MulRan: Multimodal Range Dataset for Urban Place Recognition

Giseop Kim[1], Yeong Sang Park[1], Younghun Cho[1], Jinyong Jeong[1], and Ayoung Kim[1*]

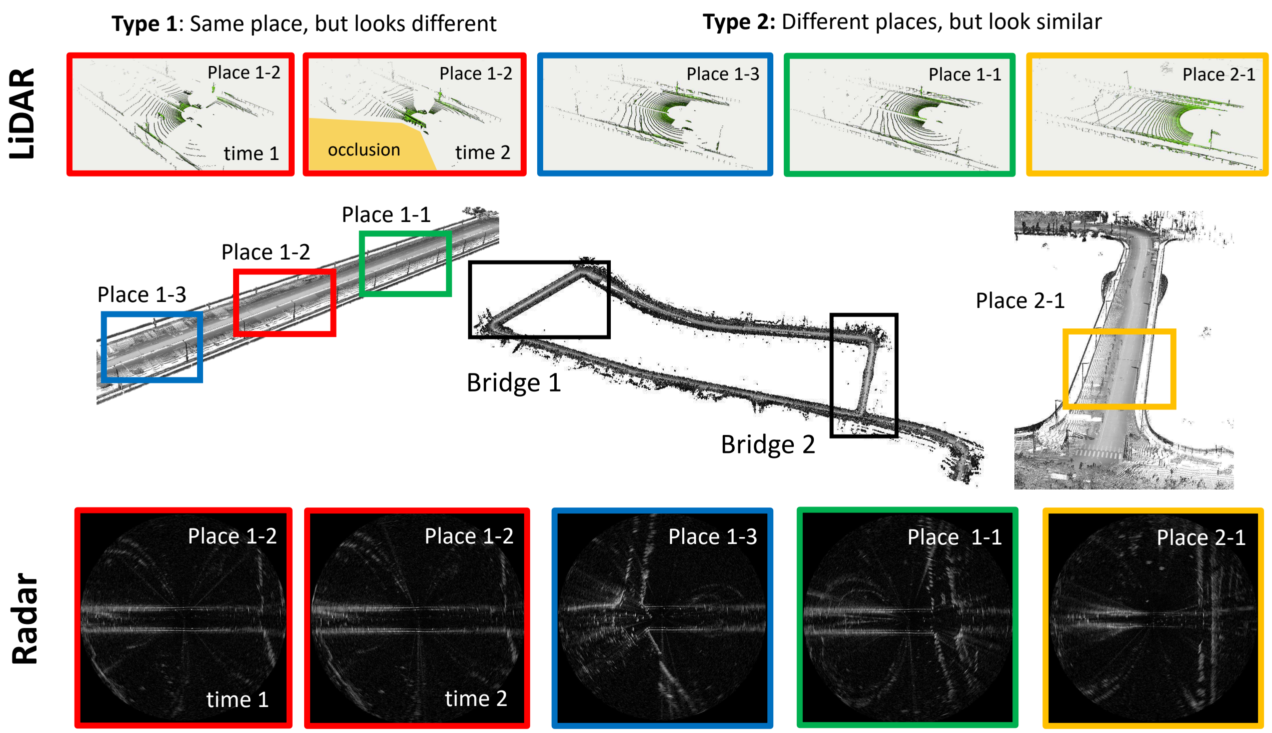


Fig. 1: (Top row) LiDAR's structural perception aliasing. These examples are from our dataset's `Riverside 01` sequence. Urban sites contain places with few structural features (e.g., bridges) and these types of places are highly repetitive. From this reason, two types of structural perception aliasing frequently occur, resulting in catastrophic place recognition false alarms. (Bottom row) Radar has a longer perceptible range and is less susceptible to occlusion than LiDAR. The examples at the bottom show that radar suffers less from the aforementioned two types of perception aliasing but has other inherent challenges, such as noise and artifacts. Through this dataset, we aim to provide a dataset for structural place recognition studies using those range measurements.

Abstract— **This paper introduces a multimodal range dataset namely for radio detection and ranging (radar) and light detection and ranging (LiDAR) specifically targeting the urban environment. By extending our workshop paper [1] to a larger scale, this dataset focuses on the range sensor-based place recognition and provides 6D baseline trajectories of a vehicle for place recognition ground truth. Provided radar data support both raw-level and image-format data, including a set of time-stamped 1D intensity arrays and 360° polar images, respectively. In doing so, we provide flexibility between raw data and image data depending on the purpose of the research. Unlike existing datasets, our focus is at capturing both temporal and structural diversities for range-based place recognition research. For evaluation, we applied and validated that our previous location descriptor and its search algorithm [2] are highly effective for radar place recognition method. Furthermore, the result shows that radar-based place recognition outperforms LiDAR-based one exploiting its longer-range measurements. The dataset is available from https://sites.google.com/view/mulran-pr**

[1]G. Kim, Y. S. Park, Y. Cho, J. Jeong and A. Kim are with the Department of Civil and Environmental Engineering, KAIST, Daejeon, S. Korea `[paulgkim, pys0728k, lucascho, jjy0923, ayoungk]@kaist.ac.kr`

This work was supported by [Deep Learning based Camera and LIDAR SLAM] project funded by Naver Labs Corporation and by the National Research Foundation of Korea (NRF) grant (NRF-2019K2A9A1A06070173).

