Appendix

Narrowing your FOV with SOLiD: Spatially Organized and Lightweight Global Descriptor for FOV-constrained LiDAR Place Recognition

June 30, 2024

Hogyun Kim[1], Jiwon Choi[1], Taehu Sim[2], Giseop Kim[3], and Younggun Cho[1†]
hg.kim@inha.edu, jiwon2@inha.edu, xogn2359@gmail.com, giseop.kim@naverlabs.com, yg.cho@inha.ac.kr

[1]Inha University, Electrical and Computer Engineering
[2]Inha University, Electrical Engineering
[3]Vision Group of NAVER LABS, Seongnam, Gyeonggi-do, 13561, Republic of Korea

1 Full View (360○) Results

We evaluate Recall@1, AUC score, F1 max score, and time as shown in Table 1. Ring++ outperforms in all performance evaluation metrics, followed by LoGG3D and ours. However, Ring++ lags in terms of speed, making it unsuitable for real-time SLAM or multi-robot systems. On the other hand, our method can achieve both speed and performance. Especially in terms of AUC score per unit weight, as represented in Fig. 1 of this paper, our method also overwhelmingly outperforms the results of the full 360-degree view.

Table 1: Peformance Evaluation for Full View LiDAR Place Recognition

Method	Size [bytes]↓	Sequence											
		00 (4541 frames)				02 (4661 frames)				05 (2761 frames)			
		Recall@1 ↑	AUC ↑	F1 max ↑	Time [Hz] ↑	Recall@1 ↑	AUC ↑	F1 max ↑	Time [Hz] ↑	Recall@1 ↑	AUC ↑	F1 max ↑	Time [Hz] ↑
Hist. (C)	**928**	0.817	0.983	0.808	0.58	0.433	0.963	0.427	0.58	0.544	**0.976**	0.598	0.58
Logg3D (G)	2176	**0.918**	0.982	0.879	15.8	**0.881**	**0.978**	**0.801**	15.8	**0.912**	0.965	**0.836**	16.5
M2DP (C)	1664	0.829	0.987	0.839	3.31	0.677	0.968	0.706	3.31	0.625	0.965	0.673	3.34
SC (C)	10016	0.893	0.979	0.876	3.04	0.794	0.946	0.778	3.04	0.742	0.968	0.785	3.09
OT (G)	1152	0.904	0.981	0.863	44.8	0.782	0.962	0.759	44.6	0.857	0.960	0.750	47.6
CVT-NET (G)	3200	0.845	0.988	0.875	8.08	0.811	0.954	0.807	8.07	0.702	0.975	0.753	8.17
Ring++ (G)	345728	0.954	0.997	0.937	0.44	0.959	0.988	0.877	0.43	0.962	0.988	0.938	0.72
Ours (C)	448	0.863	**0.996**	**0.900**	**20.8**	0.765	0.965	0.799	**20.8**	0.776	0.961	0.767	**22.1**

* **first rank** is blue and **second rank** is bold.

2 Downsampling

We evaluated the performance through an ablation study, as shown in the Table 2.

Table 2: Performance Evaluation according to voxel size of downsampling in KITTI 00

Method	Voxel Size	FOV								
		60°			120°			180°		
		Recall@1 ↑	AUC ↑	F1 max ↑	Recall@1 ↑	AUC ↑	F1 max ↑	Recall@1 ↑	AUC ↑	F1 max ↑
Ours (Modified)	0.1m	0.765	0.971	0.762	0.791	0.983	0.807	0.793	0.982	0.799
Ours	0.5m	0.800	0.986	0.830	0.822	0.994	0.867	0.840	0.994	0.868
Ours (Modified)	1.0m	0.798	0.978	0.778	0.830	0.987	0.822	0.840	0.987	0.826

* **first rank** is blue.

We found that 0.5m and 1.0m perform the best, in that order. Note that the proposed approach performs similarly under different voxel resolutions (0.5m and 1m). We also tested our method with a minimal (0.1m) voxel size. These cases yielded weaker performance than the best cases, but the metrics did not drop significantly.

