicate the first and second best results.
[2] PN++ and RN mean PointNet++ and RandLA-Net, respectively.
[3] LOAM and GICP are implemented by ourselves. LeGO-LOAM and SuMa show the results with parameters tuned by Optuna in [6].

TABLE II

AVERAGE ROTATIONAL RTES [°/ m] FOR KITTI DATASET [7]

Seq. ID	Training dataset 01	03	05	07	09	Avg.	Test dataset 00	02	04	06	08	10	Avg.
LOAM	0.50	1.87	2.30	1.24	2.33	1.65	1.48	3.41	1.50	0.41	1.89	2.06	1.79
LeGO-LOAM	0.95	1.03	0.59	0.75	0.71	0.81	0.92	0.81	0.95	0.65	0.80	0.85	0.83
GICP	0.55	0.58	0.45	0.45	0.55	0.52	0.52	0.55	0.67	0.36	0.47	0.61	0.53
SuMa	0.83	0.85	0.47	0.62	0.50	0.65	0.51	0.59	0.23	0.27	0.63	0.86	0.52
Generalized LOAM (PN++)	0.51	0.55	0.41	0.41	0.50	0.48	0.49	0.52	0.62	0.35	0.43	0.54	0.49
Generalized LOAM (RN)	0.45	0.51	0.37	0.34	0.43	0.42	0.43	0.46	0.53	0.33	0.40	0.45	0.44

[1] Red and blue respectively indicate the first and second best results.
[2] PN++ and RN mean PointNet++ and RandLA-Net, respectively.
[3] LOAM and GICP are implemented by ourselves. LeGO-LOAM and SuMa show the results with parameters tuned by Optuna in [6].