I. Introduction

Place recognition has long been an important problem for robotics application. So far most of the place recognition approaches focus on vision-based methods. Visual place recognition, or appearance simultaneous localization and mapping (SLAM), solves for the matching image given a query image in the stream. For instance, FAB-MAP [11] solved for a large scale appearance SLAM via probabilistic formulation. More recently since the advent of the bag-of-words representation of an image, visual localization using visual words has been introduced in DBoW [12].

TABLE I: Dataset comparison with respect to range sensor configuration and fitness for place recognition study

Datasets		Freiburg [3]	Ford Campus [4]	KITTI [5]	NCLT [6]	Complex Urban [7]	nuScenes [8]	Marulan [9]	Oxford Radar Robotcar [10]	**Ours (MulRan)**
LiDAR	3D Horizontal	✓	✓	✓	✓		✓		✓	✓
	2D Push-broom				✓	✓		✓	✓	
	3D Tilted					✓				
Radar	Scanning						△ (not for PR)	✓ (short: 40m)	✓	✓
Structural Diversity			★	★★	★	★★★	★★	★	★	★★★
Temporal diversity					★★★	★★			★★	★★
Spatial Scale of a Sequence		★	★	★★	★★	★★★	★★	★	★★	★★★
Loop	Frequency	★	★	★★	★★	★★	★	★	★★	★★★
	Reverse			★	★	★★			★	★★★

Among range-measuring sensors, LiDAR has been widely adopted in robotics including autonomous car research. The robustness to illumination change and measurement precision has allowed many LiDAR-based methods to outperform vision-based methods [13]. Recently, loop-closure detection methods using point cloud descriptors have been introduced in many studies [14, 15, 16].

Radar is another promising range sensor that is robust to occlusions and has a perception range longer than LiDAR, but previous methods have mostly exploited radar for object detection. The angular ambiguity prohibits direct application to point cloud-like measurements when using the radar sensor. Existing radar sensors in aerial application are mainly used for imaging radar on a large scale [17]. On a mobile platform, radar has been used for dynamic object detection [18, 19] as an auxiliary sensor and has not been studied much in SLAM research because the major challenge of using radar arises from sparse, noisy sensor characteristics compared to LiDAR. Recently, however, radar has been getting attention from academia related to SLAM, despite the aforementioned challenges, because radar is robust to occlusions and has long-range capturing capability. In particular, a frequency-modulated continuous-wave (FMCW) scanning radar was proven in [20, 21, 22]; the sensor is useful for robot motion estimation in challenging environmental conditions (e.g., forests and off-road Iceland).

From those recent promising works, outdoor robot navigation with scanning radar can feasibly succeed where LiDAR or cameras fail. However, there are few radar place recognition algorithms [23, 24, 25] and no dataset particularly designed to quantitatively validate the radar's place recognition capability despite the fact that a place recognition method is fundamentally required to complete a full radar SLAM framework.

In this paper, we present a multimodal range measurement dataset including the radar and LiDAR sensor and targeting for urban place recognition. To the best of our knowledge, radar's urban place recognition capability has not been evaluated and few methods exist. To fill this gap, we provide a LiDAR-radar data pair during urban driving as an extended version of our previous work [1], and report some LiDAR and radar place recognition methods' performance over our dataset. Our contribution points are:

- We provide a multi-environment, multi-sesson, and multimodal range (i.e., radar-LiDAR) dataset. Our dataset has both radar and 3D LiDAR (see Table II) and includes multi-session sequences along a repeated trajectory within a changing city with month-level temporal gaps as shown in Table III and Fig. 4.
- Our dataset has various types of revisit events. We deliberately designed sequences' routes to have multiple revisits for the same places (both in-session and multi-session with time gaps) so that the number of queries is enough for scalable place recognition evaluation. Furthermore, our sequences contain monthly revisits so as to capture the temporal diversity of the environment toward long-term robust place recognition.
- As a validation of the dataset, we present an initial ranged-based place recognition evaluation. We applied a place descriptor from our previous work [2] to both radar and LiDAR measurements to prove that the descriptor is applicable for general range measurements. Furthermore, radar outperforms the LiDAR in regards to place recognition, especially when a longer range is preferred (e.g., riverside).

II. Related Work

A. Existing Range Datasets

We summarized existing LiDAR and radar datasets, targeting place recognition studies. For each dataset, the reverse loop detection capability and long-term robustness are listed in Table I.

