

RaPlace: Place Recognition for Imaging Radar using Radon Transform and Mutable Threshold

Hyesu Jang¹, Minwoo Jung¹, and Ayoung Kim^{1*}

Abstract—Due to the robustness in sensing, radar has been highlighted, overcoming harsh weather conditions such as fog and heavy snow. In this paper, we present a novel radar-only place recognition that measures the similarity score by utilizing Radon-transformed sinogram images and cross-correlation in frequency domain. Doing so achieves rigid transform invariance during place recognition, while ignoring the effects of radar multipath and ring noises. In addition, we compute the radar similarity distance using mutable threshold to mitigate variability of the similarity score, and reduce the time complexity of processing a copious radar data with hierarchical retrieval. We demonstrate the matching performance for both intra-session loop-closure detection and global place recognition using a publicly available imaging radar datasets. We verify reliable performance compared to existing stable radar place recognition method. Furthermore, codes for the proposed imaging radar place recognition is released for community <https://github.com/hyesu-jang/RaPlace>.

I. INTRODUCTION

Radar is an essential range sensor for mobile robotics, playing an important role in safety and environment detection. Although camera and Light Detection and Ranging (LiDAR)-based scene understanding are widely adopted in autonomous vehicles, these sensors may be limited in harsh weather, such as thick fog, sandstorms, or heavy snow. On the contrary, radar inherently surmounts the abovementioned challenges with a more extended wavelength signal. Since radar has a long-range capability and penetrates small particles, sensor data is impervious to occlusion. Despite these advantages, radar place recognition is at an incipient stage compared to camera and LiDAR-based place recognition due to the major bottleneck in radar-based navigation originating from peculiar radar characteristics.

Radar image pixel values are represented with Radar Cross Section (RCS), radar wave reflection criterion. RCS differs from every medium that radar wave reflects, which can be used as feature points. Unfortunately, intensive noise level and false vivid objects also possess high RCS values, requiring the ambiguous points removal. Hence, the main challenge of imaging radar has been differentiating noise and feature points, obtaining valid features for robot navigation.

The key to imaging radar place recognition is utilizing the image information and RCS. Therefore, extracting meaningful information from the polar radar images and simultaneously searching for the candidates with reliable

[†]This work was jointly supported by the NRF (RS-2023-20241758) and the MOTIE (20024355), Korea.

¹H. Jang, M. Jung, and A. Kim are with the Department of Mechanical Engineering, SNU, Seoul, S. Korea [dortz, moonshot, ayoungk]@snu.ac.kr

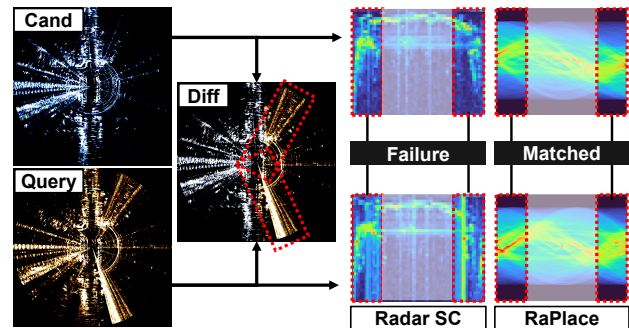


Fig. 1: A failure case in radar place recognition using the *Radar Scan Context* approach is observed due to the presence of vivid radar multipath noise and ring noises that appear as strong RCS signals. This hinders the ability of the *Radar Scan Context* approach to accurately distinguish between the noise and the feature points. However, *RaPlace* successfully overcomes the challenges posed by these noises, enabling effective place recognition performance.

accuracy is compulsory. Some tackled this issue by adopting LiDAR-based methods. For example, Hong et al. [1] and Kim et al. [2] replaced LiDAR height information with radar RCS to implement radar place recognition methods. The other approaches were learning-based methods [3, 4], demonstrating reliable results by improving the well-known visual place recognition, NetVLAD [5].

While these methods are groundbreaking in imaging radar place recognition, they do not contain a metric algorithm that takes into account radar noise and RCS properties. Despite being ranging sensors, LiDAR-based methods are insufficient in comprehensively encompassing radar data and its accompanying noise. The learning-based methods also exhibit these limitations, and treat radar images as camera images or LiDAR point clouds. Additionally, learning-based imaging radar place recognition approaches are not publicly accessible, thereby algorithm reproducing for verification is challenging. While prior research has established the reliability, it is obvious that a readily available metric algorithm is required.

For the recent work, Lu et al. [6] proposed the Radon Transform (RT)-based LiDAR place recognition that is both rotation and translation invariant. This work motivated us in that the RT was originally designed to analyze tomography images rather than scanned pointcloud images. In contrast to LiDAR, radar data shares certain similarities with tomography images, being generated by rays that penetrate the surrounding environment. Furthermore, tomography images obtained from X-ray imaging may exhibit undesired effects such as radiation saturation and detector noise, which are

akin to the noise observed in radar images. Hence, we adopt the RT to analyze the radar image. Additionally, since RT-applied data forms the basis for noise post-processing, applying RT to radar images provides a similar advantage as shown in Fig. 1.

In this paper, we address the imaging radar-based place recognition method using RT, with algorithmically congenial to radar image processing. Instead of eliminating the ambiguous points, we propose to stifle the influence of the futile points. The proposed contributions of our paper are listed below.

