

# RaNDT SLAM: Radar SLAM Based on Intensity-Augmented Normal Distributions Transform

Maximilian Hilger<sup>1,2</sup>, Nils Mandischer<sup>1,3</sup>, and Burkhard Corves<sup>1</sup>

**Abstract**—Rescue robotics sets high requirements to perception algorithms due to the unstructured and potentially vision-denied environments. Pivoting Frequency-Modulated Continuous Wave radars are an emerging sensing modality for SLAM in this kind of environment. However, the complex noise characteristics of radar SLAM makes, particularly indoor, applications computationally demanding and slow. In this work, we introduce a novel radar SLAM framework, RaNDT SLAM, that operates fast and generates accurate robot trajectories. The method is based on the Normal Distributions Transform augmented by radar intensity measures. Motion estimation is based on fusion of motion model, IMU data, and registration of the intensity-augmented Normal Distributions Transform. We evaluate RaNDT SLAM in a new benchmark dataset and the Oxford Radar RobotCar dataset. The new dataset contains indoor and outdoor environments besides multiple sensing modalities (LiDAR, radar, and IMU).

## I. INTRODUCTION

Radar-based Simultaneous Localization and Mapping (SLAM) is an emerging alternative to optical sensors in vision-restricted environments. Compared to LiDAR, the long wavelength in the millimeter range of Frequency-Modulated Continuous Wave (FMCW) radars can easily penetrate fog, dust, and similar volumetric disturbances [1]. Applications and existing approaches may be loosely grouped into the domains of autonomous driving [2]–[6] and rescue robotics [1], [7]–[9]. They offer different environmental characteristics. The confined indoor environments typically encountered by rescue robots lead to high noise in the radar signals. Therefore, significant computational effort has to be spent dealing with this noise, which leads to slow update rates. Outdoor radar SLAM approaches (e.g., autonomous driving) can rely on simpler techniques but are often untested in indoor environments. In this work, we focus on rescue robotics. Note, that rescue robots not only operate indoors, but can also be deployed in outdoor disaster scenarios, hence, we need to address two challenges: real-time performance and applicability in multiple environments. To this end, we propose a novel SLAM framework, RaNDT SLAM, based on scan registration using the Normal Distributions Transform (NDT), which takes not only the point entities but also the intensity of the radar signal into account.

<sup>1</sup>All authors are with the Institute of Mechanism Theory, Machine Dynamics and Robotics, RWTH Aachen University, Aachen, Germany.

<sup>2</sup>M. Hilger is also with the Chair of Perception for Intelligent Systems, Munich Institute of Robotics and Machine Intelligence, Technical University of Munich, Munich, Germany. maximilian.hilger@tum.de

<sup>3</sup>N. Mandischer is also with the Chair of Mechatronics, University of Augsburg, Augsburg, Germany. nils.mandischer@uni-a.de

This work was funded by the BMBF under grant ID 13N16482.

We use submaps storing intensity-augmented NDTs to accumulate filtered radar data. We estimate motion locally by fusing NDT registration, the yaw component of IMU measurements, and a motion model. Global consistency is ensured using a pose graph and a loop closure detector.

Our main contributions in RaNDT SLAM are:

- Novel modular estimation framework using radar scans, IMU measurements, and a motion model.
- Demonstration of real-time performance of our system in indoor and outdoor environments using a new benchmark dataset and the Oxford Radar RobotCar dataset<sup>4</sup>.
- Implementation<sup>5</sup> and dataset [10] publicly available, providing a new baseline for radar SLAM in rescue robotics. To the authors' knowledge, this is the first dataset for SLAM using indurad's iSDR-C radar sensor.

In the following, we first briefly review existing radar SLAM approaches using pivoting radar sensors and introduce the NDT and its applications in radar scan registration (Section II). In Section III, we introduce and explain RaNDT SLAM. We evaluate the proposed framework profoundly in indoor and outdoor scenarios, and show its real-time capability (Section IV). Finally, the results are summarized and potential additions are discussed (Section V).

## II. RELATED WORK

Pivoting FMCW radars are a commodity for SLAM in rescue robotics. We introduce SLAM algorithms using FMCW radars and the NDT SLAM, which forms this work's base.

### A. Related Work with Pivoting Radar Sensors

We categorize the existing radar SLAM systems based on their target environment into indoor and outdoor:

1) *Indoor Environments*: The radar sensor used for the indoor SLAM of Marck et al. [7] generates scans using only the strongest return per azimuth. They estimate motion based on scan-to-scan ICP and generate a map using a grid-mapping Particle Filter. In the *SmokeBot* project, a mechanically pivoting radar (MPR) was developed [11]. Within the project, Fritsche et al. [1] combine radar and LiDAR scans for SLAM and investigate two methods: a feature-based approach using a Kalman filter, and a scan registration-based method. Further, Mielle, Magnusson, and Lilienthal [8] compare *gmapping* and NDT-OM using only radar data produced by the MPR. Mandischer et al. [12] pre-process radar scans using Otsu thresholding and beam-wise clustering. The motion is estimated by probabilistic

<sup>4</sup>We will use the shortened term "Oxford dataset" in the following.

<sup>5</sup><https://github.com/IGMR-RWTH/RaNDT-SLAM>

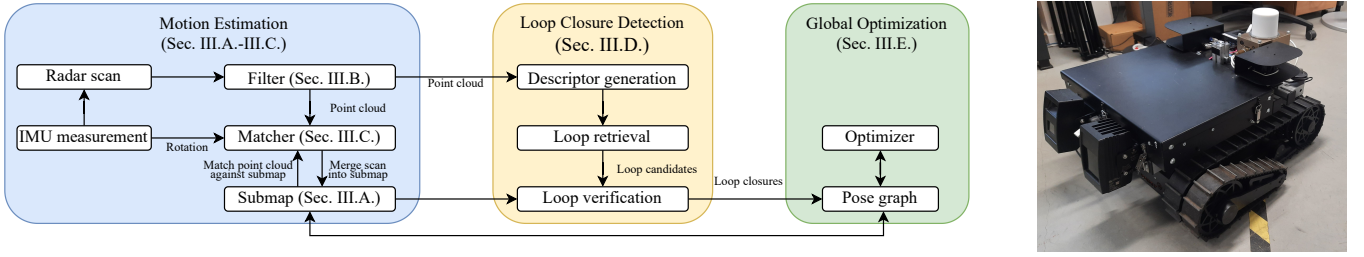


Fig. 1: Architecture of the RaNDT SLAM framework and robot used for evaluation in Section IV-A. The robot is equipped with an indurad iSDR-c pivoting radar, a phidgets IMU, and a pair of Sick TiM LiDARs.

iterative correspondences matching and an autoregression-moving-average filter for outlier rejection. Torchalla et al. [9] propose a custom sensor setup using a pivoting single-chip radar. They use Truncated Signed Distance Functions for map representation and scan registration.