From KITTI [5] to Complex Urban dataset [7], many datasets include LiDAR measurements; however, there are few available radar datasets in academia. As we pointed out in §I, there is a strong need for more radar datasets with sufficient ground truth, considering the potential of radar for robust navigation. Recently, the nuScenes dataset [8] has provided multiple radar measurements, but these non-scanning and sparse radar data are less adequate for place recognition. Marulan datasets [9] provide 2D Cartesian radar images at each sweep. The image format data are convenient

TABLE II: The sensor specification of the MulRan Dataset

Sensor	Mount type	Manufacturer	Model	Description	No.	Hz	Range
3D LiDAR	Horizontal	Ouster	OS1-64	64 channel, 360° FOV	1	10	120 m
Radar	Horizontal	Navtech	CIR204-H	0.9° and 0.06 m resolution, 360° FOV	1	4	200 m

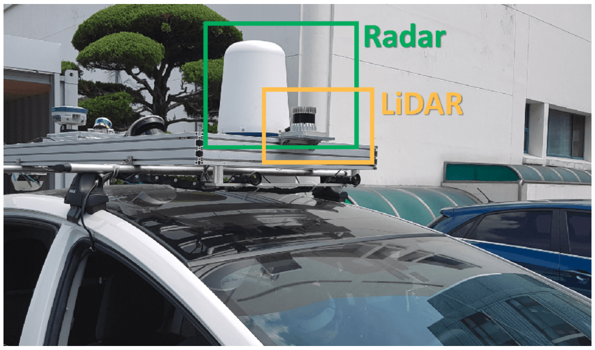


Fig. 2: The sensor system is equipped with a single radar and a 3D LiDAR sensor, providing multimodal range measurements.

because the existing vision algorithm can be directly applied. The radar sensor, however, has a relatively slow sensing speed compared to LiDAR. As a result, the image data timed only with the last angle ray (i.e., without time stamps in measurement at each azimuthal ray; see Fig. 5 for details) have relatively high data distortion due to vehicle movement. Oxford Radar RobotCar dataset [10] extended their dataset [26] by including 3D LiDARs and a scanning radar. Their dataset includes multiple repeated routes over the same area. Having a large overlap, the dataset could be applicable to place recognition; however, the dataset possesses the limited in a single place, while ours is with both temporal (monthly revisits) and structural (multi-city) diversity of the environment.

B. Place Recognition for Range Data

Compared to existing predominantly appearance-based methods [11, 12, 27], structural information-based place recognition could be more beneficial under light-condition variance and long-term scenarios [28]. In particular, LiDAR-based place recognition methods [3, 14, 15, 2, 29, 30] have been widely studied to exploit the accuracy of direct 3D measurement for distant structures (e.g., 100 m). However, existing methods are vulnerable in situations in which a raw input scan could be look different than previously experienced, such as reverse revisit and revisit with a lane-level change [2]. Existing LiDAR datasets in Table I have very few such situations, making algorithm validation to overcome the issue difficult.

In the meantime, considering the long-range capturing capability (e.g., 200 m) of LiDAR and the robustness to environmental variance (e.g., dynamic objects or occlusions), radar poses a great potential for robot missions in complex urban sites [31, 24, 21, 22] and extreme environments (e.g., fog [32]), respectively. Despite these advantages and recent studies, few place recognition methods are proposed [23, 24, 25] and there have been no public quantitative evaluations.

III. System Overview

A. Sensor Configuration

The main objective of our dataset is to boost range sensor-based place recognition researches. To meet this goal, we construct a sensor system with combining a single radar and a single 3D LiDAR as in Fig. 2 so that two different type of range sensors would capturing almost similar field of view (FOV). When installing the OS1-64 LiDAR at the front, however, it loses its FOV approximately 70 ° due to the radar behind it. The detail specification is described in Table II.

B. Sensor Calibration

We first set a vehicle base's coordinate, the same as in [7]. For the vehicle base to OS1-64 3D LiDAR calibration, we use the same extrinsic calibration pipeline for the VLP-16 3D LiDAR in [7]; the roll, pitch, and z are first found using grounds. The x, y, and yaw are then calculated using scene overlaps during a round trip.

For the extrinsic calibration between vehicle base and radar, we calculate the relative transformation between the front LiDAR (OS1-64) and the radar using phase correlation [33] as depicted in Fig. 3. We first make a LiDAR polar image from a single OS1-64 LiDAR scan's bird-eye-view of the same size as the radar polar image and using the same perception range of the radar (i.e., 3360 pixels up to 200 m $\times$ 400 pixels along 360 °). We then convert it to a binary image by assigning 1 to the pixels where the associated LiDAR points exist. By applying phase correlation as in [33], we perform the cross-modal registration and compute the relative transformation. In doing so, only relatives of x, y, and yaw are calculated because the radar sensor does not provide 3D information, but intensities over the horizontal 2D plane.

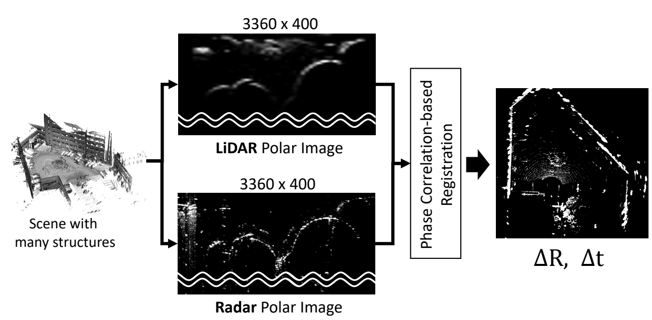


Fig. 3: Extrinsic calibration between the LiDAR and the radar using phase correlation [33].