- **A new paradigm for the radar place recognition** The proposed method is a rotational and translational invariant high-precision place recognition method for radar. The addressed approach can fill up the lacuna in reverse direction driving. To adapt the LiDAR-based method to radar, we remodeled existing similarity distance scoring algorithm and adopted hierarchical candidate extraction with data abridgment.
- **Non-learning radar place recognition criteria for wide tailorability.** The existing criteria for radar place recognition are scarce, a confident localization tool for radar simultaneous localization and mapping (SLAM) is required. The proposed approach affords high-accuracy place recognition results with low computational complexities without needing a GPU.
- **Open-source radar place recognition code** Currently few radar place recognition methods are publicly available. We release the code for radar place recognition.

II. RELATED WORK

A. Place Recognition using Cameras and LiDARs

1) *Visual Place Recognition using RGB Cameras:* Visual place recognition studies frequently utilize feature-based descriptors [7, 8, 9]. Some recent studies have leveraged specialized feature-based searching techniques. For example, Arandjelovic et al. [5] introduced a learning model architecture for visual place retrieval that combined a CNN architecture with a generalized vector of locally aggregated descriptors. Qazanfari et al. [10] employed the HSV color model and image histogram as the basis for image retrieval, utilizing color discrepancies and image histogram as the retrieval features. Panek et al. [11] generated a 3D mesh by leveraging 2D image depth and employed it for place recognition. While these techniques have yielded promising results, they remain constrained by issues stemming from the camera's inherent limitations, such as variations in illumination and challenging weather conditions. Notably, rendering techniques are unsuitable for 2D imaging radar.

2) *LiDAR Place Recognition:* Although numerous place recognition studies have employed visual sensors, utilizing LiDAR is also prevalent. Kim and Kim [12] proposed the *Scan Context* algorithm, a versatile place recognition approach for LiDAR data. The cosine similarity between two context representations of the data is used to compute the similarity distance between two images. An improved

version of this algorithm, *Scan Context++* [13], was introduced to achieve translational invariance by incorporating Cartesian context and to enhance performance through the use of hierarchical clustering.

Uy and Lee [14] improved the network architecture of PointNet [15] and NetVLAD [5] to conduct place recognition in the pointcloud data. Authors utilized descriptors with lazy triplet and quadruplet loss that were found to effectively distinguish pointcloud data, resulting in a high retrieval rate.

Ding et al. [16] and Lu et al. [6] employed RT and Discrete Fourier Transform (DFT) to compress image and implement a global searching algorithm. Especially, [6] derived the rotation and translation invariance of the proposed place retrieval method and estimated the orientation and translation differences. The extension version of RT-based LiDAR place recognition [6] was recently addressed by Xu et al. [17], which features the application of multiple sensors.

We conduct an adaptation process of RT-based algorithm for radar images. As demonstrated above, RT yields useful elements for reducing noise in tomography images with saturation. We address the issue of radar-induced noise by employing RT, where radar images are utilized as tomography images. To facilitate this adaptation, we conduct preprocessing on the radar images to address their inherent limitations and to produce a highly accurate query-pursuant scoring descriptor.

B. Place Recognition with Radars

Radar sensors can be divided into two types: (i) single-chip automotive radar, which provides point clouds with velocity value, and (ii) scanning imaging radar, which generates high-resolution image data.

Automotive radar, also known as single-chip radar, is a commonly used sensor for vehicles. As a result of the limited density of radar point cloud data, the research in the area of localization [18, 19, 20, 21] has advanced further than that in place recognition. Recently, Cait et al. [22] demonstrated the state-of-the-art (SOTA) outcomes in automotive radar place recognition. The authors conducted radial velocity-based dynamic points removal to produce stable candidates, encoded spatial and temporal information for retrieval, and enhanced accuracy through RCS histogram reranking.

In contrast to the sparse implementation of automotive radar, scanning radar is more frequently employed for the kidnapped vehicle problem. Kim et al. [2] presented a ranging sensor dataset with radar place recognition evaluation. Instead of using LiDAR elevation values, they utilized radar RCS values to create a radar scan context. Tang et al. [23] attempted to localize satellite images using radar image data. They estimated the rotation value between the radar and satellite images, and constructed a neural network to estimate the translation value and calibrate the images. Săftescu et al. [24] proposed a learning-based approach for topological localization to achieve metric pose estimation. The authors leveraged the fact that the position information of the polar radar image is independent of the heading. Wang et al. [25] addressed a 6-DOF global vehicle pose estimation network,

