

Robust Multisensor Fusion for Reliable Mapping and Navigation in Degraded Visual Conditions

Moritz Torchalla, Marius Schnaubelt, Kevin Daun and Oskar von Stryk

Abstract—We address the problem of robust simultaneous mapping and localization in degraded visual conditions using low-cost off-the-shelf radars. Current methods often use high-end radar sensors or are tightly coupled to specific sensors, limiting the applicability to new robots. In contrast, we present a sensor-agnostic processing pipeline based on a novel forward sensor model to achieve accurate updates of signed distance function-based maps and robust optimization techniques to reach robust and accurate pose estimates. Our evaluation demonstrates accurate mapping and pose estimation in indoor environments under poor visual conditions and higher accuracy compared to existing methods on publicly available benchmark data.

I. INTRODUCTION

To perform missions within unknown, GNSS-denied and degraded environments, autonomous mobile rescue robots need to localize themselves and create a map of the environment using a Simultaneous Localization and Mapping (SLAM) approach. The ability to create such a map and locate the robot's pose in it are a key prerequisite for many higher-level autonomous functions such as navigation or exploration.

Current SLAM systems typically strongly depend on accurate visual sensors, e.g. Light Detection and Ranging (lidar) [1] or cameras [2]. However, these sensors are prone to fail under visually degraded conditions such as smoke, dust, or fog, leading to failure in both localization and mapping. These conditions are common in disaster scenarios, such as dense smoke in a burning building or dust after a building collapse. Therefore, being able to localize under visually degraded conditions is a crucial capability to efficiently support first responders.

In contrast to lidar and camera, radio detection and ranging (radar) sensors are mostly unaffected by visually degraded conditions and therefore well suited for applications in such environments. However, they typically have an overall reduced data accuracy and density, which makes localization and mapping challenging. Current methods [3, 4] often use high-end sensors or are tightly coupled to specific sensors.

We propose a holistic approach that covers the aspects of sensor setup, low-level data processing, and SLAM. To

All authors are with the Simulation, Systems Optimization and Robotics Group, Technical University of Darmstadt, Hochschulstr. 10, 64289 Darmstadt, Germany {torchalla, schnaubelt, daun, stryk}@sim.tu-darmstadt.de

This work has been co-funded by the LOEWE initiative (Hesse, Germany) within the emergenCITY center. Research presented in this paper has been supported in parts by the German Federal Ministry of Education and Research (BMBF) within the subproject “Autonomous Assistance Functions for Ground Robots” of the collaborative A-DRZ project (grant no. 13N14861).

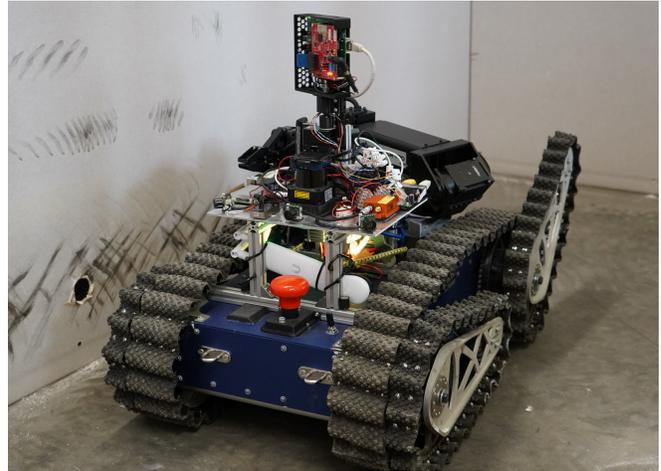


Fig. 1: The tracked evaluation platform “Asterix” carrying the navigation module with rotating radar, IMU and lidar. The lidar is only used for evaluation purposes.

reach a high sensor coverage, we mount the radar sensor on an actuated axis. The raw Multiple Input Multiple Output (MIMO) radar measurements are processed by Fast Fourier Transformations (FFTs), to reach accurate distance and velocity profiles of the environment. The data is then filtered by Constant False Alarm Rates (CFARs), to achieve a robust estimation of 2D range observations, similar to a pointcloud from lidar sensors, which we use as input for the SLAM system. Our SLAM approach builds upon the truncated signed distance functions (TSDF)-based lidar SLAM approach by Daun et. al [5] which is based on Cartographer [6], an open-source SLAM system that implements scan-to-map matching and loop closure detection. We extend the Cartographer TSDF approach with a novel forward sensor model to enable building TSDF maps from radar data and introduce robust scan matching algorithms to handle sparse, noisy, and outlier-rich radar range measurements. The key contributions are

- Design of an actuated system for optimal sensor coverage
- Novel forward sensor model to generate TSDF from radar data
- Robust scan matching method for sparse, noisy and outlier-rich radar range measurements

We evaluate our approach in room-scale indoor environments, demonstrating accurate mapping and pose estimation. On publicly available benchmark data, our approach yields higher accuracy compared to existing methods. Our imple-

mentation and experimental data is publicly available¹.