TABLE III: The sequence details

Target Problem	Sequence Name	Sequence Index (date) 01	02	03	Length (avg)
Online place recognition (in-session localization)	DCC	2019-08-02	2019-08-23	2019-09-03	4.9 km
	KAIST	2019-06-20	2019-08-23	2019-09-02	6.1 km
	Riverside	2019-08-02	2019-08-16	2019-08-23 (reverse)	6.8 km
Global localization (localization between multi-session)	Sejong City	2019-06-20	2019-08-20	2019-08-20 (reverse)	23.4 km

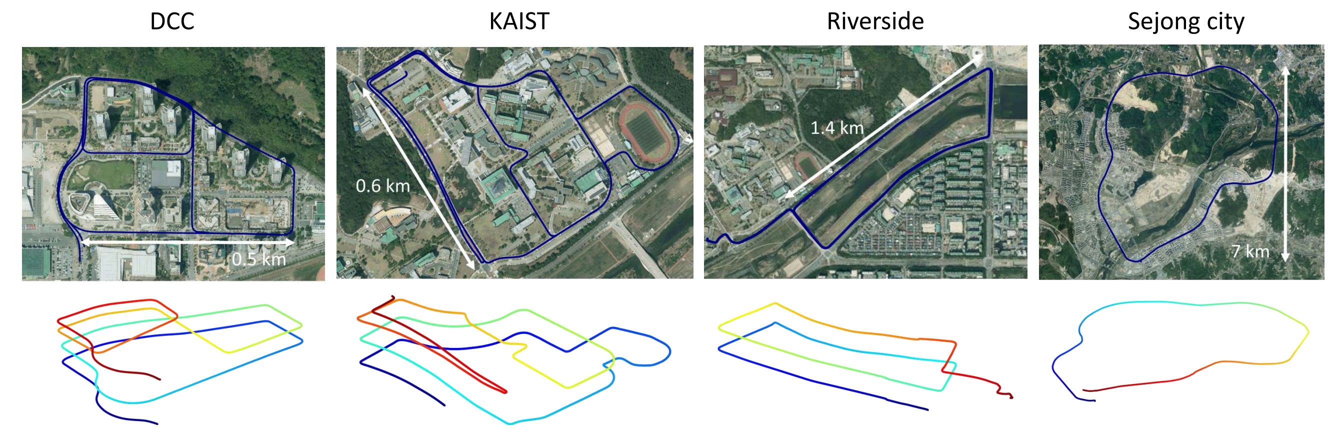
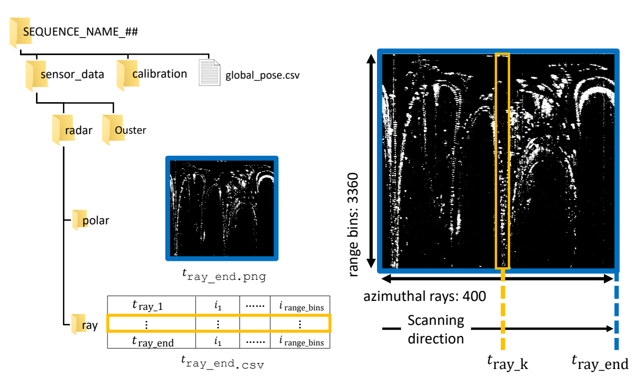


Fig. 4: Aerial-map-overlaid trajectories of the sequence `01`, `01`, `02`, and `01` for each environment, respectively. For each sequence, we tried to maximize the loop-closure candidates to generate enough queries for place recognition evaluation.

IV. MulRan: The Multimodal Range Dataset

Our dataset has four target environments and three sequences for each environment, having repeated trajectories over the same locations from different dates as in Table III. In addition, we provide sequences of a completely reversed route for `Riverside`, and `Sejong`, having one-way driving directions, to support a reverse-revisit detection studies. We first introduce our four target environments and then describe the provided data format.



(a) The data structure of a sequence (b) Information within radar data

Fig. 5: The data structure of a sequence in the MulRan dataset. We provide a single radar scan as two types; a polar image with a single time (blue in Fig. 5(b)) and a set of power-range spectra with their individual time stamps (yellow in Fig. 5(b)).

A. *Target Environments*

In this subsection, we briefly explain why the four experimental environments in Fig. 4 were selected and summarize their characteristics.

1) Dajeon Convention Center (DCC): This environment is relatively smaller than other sequences but is structurally diverse, with a square, narrow roads between high-rise buildings, a mountain, and crossroads.

2) KAIST: The `KAIST` sequence is a campus environment with multiple distinguishable structures and a small number of dynamic objects. For the `DCC` and `KAIST` sequences, we intentionally include multiple reverse revisits and revisits with lane-level translations, which are useful to evaluate rotation-invariant and translation-robust place recognition algorithms; those two requirements are particularly important for long-term autonomy for a mobile robot [29].