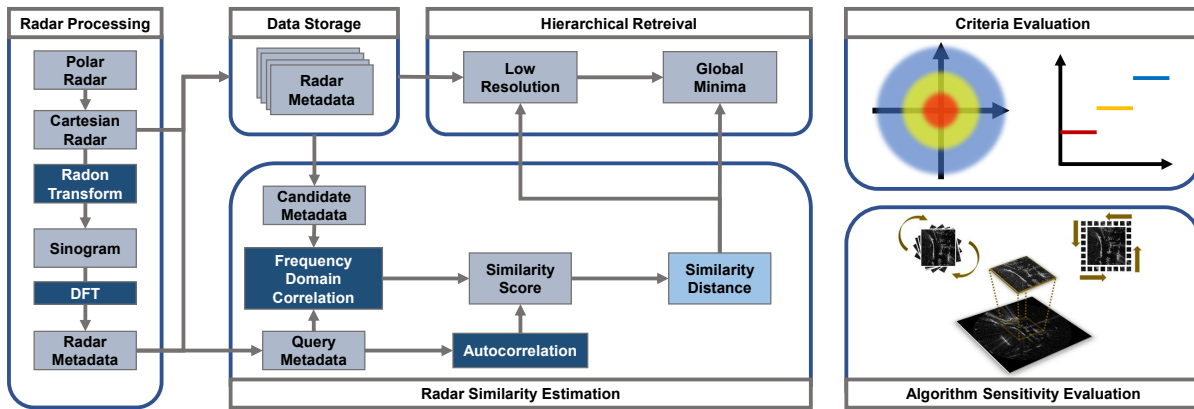


Fig. 2: The framework for RaPlace and overview of ablation studies. Initially, processed radar data is transformed into a query data and is stored in a metadata format. Candidates from metadata are evaluated with frequency domain correlation estimation. The resulting correlation scores represent the degree of similarity between the candidates and the query data, and the auto-correlation score provides a mutable threshold for determining the similarity distance. The complete framework is executed on low-resolution images and subsequently extended to high-resolution images to explore the global minima using the abovementioned steps. Ablation studies are carried out to assess the performance of the system, with a focus on criteria and sensitivity evaluation.

which employed a self-attention module for localization based on radar images. Yin et al. [26] demonstrated multi-modal sensor localization using radar on LiDAR. They used neural network-based data embedding and a differentiable Kalman filter to achieve multi-modal localization.

As summarized above, research on radar place recognition typically relies on neural networks that are challenging to reproduce. Furthermore, the solutions to the kidnapped problem using simple metric algorithms are insufficient. We propose an imaging radar-based non-learning approach to place recognition that is both resilient to noise and capable of overcoming the inherent limitations of radar sensors.

III. METHOD

Being a non-learning-based method, we seek a descriptor that can differentiate between candidates analogous to the query data. Our proposed model aims to identify places that include the vicinity of the target location, thus extending the scope of detection and smoothing the disclosure of identical places. Fig. 2 depicts the entire pipeline for our place recognition algorithm. We generate backward-warping Cartesian images from the polar image data and obtain sinogram images using the RT on the Cartesian radar data. These acquired sinograms are then converted into FFT images, and utilized as candidates. To determine the similarity score, we adopt the cross-correlation in frequency domain and mutable threshold for all previous keyframes when query images are published. The candidate location is determined as the keyframe with the lowest distance, assessed from the similarity scores. The following subsections provide a detailed account of the procedures.

A. Radar Data Preparation

FMCW radar generates 4Hz polar radar images with RCS values. Polar images are irrelevant to rotational changes however, tracing translational variations is challenging. To account for translational differences, Cartesian images are required. Unlike LiDAR bird-eye-view images, radar can

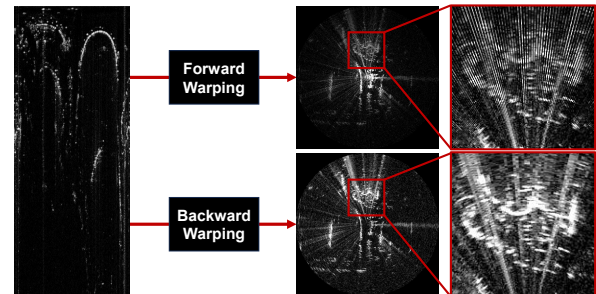


Fig. 3: The outcome of Cartesian imagery varies according to the warp methodologies. The process of forward warping leads to the generation of empty spaces amongst pixels. Contrastingly, backward warping exhibits all-occupied pixel spaces.

penetrate obstacles, which allows RCS information to be fully captured up to the range limit. To fully utilize this characteristics, we convert the polar images using a backward warping technique. As illustrated in the Fig. 3, the process of forward warping directly registers the polar pixel value in Cartesian space. As a consequence of this approach, images manifest as a pattern of discrete rays, hindering the complete utilization of the converted image due to the presence of vacant pixels. On the contrary, backward warping initiates with an empty Cartesian space and progressively populates all the xy pixels corresponding to the closest polar pixel. This conversion offers substantial assistance in pixel-level analysis of translational disparities.

B. Radon Transform for Radar

Summarizing the image data is a crucial step in generating a descriptor. The Radon transform is a commonly used image processing and restoration method in medical image treatment. The representative utilization is Computed Tomography (CT), where rotating light emissions generate the projection vectors for every angle. The column-wise enumeration of these vectors forms a sinogram that encapsulates the entire image data. We decide to employ this image data compression technique for our robotics problem. Fig. 4

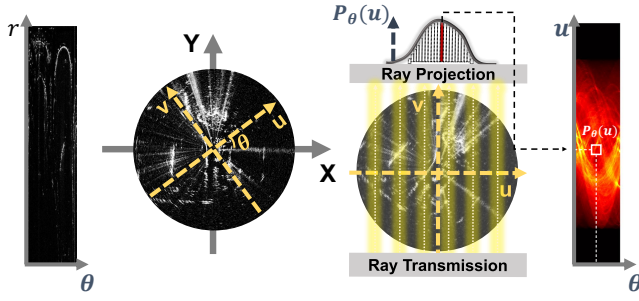


Fig. 4: The fundamental concept behind the RT. For a given input of a raw radar image, there exists a rotating coordinate system (u, v) that propagates a ray along the v axis. These rays penetrate the radar contents and generate projection vectors that ultimately become the column vectors of the sinogram.

is a schematized image of RT and sinogram. The relation between (x, y) and (u, v) for the varying coordinate can be expressed using the rotational transform.