2) *Outdoor Environments*: For outdoor deployment, Hong et al. [4] detect blob features in Cartesian radar images and track them with a Kanade-Lucas-Tomasi tracker. They estimate motion by matching features between scans and solving them for the transform via Singular Value Decomposition. Loop closures are detected using the *M2DP* descriptor. Wang et al. [5] find linear features in filtered point clouds. They use scan-to-map matching based on a point-to-line metric. Features that appear infrequently are erased from the map. Loop closures are detected using the *ScanContext* Descriptor [13]. Adolfsson et al. [6] use the strongest returns in each azimuth to compute sparse oriented surface points. They estimate motion using a weighted point-to-point metric and detect loop closures using a combination of *ScanContext*, an odometry similarity measure, and alignment quality metrics.

### B. Normal Distributions Transform SLAM

The NDT subdivides a scan using a regular grid. All points within a grid cell are approximated by a normal distribution with mean  $\boldsymbol{\mu}$  and covariance  $\boldsymbol{\Sigma}$  defined as

$$\boldsymbol{\mu} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix}, \boldsymbol{\Sigma} = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}. \quad (1)$$

The indices  $x$  and  $y$  denote the geometrical dimensions. Two paradigms are used to find the affine transformation matrix

$$\mathbf{T} = \begin{pmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{pmatrix}, \quad (2)$$

with rotation matrix  $\mathbf{R}$  and translation vector  $\mathbf{t}$  that aligns a moving scan (MS) to a fixed scan (FS): Point-to-Distribution (P2D) and Distribution-to-Distribution matching (D2D). In P2D [14], the FS is converted into a NDT. Each point of the MS is evaluated on the NDT. The sum of all evaluated points is used as score function. The resulting transform is obtained by optimizing this score function. In D2D [15], the  $L_2$  distance between NDT  $\mathcal{M}$  of the MS and NDT  $\mathcal{F}$  of the FS is minimized. This yields the cost function

$$C_{d2d} = -d_1 \sum_{i=1}^{n_{\mathcal{M}}} \sum_{j=1}^{n_{\mathcal{F}}} \exp\left(-\frac{d_2}{2} \boldsymbol{\mu}_{ij}^T (\boldsymbol{\Sigma}_i + \mathbf{R} \boldsymbol{\Sigma}_j \mathbf{R}^T) \boldsymbol{\mu}_{ij}\right), \quad (3)$$

where  $d_1$  and  $d_2$  are positive regularization parameters, and  $\boldsymbol{\mu}_{ij} = \mathbf{R} \boldsymbol{\mu}_i + \mathbf{t} - \boldsymbol{\mu}_j$  is a short notation for the difference of the means of two distributions. The number of elements in the moving and fixed NDT are given by  $n_{\mathcal{M}}$  and  $n_{\mathcal{F}}$ , respectively. While developed for LiDAR, the NDT is frequently used for radar scan registration. Some approaches aim to use the radar's intensity information to improve registration accuracy. Heller, Petrov, and Yarovoy [16] use the intensity to weight each point in both, distribution calculation and alignment optimization. Kung, Wang, and Lin [17] extend their approach by shifting the returned power, further increasing the influence of high-intensity points. In addition, the covariance of each cell is set depending on the distance of the points to the sensor. Our approach is most similar to [16], but by using the intensity as an additional dimension in the normal distributions, we aim to increase the robustness, particularly in indoor environments.

### III. METHODOLOGY

In the following, we propose the framework **Radar Normal Distributions Transform (RaNDT) SLAM**. The information is organized in three hierarchical layers: states, keyframes, and submaps. Every scan at time  $t$  is associated with a state  $\mathbf{x}_t = \{\mathbf{T}_t, \mathbf{v}_t, b_t\}$ , with the 2D pose  $\mathbf{T}_t \in SE(2)$  given in the submap frame, the velocity  $\mathbf{v}_t \in \mathbb{R}^3$  given in the robot frame, and an IMU bias  $b_t$ . Each scan is filtered and subsequently used for local motion estimation. For further processing, we select keyframes. These are used to detect loop closures whose poses  $\mathbf{P}_k$  form the nodes of the pose graph. Lastly, small submaps  $\mathcal{F}$  are used to accumulate the measurements of the keyframes. This overcomes the possible sparseness of individual scans. An overview of the frameworks' architecture is given in Figure 1.

#### A. Map Representation

The map generated by our framework is a collection of small submaps. These are represented as NDTs. In contrast to the aforementioned approaches (see Section II), we augment the distribution's parameters introduced in Equation 1 with an intensity dimension  $p$ . The augmented parameters are

$$\tilde{\boldsymbol{\mu}} = \begin{pmatrix} \mu_x \\ \mu_y \\ \mu_p \end{pmatrix}, \tilde{\boldsymbol{\Sigma}} = \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} & \Sigma_{xp} \\ \Sigma_{yx} & \Sigma_{yy} & \Sigma_{yp} \\ \Sigma_{px} & \Sigma_{py} & \Sigma_{pp} \end{pmatrix}. \quad (4)$$

To obtain the distributions, we use the sample mean and covariance of all intensity-augmented points  $\tilde{\mathbf{p}} = [p_x \ p_y \ p_p]^T$

falling into a cell as proposed by Magnusson et al. [18]. With the additional intensity value in the augmented mean vector  $\hat{\boldsymbol{\mu}}$ , incorrect correspondences with largely differing intensities can be suppressed: large values inside the exponent of Equation 3 are treated as outliers and have no influence in the optimization. The augmented covariance matrix  $\hat{\boldsymbol{\Sigma}}$  captures not only the geometric distribution of the points, but expresses also the covariance between the intensity and the geometric coordinates. Hence, the covariance implicitly models the target material's reflection characteristics. Using both, augmented mean and covariance, we aim to improve the registration accuracy in ambiguous indoor applications, where materials reflect with different intensities. The submaps are built using the scans of all keyframes. New sensor data is integrated using the recursive update formulation by Saarinen et al. [19]. As the submaps only contain a small part of the environment, we do not explicitly track occupancy values within each cell.

### B. Radar Filtering

The indoor application of our framework makes aggressive radar filtering necessary. We propose using a two-step procedure. First, we discard low-intensity and far/close points. Afterwards, we extract the cluster around the most intense point of each radar beam. For a point to belong to the cluster, two requirements have to be fulfilled:

- The point has a smaller distance than a threshold to another point of the cluster.
- The point's intensity is lower than the intensity of the neighboring point closer to the point of highest intensity.

The choice of the intensity threshold should reflect the target environment and the sensor specifics. In indoor applications, higher values can suppress the stronger multi-path reflections. Outdoors, the threshold can be set lower. This efficiently reduces the scan size of the radar from initially approx. 10000 to 1500 points. While this aggressive filtering approach is susceptible to missing important secondary landmarks, our experiments show that competitive motion estimation accuracy performance can be achieved. The radar point cloud before and after filtering is depicted in Figure 2.