3) Riverside: The `Riverside` sequences contain straight runs along a river and two bridges where structural features are repetitive, as shown in Fig. 1.

4) Sejong City: Sejong City in South Korea is an entirely planned and still-developing city [35], as seen in Fig. 4. Therefore, `Sejong` sequences are structurally diverse, ranging from rural areas to urban sites with a variety of structures (e.g., bridge, tunnel, or overpass) as seen in Fig. 6; we believe the `Sejong` sequences help to develop environment-independent algorithms for place recognition. A single `Sejong` sequence has few in-session loop candidates, but we designed `Sejong` sequences for not only online place recognition but global localization problems, such as when a map or prior experiences exist (i.e., multi-session or long-term place recognition [36, 29]). This capability is required

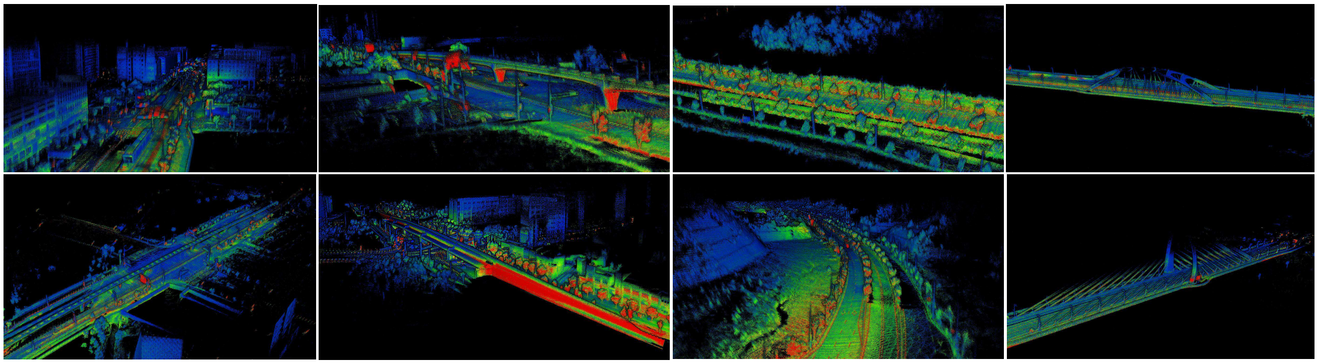
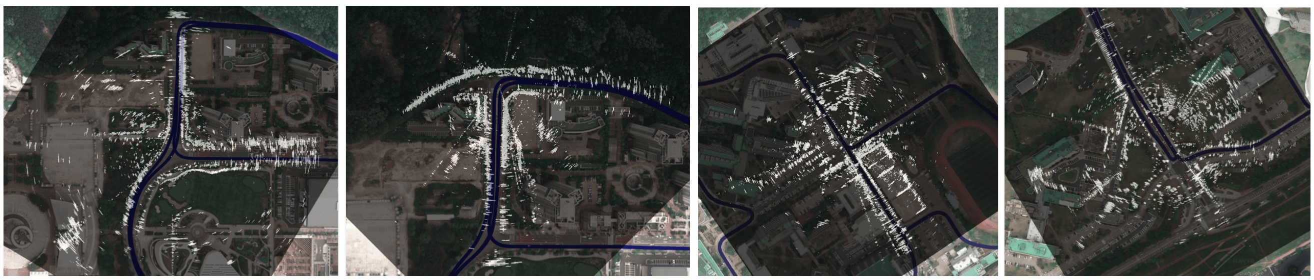


Fig. 6: Structural diversity in a single sequence (`Sejong 01`). These captures are taken from a point cloud map using our 3D LiDAR. The color map represents an intensity value of each point (red is high).



(a) `DCC 01` (b) `KAIST 01`

Fig. 7: A set of examples of a single radar scan on the corresponding aerial map. For clarity of visualization, we show them as binary images by removing pixels of low intensity. Radar can perceive a broad amount of space ($\pm$ 200 m) and suffers less from data loss from occlusions, so it is efficient at capturing structural information such as building and road shapes, which could be important cues for place identification. Multipath effects [34, 20], however, are easily observed (e.g., virtual lines penetrate where no structures exist in Fig. 7(b)); this is caused by the multiple reflections of radar signals within structures and is the crucial limitation of the radar data.

for a changing city such as Sejong, where new structures arise for every month. Therefore, we note that a `Sejong` sequence has almost 100 % more loop candidates than the other date sequence, and providing enough test cases to validate long-term place recognition algorithms.