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \quad (1)$$

We use line integration to measure the image density $P_\theta(u)$. Densities are calculated with the rays penetrating the radar detections.

$$\begin{aligned} P_\theta(u) &= \int f(x, y) dv \\ &= \int f(u \cos \theta - v \sin \theta, u \sin \theta + v \cos \theta) dv \end{aligned} \quad (2)$$

The two-dimensional slice that represents the data penetration is denoted by $f(x, y)$. Expanding the equation from a line to the entire area allows us to generate a sinogram image $\mathcal{S}(\theta, u)$, the result of the RT. The total densities along the v -axis direction compose the compressed data.

$$\mathcal{S}(\theta, u) = \begin{bmatrix} P_0(l_{min}) & \cdots & P_\pi(l_{min}) \\ \vdots & \ddots & \vdots \\ P_0(l_{max}) & \cdots & P_\pi(l_{max}) \end{bmatrix} \quad (3)$$

The non-zero boundary of $P_\theta(u)$ is denoted by l . In our case, we restrict the interval $[l_{min}, l_{max}]$ as the diagonal length of the input Cartesian radar image. As demonstrated by Ding et al. [16], utilizing sinograms with DFT can provide both rotational and translational invariance. To achieve this, the RT is used to generate sinograms for each radar image, and applying DFT to these sinogram images (4) ensures translational invariance.

$$I_i(\theta, f) = \mathcal{F}(\mathcal{S}_i(\theta, l)) = \sum_l \mathcal{S}(\theta, l) \cdot e^{-2\pi i \left(\frac{fl}{M}\right)} \quad (4)$$

C. Radar Similarity Appraisal with Mutable Threshold

The primary derivation of the distance score in our method is combining two techniques: cross-correlation and mutable threshold. Cross-correlation is commonly used to measure the similarity between two sets of data. However, calculating the cross-correlation directly in image data can be time-consuming. Therefore, we utilize the frequency domain

cross-correlation algorithm to generate the basis for the place similarity score. The cross-correlation algorithm in frequency domain is described in (5).

$$C_{q,i} = \sum_\theta \mathcal{F}^{-1}(\mathcal{F}(I_q)\mathcal{F}^*(I_i)) \quad (5)$$

We apply the DFT to candidate and query matrices I_i and I_q , and convert them into the frequency domain. The candidate and query matrices are constructed as lists of the discrete Fourier transformed vectors for each theta in sinogram \mathcal{S} (4). Next, the cross-power spectrum is computed by multiplying the DFT of the query matrix with the complex conjugate of the DFT of the candidate matrix. Finally, the inverse Fourier transform is applied to the cross-power spectrum data and summed along one axis, resulting in a one-dimensional array of cross-correlation $C_{q,i}$.

To account for differences between the query image and candidates, we calculate the auto-correlation value C_{auto} as a mutable threshold. Because the radar image captures the full surrounding environment, the score obtained from a single cross-correlation array may not always accurately reflect the information in the query radar image. To address this issue, we improve the standard of the similarity score by incorporating two cross-correlation results. Both the image similarity score for the auto-correlation and the query-candidate are determined by selecting the maximum value from the cross-correlation array. To identify the answer from the candidate images, we consider the candidate whose obtained score is closest to the auto-correlation value. Given that mutable query thresholds, the final place similarity distance d is calculated according to (6)

$$d_i = |\max(C_{auto}) - \max(C_{q,i})| \quad (6)$$

$$X_q = \underset{i}{\operatorname{argmin}}(d_i) \quad (7)$$

The output of the place recognition model is represented by X_q in (7), which is a candidate minimizing the similarity distance d .

D. Place Information Retrieval

Place decisions are conducted based on the place similarity distance, but as the interminable driving, the number of place candidates increases, leading to performance degradation. To enable efficient place retrieval across a wide range of fields without sacrificing performance, an efficient searching algorithm is required. To this end, we adopt a hierarchical method for agile minima access. Since the scores for neighboring scenes are analogous, we extract representative candidates from the cognate image collections. Under the assumption that resolution variances are irrelevant to the distance score tendency, we first obtain keyframe scores from low-resolution data, and then refine the candidate set using the highly rated group and neighboring data to achieve accurate place retrieval.

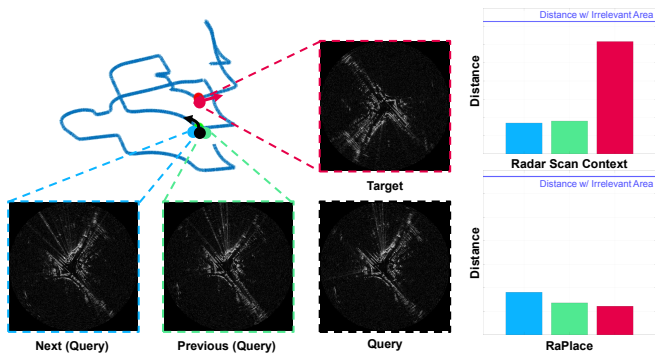


Fig. 5: Time elevation graph and place matching example in the KAIST sequence. While both Radar Scan Context and RaPlace typically indicate high similarity between neighboring frames, there are certain scenarios where they exhibit different behavior. For instance, when the vehicle undergoes a full rotation and there is a distance of approximately 8 meters between the two positions, Radar Scan Context detects a substantial difference from the neighboring frames, whereas RaPlace remains stable.

IV. EXPERIMENT

A. Radar Datasets

To ensure the reliability of verification, we evaluate our algorithm on the two datasets: *Oxford Radar Robotcar* [27] and *MulRan* [2]. Table. I illustrates a comparison of key characteristics for the dataset. The *Oxford* dataset is a short-term radar dataset that spans a week and covers the same area, making it suitable for straightforward intra-loop closure verification and consecutive global place recognition. In contrast, *MulRan* is a long-term dataset that allows for temporal-global place recognition. *MulRan* is pertinent for substantiating both single-session and multi-session loop closure, as each session drives different routes. We test our single-session loop closure algorithm on two *Oxford* datasets and four locations in *MulRan* and conduct multi-session verification in each representative location.