### C. Motion Estimation

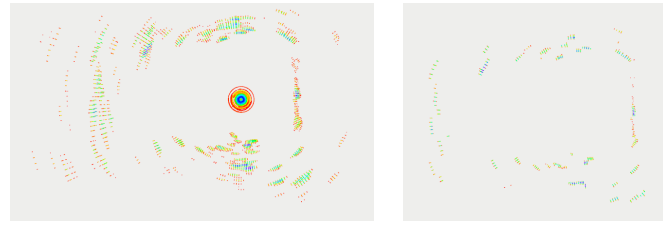
Local Motion estimation is necessary to build the submap and obtain an accurate estimate of the last state. Our state estimator uses three components to pose a least squares optimization problem (according indices annotated):

- Motion model (*mm*)
- Relative rotations obtained by an IMU (*imu*)
- NDT registration between scan and submap (*ndt*)

We explicitly not use wheel odometry as rescue robots usually employ a tracked drive train, leading to poor odometry. All components are jointly optimized using the cost function

$$C(\mathbf{x}_{t:t}) = \sum_{t_l}^t C_{ndt}(\mathbf{x}_i) + C_{mm}(\mathbf{x}_{i-1}, \mathbf{x}_i) + C_{imu}(\mathbf{x}_{i-1}, \mathbf{x}_i), \quad (5)$$

where  $t_l = t - l + 1$ . The cost function is optimized over  $l$  time steps. This optimization over a window of states forces



(a) Raw radar scan (indurad iSDR-C). (b) Filtered radar scan.

Fig. 2: Effect of the filter on the radar scan. Intensities are denoted red/low to purple/high; radome is circular entity in a.

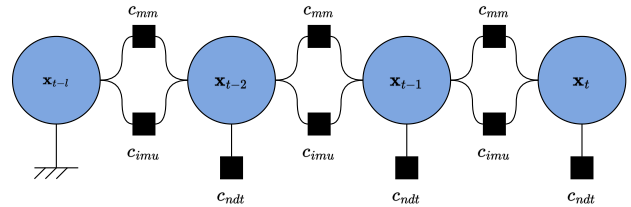


Fig. 3: Factor graph optimizing over  $l = 3$  states.

consistency over multiple poses. In ambiguous situations, this can prevent incorrect convergence. New scans are inserted into the map after leaving the optimization window. Selecting  $l$  is a trade-off between registration accuracy and computational efficiency. We experimentally determine that  $l = 3$  works well across multiple domains. As  $l$  also determines the delay before scan insertion, an improper value can lead to insufficient scan overlap, impairing the registration. Figure 3 depicts the factor graph corresponding to the cost function. Following, the individual components are introduced:

1) *Motion Model*: The motion model generates the pose prediction

$$\hat{\mathbf{T}}_t = \mathbf{T}_{t-1} \text{Exp}(\mathbf{v}_{t-1} \Delta t). \quad (6)$$

We use the notation in [20], where  $\text{Exp}(\cdot), \mathbb{R}^3 \rightarrow SE(2)$  denotes the exponential map, lifting an element of the tangent space to the group element. Accordingly, the logarithmic map  $\text{Log}(\cdot), SE(2) \rightarrow \mathbb{R}^3$  gives the tangent vector that corresponds to the group element. The motion cost is obtained as the Mahalanobis distance between prediction and estimation

$$C_{mm} = \left\| \begin{pmatrix} \text{Log}(\hat{\mathbf{T}}_t^{-1} \mathbf{T}_t) \\ \mathbf{v}_t - \mathbf{v}_{t-1} \end{pmatrix} \right\|_{\boldsymbol{\Omega}_{mm}}^2, \quad (7)$$

where  $\boldsymbol{\Omega}_{mm}$  is the information matrix of the motion model. The values of  $\boldsymbol{\Omega}_{mm}$  are selected empirically.

2) *IMU Data*: The measurements of the IMU are fused as a relative rotation  $\Delta \mathbf{R}_t$ , computed by a complementary filter [21]. The bias accounts for the drift of the angular velocity measured by the IMU's gyroscope. This gives the rotational cost function

$$C_{imu} = \left\| \begin{pmatrix} \text{Log}((\mathbf{R}_{t-1} \Delta \mathbf{R}_t \text{Exp}(b_{t-1} \Delta t))^{-1} \mathbf{R}_t) \\ b_t - b_{t-1} \end{pmatrix} \right\|_{\boldsymbol{\Omega}_{imu}}^2, \quad (8)$$

where  $\boldsymbol{\Omega}_{imu}$  is the information matrix of the IMU model.

3) *NDT Registration*: Two adjustments have to be made to integrate the NDT into the estimator. First, the parameters to optimize for ( $\mathbf{R}$  and  $\mathbf{t}$ ) have to be adjusted to be compliant with the augmented NDT cells. The augmented rotation matrix and the modified translation vector are given by

$$\tilde{\mathbf{R}} = \begin{pmatrix} \mathbf{R} & \mathbf{0} \\ \mathbf{0} & 1 \end{pmatrix}, \tilde{\mathbf{t}} = \begin{pmatrix} \mathbf{t} \\ 0 \end{pmatrix}. \quad (9)$$

This results in the augmented D2D-NDT cost function between the scan's NDT  $\mathcal{M}$  and submap  $\mathcal{F}$ , given by

$$C_{d2d} = -d_1 \sum_{i=1}^{n_{\mathcal{M}}} \sum_{j=1}^{n_{\mathcal{F}}} \exp\left(-\frac{d_2}{2} \tilde{\boldsymbol{\mu}}_{ij}^T (\tilde{\boldsymbol{\Sigma}}_i + \tilde{\mathbf{R}} \tilde{\boldsymbol{\Sigma}}_j \tilde{\mathbf{R}}^T) \tilde{\boldsymbol{\mu}}_{ij}\right). \quad (10)$$

Commonly, this function is optimized using Newton's method [15]. To express this in a least squares formulation, the second adjustment has to be made. Note, that the NDT cost function has the same shape and gradient as the Welsch loss [22]

$$\rho(r^2, c) = 1 - \exp\left(-\frac{r^2}{2c^2}\right), \quad (11)$$

with an offset of 1. Therefore, the NDT cost can equally be described by the sum of multiple quadratic error terms

$$r^2 = \tilde{\boldsymbol{\mu}}_{ij}^T (\tilde{\boldsymbol{\Sigma}}_i + \tilde{\mathbf{R}} \tilde{\boldsymbol{\Sigma}}_j \tilde{\mathbf{R}}^T) \tilde{\boldsymbol{\mu}}_{ij} \quad (12)$$

and a robust loss function  $\rho(r^2, c)$ , where  $c$  sets the scale of the loss function and  $r$  denotes the residual. Also other robust loss functions can be used. Each loss can be interpreted as resulting from a specific outlier distribution. Adaptive loss functions can interpolate between multiple variants and, therefore, allow for tailored tuning on different environment properties [22].