TABLE III

COMPARISON OF THE TYPES OF SCAN MATCHING

<table>
<tr><th rowspan="3">Seq.</th><th colspan="4">LOAM</th><th colspan="4">GICP</th><th colspan="4">Generalized LOAM (PN++)[3]</th><th colspan="4">Generalized LOAM (RN)[3]</th></tr>
<tr><th colspan="2">t_{RTE}[1] [%]</th><th colspan="2">r_{RTE}[1] [°/m]</th><th colspan="2">t_{RTE} [%]</th><th colspan="2">r_{RTE} [°/m]</th><th colspan="2">t_{RTE} [%]</th><th colspan="2">r_{RTE} [°/m]</th><th colspan="2">t_{RTE} [%]</th><th colspan="2">r_{RTE} [°/m]</th></tr>
<tr><th>F-F[2]</th><th>F-M[2]</th><th>F-F</th><th>F-M</th><th>F-F</th><th>F-M</th><th>F-F</th><th>F-M</th><th>F-F</th><th>F-M</th><th>F-F</th><th>F-M</th><th>F-F</th><th>F-M</th><th>F-F</th><th>F-M</th></tr>
<tr><td colspan="17">Training dataset</td></tr>
<tr><td>01</td><td>19.46</td><td>2.33</td><td>3.21</td><td>0.50</td><td>3.16</td><td>3.18</td><td>0.66</td><td>0.55</td><td>2.92</td><td>3.16</td><td>0.63</td><td>0.51</td><td>2.64</td><td>2.91</td><td>0.57</td><td>0.45</td></tr>
<tr><td>03</td><td>24.82</td><td>4.18</td><td>10.38</td><td>1.87</td><td>1.15</td><td>1.06</td><td>0.68</td><td>0.58</td><td>1.10</td><td>1.04</td><td>0.62</td><td>0.55</td><td>1.08</td><td>0.95</td><td>0.75</td><td>0.51</td></tr>
<tr><td>05</td><td>21.95</td><td>5.34</td><td>9.92</td><td>2.30</td><td>0.91</td><td>0.90</td><td>0.47</td><td>0.45</td><td>0.84</td><td>0.82</td><td>0.41</td><td>0.41</td><td>0.78</td><td>0.76</td><td>0.48</td><td>0.37</td></tr>
<tr><td>07</td><td>12.95</td><td>2.02</td><td>8.43</td><td>1.24</td><td>0.67</td><td>0.57</td><td>0.44</td><td>0.45</td><td>0.67</td><td>0.53</td><td>0.41</td><td>0.41</td><td>0.66</td><td>0.53</td><td>0.49</td><td>0.34</td></tr>
<tr><td>09</td><td>27.83</td><td>7.35</td><td>9.39</td><td>2.33</td><td>1.15</td><td>1.26</td><td>0.46</td><td>0.55</td><td>1.10</td><td>1.15</td><td>0.44</td><td>0.50</td><td>0.97</td><td>1.00</td><td>0.36</td><td>0.43</td></tr>
<tr><td>Mean</td><td>21.40</td><td>4.24</td><td>8.26</td><td>1.65</td><td>1.41</td><td>1.40</td><td>0.54</td><td>0.52</td><td>1.33</td><td>1.34</td><td>0.50</td><td>0.48</td><td>1.23</td><td>1.23</td><td>0.53</td><td>0.42</td></tr>
<tr><td colspan="17">Test dataset</td></tr>
<tr><td>00</td><td>16.87</td><td>3.51</td><td>7.73</td><td>1.48</td><td>0.98</td><td>1.11</td><td>0.49</td><td>0.52</td><td>1.06</td><td>1.04</td><td>0.48</td><td>0.49</td><td>0.98</td><td>0.94</td><td>0.52</td><td>0.43</td></tr>
<tr><td>02</td><td>22.55</td><td>10.80</td><td>7.43</td><td>3.41</td><td>1.66</td><td>1.54</td><td>0.54</td><td>0.55</td><td>1.51</td><td>1.45</td><td>0.49</td><td>0.52</td><td>1.43</td><td>1.33</td><td>0.48</td><td>0.46</td></tr>
<tr><td>04</td><td>31.74</td><td>3.40</td><td>6.72</td><td>1.50</td><td>1.88</td><td>1.07</td><td>1.00</td><td>0.67</td><td>1.82</td><td>1.01</td><td>0.92</td><td>0.62</td><td>1.51</td><td>0.87</td><td>0.84</td><td>0.53</td></tr>
<tr><td>06</td><td>8.00</td><td>1.19</td><td>2.59</td><td>0.41</td><td>0.69</td><td>0.72</td><td>0.41</td><td>0.36</td><td>0.67</td><td>0.68</td><td>0.39</td><td>0.35</td><td>0.65</td><td>0.64</td><td>0.41</td><td>0.33</td></tr>
<tr><td>08</td><td>21.88</td><td>4.66</td><td>9.06</td><td>1.89</td><td>1.27</td><td>1.15</td><td>0.53</td><td>0.47</td><td>1.22</td><td>1.10</td><td>0.52</td><td>0.43</td><td>1.30</td><td>1.06</td><td>0.49</td><td>0.40</td></tr>
<tr><td>10</td><td>3.71</td><td>4.08</td><td>1.37</td><td>2.06</td><td>0.98</td><td>1.40</td><td>0.48</td><td>0.61</td><td>0.89</td><td>1.24</td><td>0.37</td><td>0.54</td><td>0.89</td><td>0.99</td><td>0.48</td><td>0.45</td></tr>
<tr><td>Mean</td><td>17.46</td><td>4.61</td><td>5.82</td><td>1.79</td><td>1.24</td><td>1.17</td><td>0.58</td><td>0.53</td><td>1.19</td><td>1.09</td><td>0.53</td><td>0.49</td><td>1.13</td><td>0.97</td><td>0.54</td><td>0.44</td></tr>
</table>

[1] We adopt the RTE as a metric for evaluating pose estimation provided by the KITTI dataset [7]. t_{RTE} and r_{RTE} are the translation and rotation errors for all possible subsequences of length (100,...,800) meters.
[2] F-F and F-M mean Frame-to-Frame matching and Frame-to-Model matching, respectively.
[3] PN++ and RN mean PointNet++ and RandLA-Net, respectively.
All methods are implemented by ourselves, and common parameters are standardized.