II. RELATED WORK

To support first responders with rescue robots in disaster scenarios, Kim et al. [7] propose the fusion of Frequency-Modulated Continuous Waves (FMCW) radar and stereo infrared (IR) cameras as both sensors operate well in smoke-filled environments. Their approach however only fuses the radar range measurements with the IR cameras for a better depth perception, no pose estimation is performed. O’Toole et al. [8] presented a novel 3D sensor combining a low-power laser projector and a rolling-shutter camera that can penetrate smoke thanks to its energy efficiency.

Mandischer et al. [9] developed a custom 2D radar sensor and SLAM method for localization in heavily dusted environments using a probabilistic Iterative Correspondence (pIC) approach combined with a clustering-based radar point filter.

Lu et al. [10] overcome the sparsity of radar data by using a conditional Generative Adversarial Networks (GAN) supervisedly trained using lidar data. The GAN generates dense patches and grid maps from low-cost off-the-shelf radar scans, comparable to the ones generated using lidar scans. Instead of performing scan matching, accurate odometry is used for pose estimation, which will cause a drift over time.

Kramer et al. [11] created a millimeter-wave radar data set that is focused on evaluating robotic perception in visually degraded environments. It includes dense, high-resolution millimeter-wave radar scans from two FMCW, 3D lidar, IMU, and highly accurate ground truth data.

Besides small indoor approaches, there is also an increasing number of contributions for large-scale radar SLAM, especially in the automotive sector due to the robustness of radar in changing environmental conditions. More recent work shows the capabilities of radar-based SLAM in large-scale environments utilizing a static automotive radar sensor [4]. Keenan Burnett [12] investigated the effect of motion distortion on radar-based navigation created by spinning radar sensors.

III. METHOD

Our method for radar-based SLAM in degraded visual conditions covers three aspects: design of a sensor module for optimal sensor coverage, processing of raw radar data, and the robust SLAM approach.

A. Sensor Module

The rotating radar module is equipped with a Texas Instruments (TI) IWR1443BOOST radar sensor that streams the raw Analog-to-Digital Converter (ADC) data in real-time via Ethernet using a TI DCA1000EVM module. The radar measurements are supplemented for the 2D SLAM method by an Xsens MTi-G-710 inertial measurement unit (IMU). For better coverage and more accurate angle estimation, the

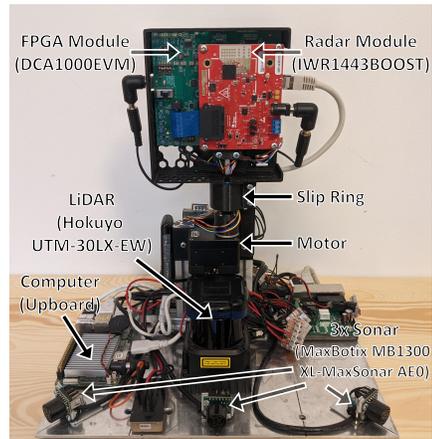


Fig. 2: Detailed view of the rotating radar module components.

radar is continuously rotated at 1 Hz by a Robotis Dynamixel XM430-W210 servo, gigabit Ethernet, and supply voltage are transmitted via a slip ring. For a direct qualitative comparison of radar and lidar SLAM methods and future multi-modal sensor data fusion, the module is equipped with a Hokuyo UTM-30LX-EW 2D lidar.

B. Radar processing

The quality of the radar data processing in terms of accuracy and number of detections is crucial for accurate and robust SLAM. An high-level overview of our pipeline is provided in Figure 3.

The streaming of the raw data enables the use of more complex methods and algorithms when compared to the more limited on-chip processing, and makes fine-tuning of the processing chain for SLAM possible. Also, different algorithms can be compared with one another in a reproducible manner. We base our implementation on the open-source framework OpenRadar² which implements the basic MIMO processing algorithms. We extend the framework by implementing a better performing CFAR algorithm and adding several post-processing methods, as well as optimize the processing chain to stream and process the raw data in real-time with more than 60 Hz.

1) *Low-level processing*: To calculate a detection threshold in range direction, the Cell Averaging Statistic Hofele (CASH)-CFAR [13] method is applied. Considering the current velocity, the number of computations can be greatly reduced by only calculating the thresholds for all points within a certain velocity around the current platform speed. After applying a training-based phase compensation for the receive (RX) channels and removing Doppler-induced phase shift, the angle is estimated for every detection. A total of two time-multiplexed transmit (TX) in combination with four RX antennas are used, resulting in an array of eight virtual antennas, with an angular resolution of 15°.