4) *Adaptive Loss Formulation*: To avoid falling into local minima, a frequently used approach is to register scans on multiple scales. In NDT matching, this is mostly accomplished by matching different map resolutions [19]. A different approach in scan matching is to adapt the scale  $c$  of the optimization problem. For a macroscopic scale, the objective function is easy to optimize, but may be affected by outliers. With increasingly smaller scales, the influence of outliers vanishes at the cost of more local minima. By solving a series of optimization problems with decreasing  $c$ , the risk of incorrect convergence can be mitigated [23]. To leverage both, the adaptive loss formulation and the multi-scale optimization, we use the adaptive loss

$$\rho(r^2, \alpha, c, \mu) = \frac{|\alpha - 2|}{\alpha} \left( \left( \frac{(r^2 / (\mu c^2))}{|\alpha - 2|} + 1 \right)^{\frac{\alpha}{2}} - 1 \right). \quad (13)$$

This formulation is derived from [22], where the shape parameter  $\alpha$  is introduced to interpolate different loss functions. In addition, the parameter  $\mu$  is used to adjust the scale. In each iteration  $n$ , the parameter is reduced using  $\mu_{n+1} \leftarrow \mu_n / k_\mu$ . In the final iteration, the control parameter reaches  $\mu = 1$ .

5) *Resulting NDT Cost Function*: The double sum in Equation 10 is computationally expensive, hence, only a subset of closest correspondences is considered. For each normal distribution in the scan's NDT, we select the four closest distributions in the submap. This yields the set of correspondences  $\mathcal{K}$ . The resulting cost function is given by

$$C_{ndt} = \frac{w_{ndt}}{n_{\mathcal{K}}} \sum_{(i,j) \in \mathcal{K}} \rho\left(\tilde{\boldsymbol{\mu}}_{ij}^T (\tilde{\boldsymbol{\Sigma}}_i + \tilde{\mathbf{R}} \tilde{\boldsymbol{\Sigma}}_j \tilde{\mathbf{R}}^T) \tilde{\boldsymbol{\mu}}_{ij}, \alpha, c, \mu\right), \quad (14)$$

where  $w_{ndt}$  is empirically tuned to set the influence of the NDT in the optimization. The final solution is obtained using *Ceres* [24] and the Levenberg-Marquardt algorithm.

#### D. Loop Closure Detection

For loop closure detection, we make use of the method proposed in [6]. Here, the filtered scan of each keyframe is used in its polar representation to generate the *ScanContext* Descriptor  $\mathbf{I}$ . The scan is binned in angular and radial direction. Each element of the descriptor is calculated as the sum of the intensities of all points divided by 20 within the bin. When creating a new descriptor, a loop closure search is triggered. Given a query scan  $q$ , the aim is to find the most likely loop candidate  $lc$ . This likelihood is expressed using a combination of the descriptor distance  $d_{sc}(\mathbf{I}^q, \mathbf{I}^{lc})$  and an odometry similarity  $d_{od}^{q,lc}$ , given by

$$lc = \underset{lc}{\operatorname{argmin}} \left( d_{sc}(\mathbf{I}^q, \mathbf{I}^{lc}) + d_{od}^{q,lc} \right). \quad (15)$$

The highest-scoring loop candidate is further refined by matching the query scan with the submap containing the loop candidate. After alignment, we use the Cauchy-Schwarz divergence between the NDTs of both, query scan and the loop candidate's submap, as gate to accept or reject the loop closure as described in [25]. This divergence between two probability density functions  $p(x)$  and  $q(x)$  is given by

$$DCS(q, p) = -\log \left( \frac{\int q(x)p(x)dx}{\sqrt{\int q(x)^2 dx \int p(x)^2 dx}} \right), \quad (16)$$

which can be expressed in closed form for Gaussian Mixture Models such as the NDT [26]. If the divergence exceeds a threshold, the loop closure is discarded. Otherwise, a constraint between the query node and the root node of the loop candidate's submap is added to the pose graph.

#### E. Global Optimization

Global consistency is ensured by regularly optimizing the pose graph. The graph consists of keyframe poses  $\mathbf{P}$ , odometry constraints  $\mathcal{C}_{od}$ , and loop constraints  $\mathcal{C}_{lo}$ . To obtain the optimal trajectory, the cost function

$$C_{pg}(\mathbf{P}) = \sum_{(a,b) \in \mathcal{C}_{od}} (\mathbf{e}_{ab}^T \boldsymbol{\Omega}_{od} \mathbf{e}_{ab}) + \sum_{(a,b) \in \mathcal{C}_{lo}} (\mathbf{e}_{ab}^T \boldsymbol{\Omega}_{lo} \mathbf{e}_{ab}) \quad (17)$$

is minimized using the Levenberg-Marquardt algorithm. Here, the information matrices  $\boldsymbol{\Omega}_{od}$  and  $\boldsymbol{\Omega}_{lo}$  are set to constant values. The pose errors  $\mathbf{e}_{ab}$  quantify the difference between the relative transform between the pose graph nodes and the relative transform stored in the constraint.

#### IV. EVALUATION

We evaluate the performance of our approach in two new benchmark scenarios [10] to demonstrate the feasibility in varying environments. We discuss the most important parameters and their configuration. The resulting trajectories are benchmarked<sup>6</sup> against LiDAR SLAM. We compare our framework to state-of-the-art radar odometry and SLAM approaches based on the Oxford dataset [27]. To evaluate RaNDT SLAM, the estimated trajectory  $\mathbf{P}_1, \dots, \mathbf{P}_n \in SE(2)$  is compared to a ground truth trajectory  $\mathbf{Q}_1, \dots, \mathbf{Q}_n \in SE(2)$ . We evaluate two performance metrics: Relative Pose Error (RPE) and Absolute Trajectory Error (ATE) [28]. In line with [9], we use the mean variant of the RPE. For the ATE, we use the Root Mean Squared Error (RMSE) of the translation errors between the estimated and the ground truth poses:

$$T\text{-RPE} = \frac{1}{n-1} \sum_{i=1}^{n-1} \|\text{trans}(\mathbf{E}_i)\|, \quad (18a)$$

$$R\text{-RPE} = \frac{1}{n-1} \sum_{i=1}^{n-1} \|\text{rot}(\mathbf{E}_i)\|, \quad (18b)$$