Dataset	Distance	Max. Interval	Sequences	Sessions	
MulRan	DCC	4.9km	33 days	3	Intra
	KAIST	6.1km	75 days	3	Intra
	Riverside	6.8km	22 days	3	Intra/Multi
	Sejong	23.4km	62 days	3	Multi
Oxford	9km	8 days	32	Intra/Multi	

TABLE I: Radar Dataset Attributes

B. Evaluation Criteria

1) *True Positives Matching Graph*: Demonstrating the indisputable accuracy of our place recognition approach, we visualize the matching results for each position. Results indicate the true-positive predictions for each location and demonstrate the robustness and stability of the algorithm. Furthermore, the graphs reveal the vulnerability of the place recognition descriptor. By creating a true-positive point trajectory for the four intra-sessions and TP detection occurrence graphs for the three multi-sessions, we compare all matching results with the state-of-the-art method for structural radar place recognition *MulRan* (Radar Scan Context) [2]. We evaluate the performance of our method and assess the effectiveness in place recognition.

Dataset	Method	Single Session		
		AUC	F1 Score	
MulRan	DCC	R-SC	0.801	0.679
		RP	0.843	0.735
	KAIST	R-SC	0.819	0.786
		RP	0.881	0.781
Riverside	R-SC	0.944	0.862	
	RP	0.959	0.897	
Sejong	R-SC	0.904	0.830	
	RP	0.924	0.858	
Oxford	19-01-16	R-SC	0.545	0.534
		RP	0.595	0.543
	19-01-18	R-SC	0.826	0.773
	RP	0.827	0.734	

TABLE II: Intra-session Place Recognition Results

2) *Assessment with Scores*: In our evaluation, we configure the place identification standards to operate within a 20-meter range, accounting for the sparsity of the imaging radar data. An alleviated matching standard is recommended for radar place recognition in contrast to LiDAR, which is better suited for detecting rough environmental changes. To evaluate the performance of our prediction model, we rely on the Area Under the Curve (AUC) value from the Precision-Recall (PR)-curve as a reliable criterion for evaluation. Additionally, we account for the F1 score-recall curve and the maximum F1 score for a comprehensive assessment of our method. When evaluating the prediction model, we calculate Recall@1 only when the exact answer exists among the candidates. However, since the matching answer may not always exist in SLAM, Recall@1 is not a practical measure for scarce loop closure instances.

C. Intra-session Link Proposal

The DCC, KAIST, and Riverside datasets from *MulRan* and each sequence from the *Oxford* dataset are rich in loop closures, making them suitable for verifying the reverse lane recognition and data attributes in our method. The intra-session place recognition evaluation results or Radar Scan Context (R-SC) and RaPlace (RP) are presented in Fig. 1, Fig. 6 and Table. II, where the true-positive match consequences are identified, allowing us to detect where the localization performances are refined.

To visually represent the results of our evaluation, we create a trajectory graph that features true-positive points. This graph includes green points that denote detections that our algorithm identifies but are missed by Radar Scan Context and red points that represent detections that our algorithm fails to detect but are identified by Radar Scan Context. This approach allows us to easily compare the performance of our method to the state-of-the-art method for structural radar place recognition. As shown in Fig. 6, the proposed method finds a significant number of true-positive matches that the Radar Scan Context overlooked. Notably, our algorithm is particularly robust in places where both rotational and translational intervals exist.

The quantitative results in Fig. 6 and Table. II include specific numerical values that indicate the precision, F1 score, and recall of the prediction models. Compared to existing methods, the proposed method shows a relaxed