3 RAC

We can create a Radial-Azimuthal bin (i.e. Bird-eye view (BEV)) with elevational projection to utilize *RAC (Radial-Azimuthal Counter)*. If we use *RAC* for a place recognition problem, we create an *IAV (Implicit Azimuth Counter)* or *IRV (Implicit Range Counter)*. Similar to *R-SOLiD* with rotation invariance for loop detection, the matrix product between *RAC* and *IAV* generates *R-RAC*. The experiments were conducted with the Aeva LiDAR, which has the smallest vertical FOV, to ensure minimal impact of vertical information. Table 3 shows that vertical information is meaningful for representing narrow FOV.

Table 3: Peformance Evaluation in HeLiPR with Aeva LiDAR

Method	Size [bytes]↓	Town01-Town02			RoundAbout01-RoundAbout02			KAIST04-KAIST05		
		Recall@1↑	AUC↑	F1 max↑	Recall@1↑	AUC↑	F1 max↑	Recall@1↑	AUC↑	F1 max↑
R-SOLiD using RAC	448	0.733	0.907	0.733	0.697	0.830	0.605	0.872	0.936	0.877
R-SOLiD using REC	448	**0.781**	**0.934**	**0.783**	**0.729**	**0.886**	**0.646**	**0.882**	**0.955**	**0.886**

* **first rank** is blue.

4 3-D matching graph in HeLiPR datasets

We draw a 3-D graph when the F1 score is maximized and when the precision is 100%. For methods where precision does not reach 100%, we base the graph on the highest precision achieved. In the Aeva sequence, the proposed method exhibits the densest matching performance, while in the Avia sequence, RING++ demonstrates the densest matching performance.

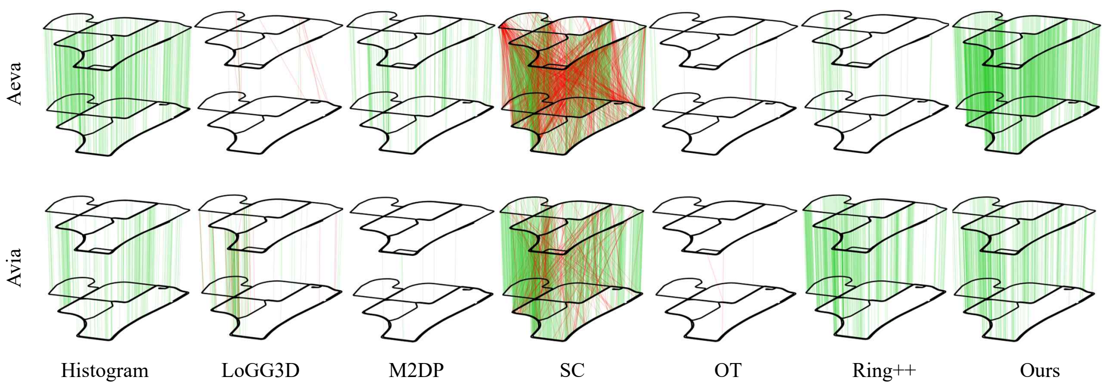

Figure A: 3-D matching pair graph when each descriptor's precision is highest in *HeLiPR* dataset (source: *KAIST04*, target: *KAIST05*). Green represents true matching pairs, and red represents false matching pairs.

5 Constant IEV

To check the impact of IEV, we compare the results when using constant IEV as the ablation study.

Table 4: Peformance Evaluation in KITTI 05 with Constant IEV

Method	Size [bytes]↓	FOV 60°			FOV 120°			FOV 180°		
		Recall@1 ↑	AUC ↑	F1 max ↑	Recall@1 ↑	AUC ↑	F1 max ↑	Recall@1 ↑	AUC ↑	F1 max ↑
Ours-Constant IEV	448	0.635	0.988	0.731	0.652	0.987	0.732	0.657	0.965	0.693
Ours	448	**0.639**	**0.992**	**0.732**	**0.656**	**0.989**	**0.737**	**0.659**	**0.974**	**0.710**

* **first rank** is blue.