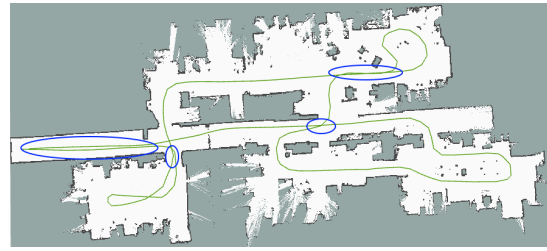
$$\text{ATE} = \sqrt{\frac{1}{n} \sum_{i=1}^n \|\text{trans}(\mathbf{F}_i)\|^2}, \quad (18c)$$

with the translational ( $\text{trans}(\cdot)$ ) and a rotational ( $\text{rot}(\cdot)$ ) component, the mean of all individual errors  $\mathbf{E}_i$ , and  $\mathbf{F}_i = \mathbf{Q}_i^{-1}\mathbf{P}_i$ . The poses of the ground truth trajectory are obtained at a different rate than the estimated poses. For every ground truth pose, we interpolate linearly between the two closest estimated poses. For the evaluation on the Oxford dataset (Section IV-E), we, additionally, employ the KITTI odometry error [29]. This metric measures the drift by computing the translation error and rotation error over trajectory segments of length 100, 200, ... , 800 m. The drift is obtained by averaging over all evaluated segments. This metric is not applicable to the RaNDT SLAM benchmark, as the trajectories are not sufficiently long.

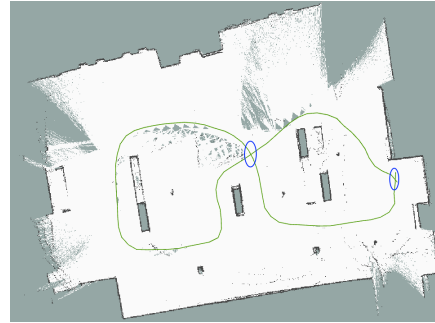
##### A. RaNDT SLAM Dataset

The first scenario in the RaNDT SLAM dataset is recorded in a lab space (Figure 4a). The test environment contains a long corridor, two labs, and a cluttered workshop. Multiple loops occur in the trajectory. The total trajectory is about 137 m long and is traversed in 308 s. The low velocity accounts for the expected slow operating speed in rescue robotics applications. To assess the outdoor performance, a second scenario is recorded in a sparse courtyard (Figure 4b). The robot drives in an eight-like shape, containing two possible loop closures. Compared to the indoor scenario, the outdoor setting offers less features and has varying inclination angles. Latter violates the 2D assumption and further tests the robustness. All scenarios are recorded using a tracked robot equipped with an iSDR-C radar sensor by indurad and a Phidgets Spatial IMU. The radar sensor offers a range resolution of 16.1 mm and an angular resolution of 1.64°. The radar spins at 5 Hz and has a maximum range of 16 m. The benchmark dataset is publicly available [10].

<sup>6</sup>All experiments are performed on a PC running Ubuntu 20.04 in Docker with a Ryzen 2600X CPU at 3.6 GHZ and 16 GB RAM, without GPU.



(a) Indoor environment.



(b) Outdoor environment.

Fig. 4: RaNDT SLAM scenarios (green: LiDAR trajectory, blue: potential loop closures); maps generated using LiDAR.

##### B. Parameters

We prepare three parameter configurations using an extensive parameter study: A specialized configuration for each, indoor and outdoor, application, and a mixed configuration suitable for both. Most prominently, the resolution of the map has to be set properly for the outdoor application. Indoors, a resolution of 0.5 m yields great results, but fails outdoors. Hence, the outdoor resolution is set to 1.2 m. The mixed configuration uses a resolution of 1 m to trade off accuracy in both, indoor and outdoor environment. In addition, we observe that the outdoor configuration requires a larger shape  $\alpha$  and a larger scale  $c$  compared to the indoor one. This is in line with the expected higher outlier ratio indoors produced by stronger multi-path reflections. For indoor usage, we use  $\alpha = -2.0$  and  $c = 1.5$ . Outdoors, we choose  $\alpha = -1.0$  and  $c = 2.0$ . The mixed configuration is set to  $\alpha = -1.5$  and  $c = 2.0$ . The complete parameter configurations may be found in our code publication of RaNDT SLAM<sup>5</sup>. While the parameter settings for the Oxford dataset are based on the outdoor configuration, several adjustments need to be made: Firstly, the intensity threshold is set to  $p = 70$ , as the Navtech sensor returns higher intensity values. In addition, we set the resolution to 3.5 m to account for the larger scale of the environment. Finally, we reduce the number of scans that are inserted into a submap. This accounts for the higher velocity of the robot in autonomous driving scenarios.

##### C. Evaluation with RaNDT Radar Benchmark

Within our new dataset, we use *SlamToolbox* [30] with LiDAR data to generate a ground truth trajectory. The comparison of radar with LiDAR trajectories is commonly applied in absence of more accurate ground truth measures,

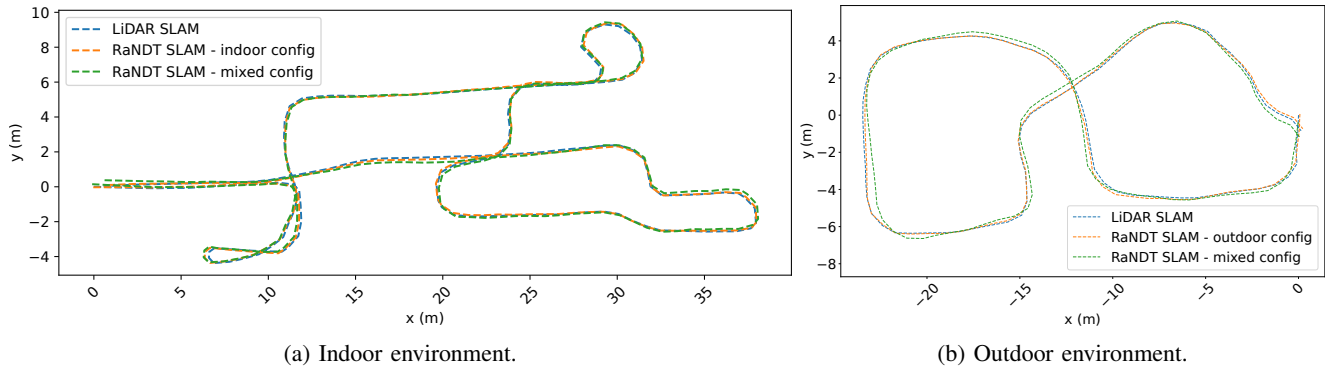


Fig. 5: Estimated trajectories in different environments, using specific and mixed parameter configurations.

see, e.g., [9]. However, it should be noted that the state of the art in radar SLAM approaches the accuracy of LiDAR SLAM [6]. We perform ten runs for each scenario and in each parameter configuration. This accounts for slight variations in the performance measures due to the multi-threaded implementation. The indoor environment is tested with the indoor and mixed configuration, and the outdoor environment with the outdoor and mixed configuration. Using the indoor configuration on the outdoor environment and vice versa proved to be infeasible (compare Section IV-B). In addition, we evaluate the specialized configurations without loop closure to demonstrate the improved performance enabled by the loop closure module. Table I lists ATE and RPE measures of the tests using the LiDAR trajectory as ground truth.