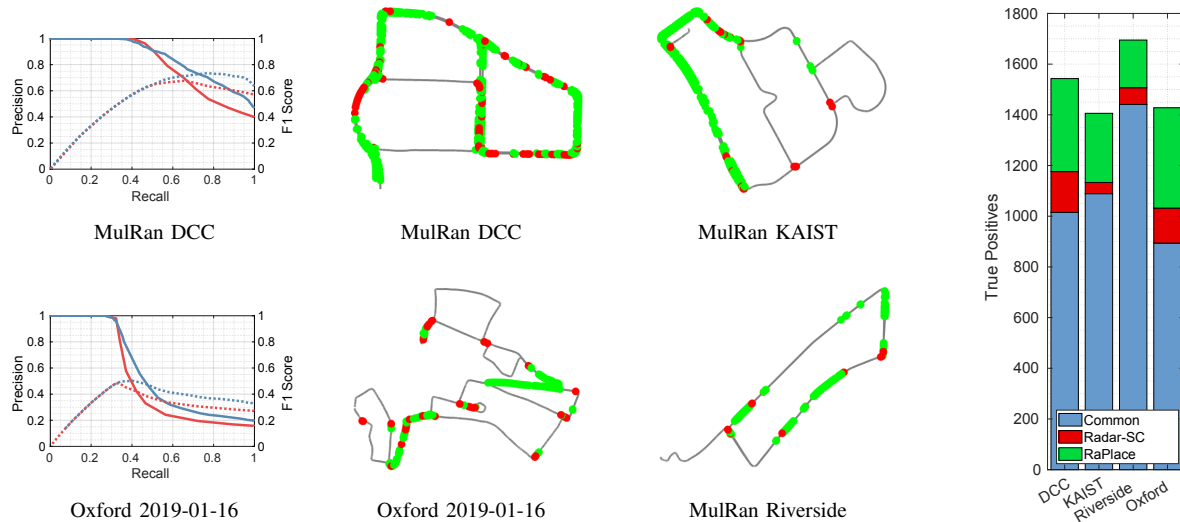


Fig. 6: The figure on the left illustrates the PR and F1-recall curves for the intra-session evaluation. The solid lines represent the precision values, while the dotted lines represent the F1 score for each recall. The results obtained by the proposed method (blue line) are substantially superior to those of Radar Scan Context (red line), with the proposed method detecting more latent links that improve the precision values at high recalls. However, for the Oxford dataset, the low precision scores for both methods are attributed to the scarcity of revisited locations. The middle part of the figure depicts the trajectory, with points where only the proposed method (green) or Radar Scan Context (red) detect the same place. The statistical graph on the right indicates that our proposed method detects more true positive results than the Radar Scan Context.

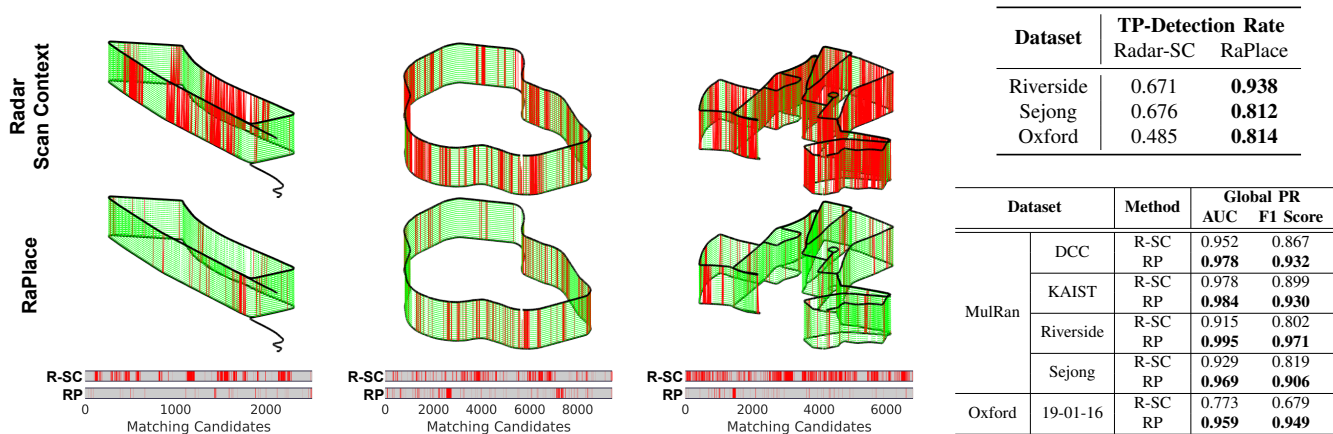


Fig. 7: The global place recognition results for three large-scale sequences. We represent all matching results obtained by the Radar Scan Context and RaPlace as green lines and the detection failures of each method as red lines. The Radar Scan Context fail to detect certain places due to the rigid transformation information being different from the original path. However, our proposed method is robust in detecting the surrounding places, resulting in a high rate of true positive detection.

decrease in precision. RaPlace exhibits invariance for image information in matching, which allows it to identify obscure place pairs and increase overall precision. The proposed method outperforms existing structural methods in terms of the Area Under the Curve (AUC). Since the precision is supplemented for the ample recall in our algorithm, the maximum value of the F1 score is extended and located at a higher recall value.

D. Multi-session Place Recognition

The results of the multi-session evaluation are presented in Fig. 7, and they are consistent with the intra-session results.

In the multi-sequence trajectory graphs, green lines represent matched links between two sequences using all place recognition methods, while red lines denote instances where each method failed to detect a match. Our proposed method outperforms Radar Scan Context regarding detection rates, with fewer missed detections overall. As shown in the table in Fig. 7, the density of entire true-positive detections is significantly improved with the proposed method, demonstrating accuracy in global place recognition. The quantitative global place recognition results for all sequences are summarized in the tables in Fig. 7. We achieve comprehensive precision values in adequate recalls by reducing the false-positive ratio.

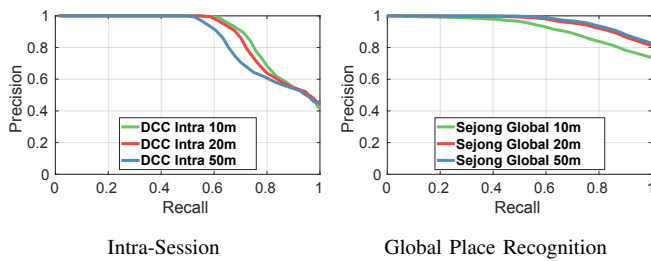


Fig. 8: The figure illustrates the results for the variance of the loop detection criteria, which are set at 10m, 20m, and 50m. For the intra-session evaluation, the precision values decrease as the criteria was relaxed. However, for the global place recognition results, expanding the criteria improves the detection accuracy.

The performance of global place recognition is significantly improved by increasing the absolute true-positive quantity.