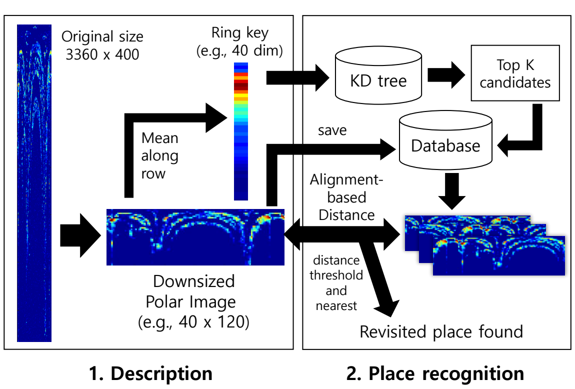


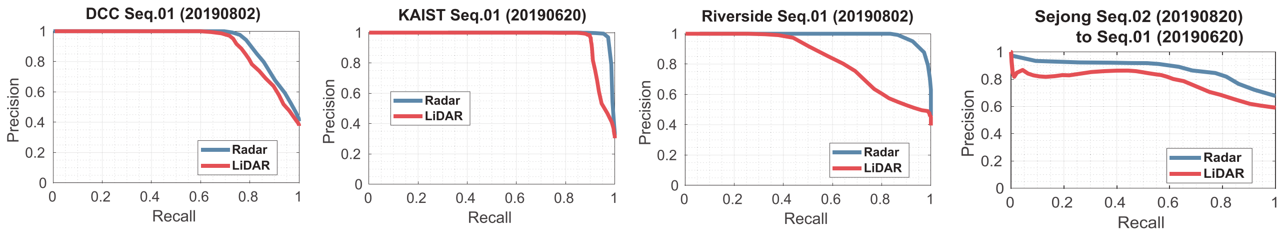
Fig. 8: A pipeline of Scan Context-based [2] radar place recognition method. We use the jet color map for visualization clarity; we use the intensity polar image directly, only downsampling with pixel interpolation.

B. Ground Truth Trajectory

The baselines (i.e., 6D ground truth trajectory) of the aforementioned sequences are made via SLAM using fiber optic gyro (FOG), Virtual Reference Station GPS (VRS-GPS), and Iterated Closest Point (ICP). This process is exactly the same as Complex Urban Dataset's ground truth trajectory-making procedure [7], so please refer to it for details.

C. Data Description and Format

The overall directory structure of a sequence is depicted in Fig. 5. The *Ouster* directory contains `<time stamp.bin>` binary files and a single file corresponds to a single scan, which contains x, y, z, and intensity information for all points in the scan. The format is identical to KITTI [5]'s Velodyne scan file. For the radar data, as details are visualized in Fig. 5, we provide not only polar images as a default format but also a data directory *ray*, which contains individual rays' time for more precise data processing such as motion distortion compensation [37] or 1D signal requiring methods (e.g., constant false alarm rate (CFAR)). Those two types of radar data format have exactly the same information, except for the individual time stamps in a scan. We note again that radar's timestamps are provided both per scan and per



(a) Online place recognition (b) Global localization

Fig. 9: Radar outperforms LiDAR for urban place recognition.

ray whereas LiDAR's are given in per scan.

For the users' convenience, we also provide a Robot Operating System (ROS) player program, which reads individual radar and LiDAR files in the directory and publishes corresponding ROS topics according to their times.

V. Evaluation of MulRan Dataset

In this section, we evaluate place recognition methods for LiDAR and radar of the proposed dataset. To the best of our knowledge, however, there is no common agreement on radar place recognition methods or widely used methods, such as DBoW [12], in the visual domain. Thus, in this paper, we also introduce a practical and effective radar place recognition algorithm called Radar Scan Context (RSC).

A. Radar Place Recognition

We found that our previous work Scan Context [2], which was originally proposed for LiDAR place recognition, would be also a good choice for radar data (i.e., an intensity image in a polar coordinate) for place recognition and even outperforms a LiDAR sensor. The overall place recognition pipeline is described in Fig. 8. We first downsize an original radar polar image to a small image. We use a mean function rather using L_0 norm, which was originally proposed in [2], for making a ring key because a radar image rarely has zero values and mean function still satisfies rotation invariance. The two-phase hierarchical search algorithm and the alignment-based distance function for pairwise comparison are identical to the original paper [2]. The computation time is practically acceptable for real-time navigation as already proven in [2].

B. Experiments

Evaluation metric. We used the precision-recall curve [38] as the evaluation metric and the results are depicted in Fig. 9. We consider a query's answer correct if the top 1 retrieved answer is in $5\,\mathrm{m}$; that is, a robot should find a previous node's index if the robot revisits the place within $5\,\mathrm{m}$ or reject localization to avoid a false-alarm.

Comparison method. For the `LiDAR` method in the result figure (Fig. 9), we used a single OS1-64 3D LiDAR scan as a query input and Scan Context [2] as a method, which showed state-of-the-art LiDAR place recognition performances for a horizontally mounted 3D LiDAR [29]. For both `LiDAR` and `Radar`, the descriptor's shape was 40×120 and 50 candidates are retrieved from ring key tree search.

C. Results

Despite LiDAR exhibiting competitive results in relation to radar for a few environments (i.e., `DCC` and `KAIST`), LiDAR's performance decreased in wide-open spaces with few structural features and many moving objects (i.e., `Riverside`), as argued in Fig. 1 and shown in the right plot of Fig. 9(a).