1) *Indoor*: In the indoor scenario, both, the indoor and the mixed configuration, show accurate performance (Figure 5a). Regarding ATE and R-RPE, the mixed configuration performs only slightly worse. The T-RPE increases by 18.8%, indicating that the local accuracy deteriorates. Here it can be observed, that the trajectory obtained with the mixed configuration deviates slightly from ground truth on long straights. However, we conclude that both configurations are suitable for indoor usage. Comparing the performance gain enabled by the loop closure, we observe an ATE reduction of 57%, underlining the importance of the loop closure. The RPEs increase slightly.

2) *Outdoor*: The outdoor trajectories generated by outdoor and mixed configuration are depicted in Figure 5b. The trajectory estimated by the outdoor configuration closely aligns to the ground truth trajectory. This is also reflected in the low ATE measure. When using the mixed configuration, the ATE increases significantly. We observe that the overall trajectory quality remains high, but the long straight (at  $x \approx -23$  m) deviates. Hence, the mixed configuration deteriorates more in outdoor environments. However, as the pure outdoor configuration is infeasible in indoor environment and vice versa, the mixed configuration is a good choice in case of the robot switching between outdoor and indoor environments, e.g., in case of large scale disaster sites with multiple buildings. In the outdoor environment, we observe an ATE reduction by 32%. The gain is less than indoors, which can be explained by the smaller amount of loop closures. In

TABLE I: Performance metrics generated by RaNDT SLAM with (w.) and without (w/o) loop closure (LC).

Environment	Parameter Config.	Indoor		Outdoor	
		Indoor	Mixed	Outdoor	Mixed
w. LC	ATE [m]	0.143995	0.169776	0.242701	0.422452
	T-RPE [m]	0.033777	0.040120	0.061772	0.063413
	R-RPE [°]	1.346297	1.390718	1.776593	1.925730
w/o LC	ATE [m]	0.337302	—	0.354598	—
	T-RPE [m]	0.032732	—	0.059621	—
	R-RPE [°]	1.314041	—	1.706450	—

TABLE II: Cycle times of individual subsystems in RaNDT SLAM (SCD: *ScanContext* Descriptor).

Subtask	Duration indoor [ms]	Duration outdoor [ms]
Pre-Processing	2.743 ± 0.923	1.448 ± 0.043
Motion Estimation	35.671 ± 15.964	18.404 ± 8.603
SCD Matching	1.118 ± 0.385	2.195 ± 0.844
Loop Refinement	14.266 ± 5.833	7.214 ± 11.484
Graph Optimization	54.833 ± 34.623	12.660

conclusion, RaNDT SLAM indicates high accuracy in both, indoor and outdoor, environments. The loop closure module improves the global consistency of the estimated trajectory significantly. When targeting a specific environment or use-case scenario, the performance can be further enhanced by deploying the according parameter configuration.

#### D. Computational Cost

To assess the real-time capability, we evaluate the average duration spent by the most important subsystems (Table II). We measure the graph optimization each time new loop closure constraints are added to the graph. No robust measure for the standard deviation can be given in the outdoor sequence, as only two loop closures are detected in the trajectory. The pre-processing and motion estimation steps are performed synchronous. Therefore, the time spent on a complete cycle of RaNDT SLAM is significantly smaller than the sensor’s measurement cycle of 200 ms. This allows the online usage for autonomous operation and is promising for future radar sensors with higher scanning rate. The increased processing times in the indoor application are caused by the finer resolution of the NDT map. This leads to NDTs with more components, which makes the optimiza-

TABLE III: Evaluation on Oxford dataset compared to state of the art.

Method	10-12-32	16-13-09	17-13-26	18-14-14	18-15-20	10-11-46	16-11-53	18-14-46	Mean	
<b>Drift (in % / (°/100 m)) on the Oxford dataset comparing SLAM and odometry methods.</b>										
Odometry	Cen2019 [31]	-	-	-	-	-	-	-	3.72/0.95	
	MC-RANSAC [32]	-	-	-	-	-	-	-	3.31/1.09	
	CC-means [33]	-	-	-	-	-	-	-	2.53/0.82	
	Under the radar [34]	-	-	-	-	-	-	-	2.06/0.67	
	R <sup>3</sup> O [35]	1.89/0.56	1.95/0.60	2.09/0.58	2.06/0.61	2.18/0.65	2.02/0.57	2.17/0.60	1.96/0.57	2.04/0.59
	MbyM [36]	-	-	-	-	-	-	-	-	1.16/0.30
	HERO [37]	1.77/0.62	1.75/0.95	2.04/0.73	1.83/0.61	2.20/0.77	2.14/0.71	2.01/0.61	1.97/0.65	1.96/0.66
	Kung [17]	-	-	-	-	-	-	-	-	1.96/0.60
	SDRO [38]	1.43/0.32	1.58/0.24	1.52/0.27	1.58/0.35	1.66/0.24	1.63/0.32	1.78/0.34	1.68/0.38	1.65/0.30
	CLEAR-3 [39]	1.20/0.36	1.24/0.40	1.23/0.39	1.35/0.42	1.24/0.41	1.22/0.39	1.39/0.40	1.39/0.44	1.28/0.40
CLEAR-3-s50 [39]	<b>1.05/0.34</b>	<b>1.08/0.34</b>	<b>1.07/0.36</b>	<b>1.11/0.37</b>	<b>1.03/0.37</b>	<b>1.05/0.36</b>	<b>1.18/0.36</b>	<b>1.11/0.36</b>	<b>1.09/0.36</b>	
ours SLAM	RadarSLAM-Full [4]	1.98/0.60	1.48/0.50	1.71/0.50	2.22/0.70	1.77/0.60	1.96/0.70	1.81/0.60	1.68/0.50	1.83/0.60
	MAROAM [5]	1.63/0.46	1.83/0.56	1.49/0.47	1.54/0.47	1.61/0.50	1.55/0.53	1.78/0.54	1.55/0.50	1.62/0.50
	TBV Radar SLAM [6]	1.17/0.35	1.15/0.35	<b>1.06/0.35</b>	1.12/0.37	1.09/0.36	1.18/0.35	1.32/0.36	<b>1.10/0.36</b>	1.15/0.36
	<b>RaNDT SLAM w/o LC</b>	1.51/0.49	1.53/0.50	1.55/0.50	1.49/0.51	1.31/0.46	1.51/0.47	2.03/0.50	1.92/0.65	1.60/0.52
<b>RaNDT SLAM w. LC</b>	1.52/0.50	1.54/0.49	1.40/0.46	1.45/0.49	1.19/0.44	1.44/0.46	1.72/0.52	1.72/0.60	1.50/0.50	
<b>ATE (in m) on the Oxford dataset comparing SLAM methods.</b>										
RadarSLAM-Full [4]	9.59	11.18	5.84	21.21	7.74	13.78	7.14	6.01	10.31	
MAROAM [5]	13.76	6.95	8.36	10.34	10.96	12.42	12.51	7.71	10.38	
TBV Radar SLAM-dyn-cov [6]	<u>4.22</u>	4.30	<b>3.37</b>	<u>4.04</u>	4.27	<u>3.38</u>	<u>4.98</u>	<b>4.27</b>	<u>4.10</u>	
TBV Radar SLAM [6]	<b>4.07</b>	<u>4.04</u>	<u>3.58</u>	<b>3.79</b>	<b>3.83</b>	<b>3.14</b>	<b>4.39</b>	<u>4.33</u>	<b>3.90</b>	
<b>RaNDT SLAM w. LC (ours)</b>	5.40	<b>4.00</b>	5.89	6.45	<u>4.03</u>	5.30	10.36	6.27	5.96	