E. Effect of Loop Detection Criteria

When conducting multi-session place recognition, we observe that the rearranged driving route and the temporal difference lead to confusion in searching for the same location. Therefore, ground truth boundary configuration is essential to evaluate the performance of the global place recognition algorithm, which must take into account the different traffic lanes containing opposite positions. We obtain the prediction results for the 10, 20, and 50 meters boundary and depict PR-curves for the DCC and Sejong from *MulRan*. Fig. 8 is the PR curves for the multiple ground truth boundaries. Our evaluation indicates that in the intra-session case, a larger ground truth boundary leads to less accurate results. This is because intra-session sequences share nearly identical environments, readily finding accurate correspondences. Additionally, intra-session sequences typically follow short paths, which limits the number of candidates that need to be considered. However, we observe that alleviating the ground truth boundary improves the accuracy of the global place recognition. In global place recognition, we need to search a large amount of data to find suitable candidates, and the environments in the candidate data may differ from the query data. In this process, an alleviated ground truth boundary helps to identify nearby environments and find the most similar representations from the candidate data. These findings support our initial assumptions regarding the significance of radar frequency and lane width for place recognition criteria. Also, the radar range is up to 200m, mitigated confines of ground truth represent desirable candidates for loop closure, resulting in reliable true-positive results.

F. Sensitivity of Place Recognition

To assess the robustness of our proposed model, we evaluate our place recognition algorithm to verify sensitivity. Although we claim to achieve both rotational and translational invariance, it is imperative to provide empirical evidence to support our assertion. To compare the performance of Radar Scan Context and RaPlace, we conduct a graph tendency analysis subject to different thresholding standards. We select a query sample from the radar dataset and generate candidate samples by rotating and rearranging both vertical

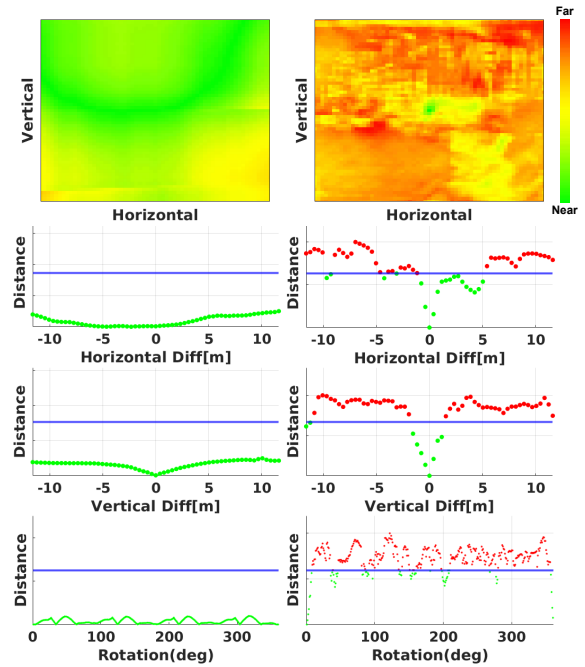


Fig. 9: The sensitivity evaluation results for virtual candidates, which are generated by rotating and cropping the query image. The distance scores are measured to verify the algorithm's coverage. Across all variations, RaPlace exhibits a more consistent and stable distance score than Radar Scan Context. While all distance values have to be below the threshold (Blue line), Radar Scan Context only provides a reliable score for a limited area.

and horizontal directions. To avoid squandering neighbor radar information, we crop the query image from the original full Cartesian image. Each rotated and translated image is occupied with real surrounding information and devoid of empty black spaces.

Fig. 9 displays the results of the similarity distance analysis. Radar Scan Context and RaPlace are used to determine the minimum distance to assure place analogy, and the algorithms are relative detection methods. To ensure a fair comparison of the performance of each algorithm, we normalize all graphs by reflecting the threshold values of each algorithm. The threshold values are calculated using the two radar data from completely different places. We observe that Radar Scan Context has a reliable score in the periphery of minima for both rotational and translational variance. However, it requires precise revisits to achieve confident place recognition, and analogous places could not attain a definitely distinguishable score. Conversely, the proposed method exhibits robustness for both rotational and translational variance, with all our scores remaining below the threshold value. Our sensitivity study provides compelling evidence as to why the RT-based algorithm identifies more places and offers inclusive localization candidates.

G. Computational Time

We evaluate the execution time of the proposed algorithm, which consists of three primary steps: polar to Cartesian conversion, Radon transform, and image retrieval. The statistical computations were performed utilizing a 2.5GHz 16-Core Intel i7-11700 CPU. Retrieving 2000 frames took

approximately 30ms, achieving real-time performance. The data loading time for all DCC data is approximately 38.4ms, and given that radar images are produced at a rate of 4Hz, our algorithm can process the offline dataset in real-time. However, the average computing time for the polar to Cartesian conversion and the Radon transform are about 85.25ms and 86.58ms, respectively, making it challenging to ensure real-time performance in long-term online environments.

V. CONCLUSION

We propose a novel radar place recognition algorithm that is rigid transformation-invariant and noise-irrelevant. Our methodology involves the computation of similarity scores and the estimation of similarity distances through the use of a mutable threshold. Results derived from experimentation and ablation studies indicate that our method produces promising outcomes. Despite our methodology utilizing RT and DFT to condense image information, the computational time requires to process high-resolution images still presents a challenge. As such, the development of a time-image resolution complementary algorithm is imperative to enable efficient storage and retrieval of radar location information. As part of our future work, we plan to modify our algorithm to be compatible with radar-based modules. Given that our method produces a greater number of potential candidates for loop closure than preexisting radar place recognition methodologies, we expect that our approach could be leveraged to achieve highly accurate radar-based SLAM.