In comparison to LiDAR, the performance of our proposed Scan Context-based radar place recognition method was consistent in all environments. RSC is robust to multipath effects, as shown in Fig. 7. Because our alignment-based distance [2] first downsamples the large (3333×400), noisy radar image to a small image (in our experiment, we used 40 $\times$ 120) and takes into account the consensus of all rays in a scan, it is less disturbed by partially noisy (from multipath effects) azimuthal rays.

Radar also outperformed LiDAR in global localization. As in Fig. 9(b), we used `Sejong 01` as a map (database), and query data is from a different date, `Sejong 02`. We note that the steep valley point at the left side of `LiDAR`'s curve in Fig. 9(b) occurred in narrow places, where LiDAR's perception ability is almost lost, usually due to dynamic objects' occlusions (e.g., a large bus). On the other hand, radar is not vulnerable to occlusion, so this phenomenon is not found.

VI. Conclusion

In this paper, we release a new dataset, called Multimodal Range (MulRan) Dataset. Our dataset is particularly designed for place recognition studies in urban sites regarding the intentional inclusion of many loop candidates within multi-cities and reverse revisits through multi-session and month-level time gaps. However, we do not restrict MulRan's potential usage for place recognition studies and expect it to be useful for other radar and radar-LiDAR fusion-based robotics research (e.g., SLAM) via accompanying 6D ground truth trajectories. Together with the dataset, we also introduced the Scan Context-based radar place recognition method and showed that radar outperforms LiDAR.

References

[1] Y. S. Park, J. Jeong, Y. Shin, and A. Kim, " Radar Dataset for Robust Localization and Mapping in Urban Environment," in *ICRA Workshop on Dataset Generation and Benchmarking of SLAM Algorithms for Robotics and VR/AR*, Montreal, May. 2019.

[2] G. Kim and A. Kim, " Scan Context: Egocentric Spatial Descriptor for Place Recognition within 3D Point Cloud Map," in *Proc. IEEE/RSJ Intl. Conf. on Intell. Robots and Sys.*, Madrid, Oct. 2018.

[3] B. Steder, G. Grisetti, and W. Burgard, "Robust place recognition for 3d range data based on point features," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2010, pp. 1400–1405.

[4] G. Pandey, J. R. McBride, and R. M. Eustice, "Ford Campus Vision and Lidar Data Set," *Intl. J. of Robot. Research*, vol. 30, no. 13, pp. 1543–1552, 2011.

[5] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the KITTI vision benchmark suite," in *Proc. IEEE Conf. on Comput. Vision and Pattern Recog.*, 2012, pp. 3354–3361.

[6] N. Carlevaris-Bianco, A. K. Ushani, and R. M. Eustice, "University of Michigan North Campus Long-Term Vision and Lidar Dataset," *Intl. J. of Robot. Research*, vol. 35, no. 9, pp. 1023–1035, 2016.

[7] J. Jeong, Y. Cho, Y.-S. Shin, H. Roh, and A. Kim, "Complex Urban Dataset with Multi-level Sensors from Highly Diverse Urban Environments," *Intl. J. of Robot. Research*, vol. 38, no. 6, pp. 642–657, 2019.

[8] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom, "nuscenes: A multimodal dataset for autonomous driving," *arXiv preprint arXiv:1903.11027*, 2019.

[9] T. Peynot, S. Scheding, and S. Terho, "The marulan data sets: Multi-sensor perception in a natural environment with challenging conditions," *Intl. J. of Robot. Research*, vol. 29, no. 13, pp. 1602–1607, 2010.

[10] D. Barnes, M. Gadd, P. Murcutt, P. Newman, and I. Posner, "The Oxford Radar RobotCar Dataset: A Radar Extension to the Oxford RobotCar Dataset," *Proc. IEEE Intl. Conf. on Robot. and Automat.*, May 2020, Accepted. To appear.

[11] M. Cummins and P. Newman, "FAB-MAP: Probabilistic Localization and Mapping in the Space of Appearance," *Intl. J. of Robot. Research*, vol. 27, no. 6, pp. 647–665, 2008.

[12] D. Gálvez-López and J. D. Tardós, "Bags of Binary Words for Fast Place Recognition in Image Sequences," *IEEE Trans. Robot.*, vol. 28, no. 5, pp. 1188–1197, October 2012.

[13] J. Zhang and S. Singh, "Low-drift and real-time lidar odometry and mapping," *Autonomous Robots*, vol. 41, no. 2, pp. 401–416, 2017.

[14] L. He, X. Wang, and H. Zhang, "M2DP: A novel 3D point cloud descriptor and its application in loop closure detection," in *Proc. IEEE/RSJ Intl. Conf. on Intell. Robots and Sys.*, 2016, pp. 231–237.

[15] M. A. Uy and G. H. Lee, "PointNetVLAD: Deep Point Cloud Based Retrieval for Large-Scale Place Recognition," *Proc. IEEE Conf. on Comput. Vision and Pattern Recog.*, 2018, In press.