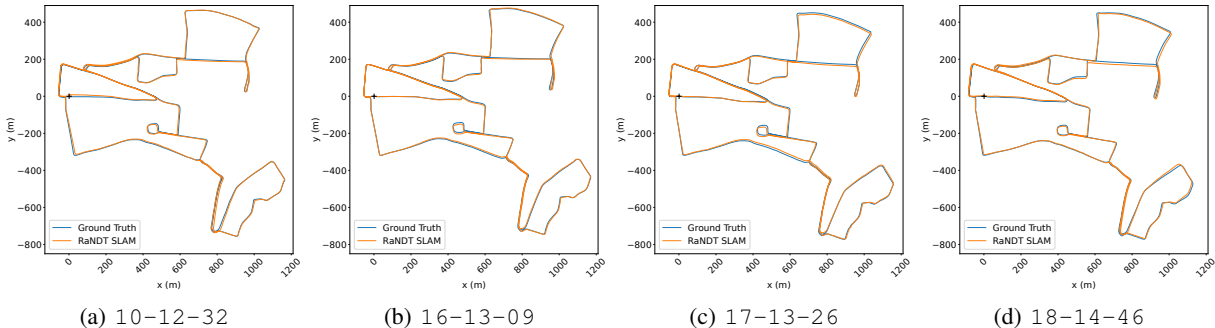


Fig. 6: Selected Oxford trajectories estimated by RaNDT SLAM compared to GT. Start positions are marked (+).

tion more time consuming. Moreover, we observe that the indoor scenario takes approximately twice as long. This is because more distributions are generated, which makes the optimization harder. Note, that the cycle times are parameter-dependent and need to be considered as a trade-off between computational times, resolution and expected noise ratio.

#### E. Evaluation with Oxford Radar RobotCar Dataset

To underline the applicability of RaNDT SLAM in different environments and with different sensor models, we test RaNDT SLAM on the Oxford dataset [27], using the provided GT obtained via visual odometry, visual loop closures, and GPS constraints. This allows us to compare our results with state-of-the-art radar SLAM and radar odometry approaches, partially introduced in Section II. Table III depicts the KITTI odometry errors and the ATEs. As the ATE largely favors algorithms using loop closure, we present only full SLAM systems and omit results obtained with radar odometry. The ATE achieved with RaNDT SLAM scores second-best just behind TBV SLAM [6]. In addition, the translational drift is able to keep up with other radar SLAM and radar odometry methods. Notably, we perform better than the other NDT-based methods regarding the translational drift [4], [38]. The rotational drift, however, is less competitive. We

attribute this to the missing motion compensation in our framework, as motion distortion compromises the measurements especially during turns of the vehicle. We conclude, that RaNDT SLAM generalizes well to different domains and sensor types, although being designed for smaller-scale and indoor environments. We expect the performance to increase further if techniques tailored for autonomous driving datasets such as motion compensation [32] or origin shifting in descriptor generation [13] are implemented.

#### V. CONCLUSION AND FUTURE WORK

In this work, we proposed a novel radar SLAM framework, RaNDT SLAM, suited for both indoor and outdoor environments. Within, we combined NDT registration, IMU measurements, and a motion model. We validated our approach using indoor and outdoor scenarios and reached high trajectory accuracy in both cases. Furthermore, we evaluated the computational effort and showed that we achieve real-time performance. This paves the way to robust, autonomous operation in small-scale, vision-denied environments. Further, we compared RaNDT SLAM to the state of the art in large-scale radar SLAM and odometry using the Oxford Radar RobotCar dataset, which underlines its good performance even outside its designed scope. Code<sup>5</sup> and benchmark dataset [10] have

been made publicly available, setting a new baseline for radar SLAM in rescue robotics environments. Possible extensions of our work may integrate more sensor data to improve the accuracy in case of low vision restriction or environments with switching conditions. The flexible estimator structure can easily be extended, e.g., using Doppler measurements. We aim to deploy RaNDT SLAM with the German Rescue Robotics Center, to access its performance in application.

## ACKNOWLEDGMENT

We like to thank indurad, particularly Matthias Rabel and Stephan Renner, for providing the iSDR-C sensors.