REFERENCES

- [1] Z. Hong, Y. Petillot, and S. Wang, "Radarslam: Radar based large-scale slam in all weathers," in *Proc. IEEE/RSJ Intl. Conf. on Intell. Robots and Sys.*, 2020, pp. 5164–5170.
- [2] G. Kim, Y. S. Park, Y. Cho, J. Jeong, and A. Kim, "Mulran: Multimodal range dataset for urban place recognition," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2020, pp. 6246–6253.
- [3] D. De Martini, M. Gadd, and P. Newman, "Kradar++: Coarse-to-fine fmcw scanning radar localisation," *IEEE Sensors J.*, vol. 20, no. 21, p. 6002, 2020.
- [4] M. Gadd, D. De Martini, and P. Newman, "Look around you: Sequence-based radar place recognition with learned rotational invariance," in *IEEE/ION Position, Location and Navigation Symposium*, 2020, pp. 270–276.
- [5] R. Arandjelovic, 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, pp. 5297–5307.
- [6] S. Lu, X. Xu, H. Yin, Z. Chen, R. Xiong, and Y. Wang, "One ring to rule them all: Radon sinogram for place recognition, orientation and translation estimation," in *Proc. IEEE/RSJ Intl. Conf. on Intell. Robots and Sys.*, 2022, pp. 2778–2785.
- [7] P. C. Ng and S. Henikoff, "Sift: Predicting amino acid changes that affect protein function," *Nucleic acids research*, vol. 31, no. 13, pp. 3812–3814, 2003.
- [8] E. Rosten, R. Porter, and T. Drummond, "Faster and better: A machine learning approach to corner detection," *IEEE Trans. Pattern Analysis and Machine Intell.*, vol. 32, no. 1, pp. 105–119, 2008.
- [9] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "Orb: An efficient alternative to sift or surf," in *Proc. IEEE Intl. Conf. on Comput. Vision*, 2011, pp. 2564–2571.
- [10] H. Qazanfari, H. Hassanpour, and K. Qazanfari, "Content-based image retrieval using hsv color space features," *Intl. J. of Comput. and Information Eng.*, vol. 13, no. 10, pp. 533–541, 2019.
- [11] V. Panek, Z. Kukulova, and T. Sattler, "Meshloc: Mesh-based visual localization," in *Proc. European Conf. on Comput. Vision*, 2022, pp. 589–609.
- [12] 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.*, 2018, pp. 4802–4809.
- [13] G. Kim, S. Choi, and A. Kim, "Scan context++: Structural place recognition robust to rotation and lateral variations in urban environments," *IEEE Trans. Robot.*, vol. 38, no. 3, pp. 1856–1874, 2021.
- [14] M. A. Uy and G. H. Lee, "Pointnetvlad: Deep point cloud based retrieval for large-scale place recognition," in *Proc. IEEE Conf. on Comput. Vision and Pattern Recog.*, 2018, pp. 4470–4479.
- [15] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in *Proc. IEEE Conf. on Comput. Vision and Pattern Recog.*, 2017, pp. 652–660.
- [16] X. Ding, X. Xu, S. Lu, Y. Jiao, M. Tan, R. Xiong, H. Deng, M. Li, and Y. Wang, "Translation invariant global estimation of heading angle using sinogram of lidar point cloud," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2022, pp. 2207–2214.
- [17] X. Xu, S. Lu, J. Wu, H. Lu, Q. Zhu, Y. Liao, R. Xiong, and Y. Wang, "Ring++: Roto-translation invariant gram for global localization on a sparse scan map," *arXiv preprint arXiv:2210.05984*, 2022.
- [18] E. Ward and J. Folkesson, "Vehicle localization with low cost radar sensors," in *Proc. IEEE Intell. Vehicle Symposium*, 2016, pp. 864–870.
- [19] F. Schuster, M. Wörner, C. G. Keller, M. Hauéis, and C. Curio, "Robust localization based on radar signal clustering," in *Proc. IEEE Intell. Vehicle Symposium*, 2016, pp. 839–844.
- [20] 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.*, 2019, pp. 1307–1314.
- [21] C. X. Lu, M. R. U. Saputra, P. Zhao, Y. Almalioglu, P. P. De Gusmao, C. Chen, K. Sun, N. Trigoni, and A. Markham, "milliego: single-chip mmwave radar aided egomotion estimation via deep sensor fusion," in *Conf. on Embedded Networked Sensor Systems*, 2020, pp. 109–122.
- [22] K. Cait, B. Wang, and C. X. Lu, "Autoplace: Robust place recognition with single-chip automotive radar," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2022, pp. 2222–2228.
- [23] T. Y. Tang, D. De Martini, D. Barnes, and P. Newman, "Rsl-net: Localising in satellite images from a radar on the ground," *IEEE Robot. and Automat. Lett.*, vol. 5, no. 2, pp. 1087–1094, 2020.
- [24] Ş. Săftescu, M. Gadd, D. De Martini, D. Barnes, and P. Newman, "Kidnapped radar: Topological radar localisation using rotationally-invariant metric learning," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2020, pp. 4358–4364.
- [25] W. Wang, P. P. de Gusmao, B. Yang, A. Markham, and N. Trigoni, "Radarloc: Learning to relocalize in fmcw radar," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2021, pp. 5809–5815.
- [26] H. Yin, R. Chen, Y. Wang, and R. Xiong, "Rall: end-to-end radar localization on lidar map using differentiable measurement model," *IEEE Trans. Intell. Transport. Sys.*, vol. 23, no. 7, pp. 6737–6750, 2021.
- [27] D. Barnes, M. Gadd, P. Murcutt, P. Newman, and I. Posner, "The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset," in *Proc. IEEE Intl. Conf. on Robot. and Automat.*, 2020, pp. 6433–6438.