[16] R. Dubé, A. Cramariuc, D. Dugas, H. Sommer, M. Dymczyk, J. Nieto, R. Siegwart, and C. Cadena, "SegMap: Segment-based mapping and localization using data-driven descriptors," *Intl. J. of Robot. Research*, 2019.

[17] A. Lonnqvist, Y. Rauste, M. Molinier, and T. Hame, "Polarimetric SAR data in land cover mapping in boreal zone," *IEEE Trans. Geosci. and Remote Sensing*, vol. 48, no. 10, pp. 3652–3662, 2010.

[18] H. Cho, Y.-W. Seo, B. V. Kumar, and R. R. Rajkumar, "A multi-sensor fusion system for moving object detection and tracking in urban driving environments," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2014, pp. 1836–1843.

[19] K.-H. Lee, Y. Kanzawa, M. Derry, and M. R. James, "Multi-target track-to-track fusion based on permutation matrix track association," in *Proc. IEEE Intell. Vehicle Symposium*, 2018, pp. 465–470.

[20] S. H. Cen and P. Newman, "Precise ego-motion estimation with millimeter-wave radar under diverse and challenging conditions," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2018.

[21] R. Aldera, D. Martini, M. Gadd, and P. Newman, "Fast Radar Motion Estimation with a Learnt Focus of Attention using Weak Supervision," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2019.

[22] S. H. Cen and P. Newman, "Radar-only ego-motion estimation in difficult settings via graph matching," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2019.

[23] J. Callmer, D. Törnqvist, F. Gustafsson, H. Svensson, and P. Carlbom, "Radar SLAM using visual features," *EURASIP Journal on Advances in Signal Processing*, vol. 2011, no. 1, p. 71, 2011.

[24] K. Werber, J. Klappstein, J. Dickmann, and C. Waldschmidt, "Interesting areas in radar gridmaps for vehicle self-localization," in *2016 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*. IEEE, 2016, pp. 1–4.

[25] E. Takeuchi, A. Elfes, and J. Roberts, "Localization and place recognition using an ultra-wide band (uwb) radar," in *Field and service robotics*. Springer, 2015, pp. 275–288.

[26] W. Maddern, G. Pascoe, C. Linegar, and P. Newman, "1 year, 1000 km: The Oxford RobotCar dataset," *Intl. J. of Robot. Research*, vol. 36, no. 1, pp. 3–15, 2017.

[27] R. Arandjelović, P. Gronat, A. Torii, T. Pajdla, and J. Sivic, "NetVLAD: CNN architecture for weakly supervised place recognition," in *Proc. IEEE Conf. on Comput. Vision and Pattern Recog.*, 2016.

[28] Y. Ye, T. Cieslewski, A. Loquercio, and D. Scaramuzza, "Place recognition in semi-dense maps: Geometric and learning-based approaches," in *Proc. Brit. Mach. Vis. Conf.*, 2017.

[29] G. Kim, B. Park, and A. Kim, "1-Day Learning, 1-Year Localization: Long-term LiDAR Localization using Scan Context Image ," *IEEE Robot. and Automat. Lett.*, vol. 4, no. 2, pp. 1948–1955, 2019.

[30] D. L. Rizzini, F. Galasso, and S. Caselli, "Geometric relation distribution for place recognition," *IEEE Robot.*

and Automat. Lett., vol. 4, no. 2, pp. 523–529, 2019.
[31] F. Schuster, C. G. Keller, M. Rapp, M. Haueis, and C. Curio, "Landmark based radar SLAM using graph optimization," in *Proc. IEEE Intell. Transport. Sys. Conf.* IEEE, 2016, pp. 2559–2564.
[32] Y. S. Park, J. Kim, and A. Kim, " Radar Localization and Mapping for Indoor Disaster Environments via Multi-modal Registration to Prior LiDAR Map ," in *Proc. IEEE/RSJ Intl. Conf. on Intell. Robots and Sys.*, Macau, Nov. 2019.
[33] Y. S. Park, Y.-S. Shin, and A. Kim, "PhaRaO: Direct Radar Odometry using Phase Correlation," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2020, Accepted. To appear.
[34] M. Adams, M. D. Adams, and E. Jose, *Robotic navigation and mapping with radar*, 2012.
[35] A. Licha, "Songdo and Sejong: Master-planned Cities in South Korea," Master's thesis, Urban Affairs Department, Sciences Po Paris, 2014.
[36] Z. Chen, L. Liu, I. Sa, Z. Ge, and M. Chli, "Learning context flexible attention model for long-term visual place recognition," *IEEE Robot. and Automat. Lett.*, vol. 3, no. 4, pp. 4015–4022, 2018.
[37] D. Vivet, P. Checchin, and R. Chapuis, "Radar-only localization and mapping for ground vehicle at high speed and for riverside boat," in *Proc. IEEE Intl. Conf. on Robot. and Automat.* IEEE, 2012, pp. 2618–2624.
[38] S. Lowry, N. S"underhauf, P. Newman, J. J. Leonard, D. Cox, P. Corke, and M. J. Milford, "Visual place recognition: A survey," *IEEE Trans. Robot.*, vol. 32, no. 1, pp. 1–19, 2015.