## REFERENCES

- [1] P. Fritsche, S. Kueppers, G. Briese, and B. Wagner, "Fusing lidar and radar data to perform slam in harsh environments," in *International Conference on Informatics in Control, Automation and Robotics*. Springer, Cham, 2018, pp. 175–189.
- [2] P. Checchin, F. Gérossier, C. Blanc, R. Chapuis, and L. Trassoudaine, "Radar scan matching slam using the fourier-mellin transform," in *Field and Service Robotics*. Springer, 2010, pp. 151–161.
- [3] F. Schuster, C. G. Keller, M. Rapp, M. Haueis, and C. Curio, "Landmark based radar slam using graph optimization," in *International Conference on Intelligent Transportation Systems*. IEEE, 2016, pp. 2559–2564.
- [4] Z. Hong, Y. Petillot, A. Wallace, and S. Wang, "Radarslam: A robust simultaneous localization and mapping system for all weather conditions," *The International Journal of Robotics Research*, vol. 41, no. 5, pp. 519–542, 2022.
- [5] D. Wang, Y. Duan, X. Fan, C. Meng, J. Ji, and Y. Zhang, "Maroam: Map-based radar slam through two-step feature selection," 25.10.2022. [Online]. Available: <https://arxiv.org/pdf/2210.13797>
- [6] D. Adolfsson, M. Karlsson, V. Kubelka, M. Magnusson, and H. Andreasson, "Tbv radar slam – trust but verify loop candidates," *Robotics and Automation Letters*, vol. 8, no. 6, pp. 3613–3620, 2023.
- [7] J. W. Marck, A. Mohamoud, E. vd Houwen, and R. van Heijster, "Indoor radar slam a radar application for vision and gps denied environments," in *European Radar Conference*, 2013, pp. 471–474.
- [8] M. Mielle, M. Magnusson, and A. J. Lilienthal, "A comparative analysis of radar and lidar sensing for localization and mapping," in *European Conference on Mobile Robots*, 2019, pp. 1–6.
- [9] M. Torchalla, M. Schnaubelt, K. Daun, and O. von Stryk, "Robust multisensor fusion for reliable mapping and navigation in degraded visual conditions," in *International Symposium on Safety, Security, and Rescue Robotics*. IEEE, 2021, pp. 110–117.
- [10] M. Hilger and N. Mandischer, "Randt radar benchmark," 2023. [Online]. Available: <https://zenodo.org/record/8199947>
- [11] P. Fritsche and B. Wagner, "Scanning techniques with low bandwidth radar for robotic mapping and localization," in *International Conference of Informatics in Control, Automation and Robotics*, 2015.
- [12] N. Mandischer, S. Charaf Eddine, M. Hüsing, and B. Corves, "Radar slam for autonomous indoor grinding," in *Radar Conference*. IEEE, 2020, pp. 1–6.
- [13] G. Kim, S. Choi, and A. Kim, "Scan context++: Structural place recognition robust to rotation and lateral variations in urban environments," *Transactions on Robotics*, vol. 38, no. 3, pp. 1856–1874, 2022.
- [14] P. Biber and W. Strasser, "The normal distributions transform: a new approach to laser scan matching," in *International Conference on Intelligent Robots and Systems*. IEEE/RSJ, 2003, pp. 2743–2748.
- [15] T. Stoyanov, M. Magnusson, H. Andreasson, and A. J. Lilienthal, "Fast and accurate scan registration through minimization of the distance between compact 3d ndt representations," *The International Journal of Robotics Research*, vol. 31, no. 12, pp. 1377–1393, 2012.
- [16] M. Heller, N. Petrov, and A. Yarovoy, "A radar-oriented approach to the normal distributions transform," in *European Radar Conference*, 2022, pp. 165–168.
- [17] P.-C. Kung, C.-C. Wang, and W.-C. Lin, "A normal distribution transform-based radar odometry designed for scanning and automotive radars," in *International Conference on Robotics and Automation*. IEEE, 2021, pp. 14417–14423.
- [18] M. Magnusson, A. Lilienthal, and T. Duckett, "Scan registration for autonomous mining vehicles using 3d-ndt," *Journal of Field Robotics*, vol. 24, no. 10, pp. 803–827, 2007.
- [19] J. Saarinen, H. Andreasson, T. Stoyanov, J. Ala-Luhtala, and A. J. Lilienthal, "Normal distributions transform occupancy maps: Application to large-scale online 3d mapping," in *International Conference on Robotics and Automation*. IEEE, 2013, pp. 2233–2238.
- [20] J. Solà, J. Deray, and D. Atchuthan, "A micro lie theory for state estimation in robotics," 2018. [Online]. Available: <http://arxiv.org/pdf/1812.01537v9>
- [21] R. G. Valenti, I. Dryanovski, and J. Xiao, "Keeping a good attitude: A quaternion-based orientation filter for imus and margs," *Sensors*, vol. 15, no. 8, pp. 19302–19330, 2015.
- [22] J. T. Barron, "A general and adaptive robust loss function," in *Conference on Computer Vision and Pattern Recognition*. IEEE/CVF, 2019.
- [23] H. Yang, P. Antonante, V. Tzoumas, and L. Carlone, "Graduated non-convexity for robust spatial perception: From non-minimal solvers to global outlier rejection," *Robotics and Automation Letters*, vol. 5, no. 2, pp. 1127–1134, 2020.
- [24] S. Agarwal and K. Mierle, "Ceres solver," 2022. [Online]. Available: <https://github.com/ceres-solver/ceres-solver>
- [25] W. Tabib and N. Michael, "Simultaneous localization and mapping of subterranean voids with gaussian mixture models," in *Field and Service Robotics*. Springer, Singapore, 2021, pp. 173–187.
- [26] K. Kampa, E. Hasanbelliu, and J. C. Principe, "Closed-form cauchy-schwarz pdf divergence for mixture of gaussians," in *International Joint Conference on Neural Networks*, 2011, pp. 2578–2585.
- [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 *International Conference on Robotics and Automation*. IEEE, 2020, pp. 6433–6438.
- [28] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of rgb-d slam systems," in *International Conference on Intelligent Robots and Systems*. IEEE/RSJ, 2012, pp. 573–580.
- [29] 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*. IEEE, 2012, pp. 3354–3361.
- [30] S. Macenski and I. Jambrecic, "Slam toolbox: Slam for the dynamic world," *Journal of Open Source Software*, vol. 6, no. 61, p. 2783, 2021.
- [31] S. H. Cen and P. Newman, "Radar-only ego-motion estimation in difficult settings via graph matching," in *International Conference on Robotics and Automation*. IEEE, 2019, pp. 298–304.
- [32] K. Burnett, A. P. Schoellig, and T. D. Barfoot, "Do we need to compensate for motion distortion and doppler effects in spinning radar navigation?" *Robotics and Automation Letters*, vol. 6, no. 2, pp. 771–778, 2021.
- [33] R. Aldera, M. Gadd, D. De Martini, and P. Newman, "What goes around: Leveraging a constant-curvature motion constraint in radar odometry," *Robotics and Automation Letters*, vol. 7, no. 3, pp. 7865–7872, 2022.
- [34] D. Barnes and I. Posner, "Under the radar: Learning to predict robust keypoints for odometry estimation and metric localisation in radar," in *international conference on robotics and automation*. IEEE, 2020, pp. 9484–9490.
- [35] D. L. S. Lubanco, A. Hashem, M. Pichler-Scheder, A. Stelzer, R. Feger, and T. Schlechter, "R3o: Robust radon radar odometry," *Transactions on Intelligent Vehicles*, pp. 1–17, 2023.
- [36] D. Barnes, R. Weston, and I. Posner, "Masking by moving: Learning distraction-free radar odometry from pose information," in *Conference on Robot Learning*, 2019.
- [37] K. Burnett, D. J. Yoon, A. P. Schoellig, and T. Barfoot, "Radar Odometry Combining Probabilistic Estimation and Unsupervised Feature Learning," in *Proceedings of Robotics: Science and Systems*, 2021.
- [38] R. Zhang, Y. Zhang, D. Fu, and K. Liu, "Scan denoising and normal distribution transform for accurate radar odometry and positioning," *Robotics and Automation Letters*, vol. 8, no. 3, pp. 1199–1206, 2023.
- [39] D. Adolfsson, M. Magnusson, A. Alhashimi, A. J. Lilienthal, and H. Andreasson, "Lidar-level localization with radar? the cfar approach to accurate, fast, and robust large-scale radar odometry in diverse environments," *Transactions on Robotics*, vol. 39, no. 2, pp. 1476–1495, 2023.