¹https://github.com/tu-darmstadt-ros-pkg/cartographer/tree/tsdf_radar-noetic

²<https://github.com/PreSenseRadar/OpenRadar>

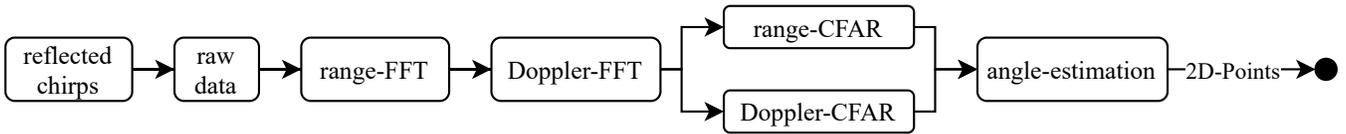


Fig. 3: Overview of the raw radar data processing pipeline: The reflected radar chirps are received, mixed with the transmitted signal, and streamed as raw data to the host. Next, two separate FFTs are applied to get a distance and velocity profile of the environment. Peaks in these signals correspond to true detections, which are filtered by combining two CFAR thresholds in range and Doppler direction. Finally, for every filtered detection the angle is estimated, resulting in 2D spatial points, including velocity, ready for post-processing.

Two Angle of Arrival (AoA) methods were investigated, Minimum Variance Distortionless Response (MVDR) beamforming (also known as Capon beamformer [14]) and a FFT-based method. Both assume a planar wavefront in their unmodified formulation, which breaks for near-field detections, reducing accuracy. For the latter method, however, the phase error between the virtual antennas arrays was minimized for detections in near-field by taking the antenna geometry into account. With that, the FFT-based angle estimation achieved overall higher accuracy at about five times the speed when compared to the Capon beamformer. The accuracy was measured with lidar as a reference and computing the average distance of all radar detections to the nearest lidar scan points while moving through different environments.

2) *High-level processing*: Since radar penetrates certain objects, this causes multiple consecutive detections in the range direction. These linear features are removed in the first post-processing step by only keeping the most intense point within a cluster of multiple range detections. Clusters are determined by defining a maximum thickness of objects, which equals to a maximum distance in the range directions for objects in the same cluster.

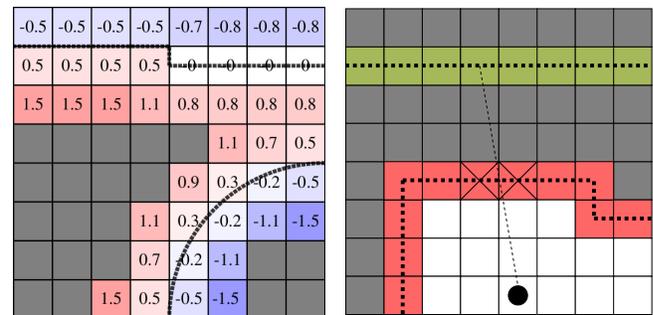
To increase the density of the radar scans, multiple scans are accumulated and then filtered for outliers. Assuming a low speed of the platform (about 1 m per second), the relative motion of two consecutive scans is small in relation to the sensor resolution and therefore negligible and not considered. After accumulating n scans, the detections are clustered by applying a Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [15] on the polar coordinate representation of the points in 2D-space. All clusters with a size less than a set number m are considered as outlier candidates. Assuming that the outliers are caused by multipath reflections or noise, they will have a relatively low intensity, when compared to other points in the scan. Therefore all outlier candidates that have intensities below the threshold $I_{min} = \mu - \lambda_n \sigma$ will be discarded, with $\lambda_n \in \{1, 2, 3\}$ being a scaling factor, σ and μ being the standard deviation and mean of all intensities in the accumulated scans.

C. TSDF-based SLAM for object-penetrating sensors with noisy data

To create a map and localize within it, we build upon the TSDF-based lidar SLAM approach by Daun et. al [5] which is based on Cartographer [6], an open-source SLAM system

that implements scan-to-map matching and loop closure detection. The TSDF grid produces larger gradients around the surfaces when compared to probability grids, resulting in a more robust and accurate scan matching for lidar-based SLAM [5]. Therefore, we extend the Cartographer TSDF approach with a novel forward sensor model to enable building TSDF maps from radar data and introduce robust scan matching algorithms to handle sparse, noisy, and outlier-rich radar range measurements.

1) *Sensor model for map update*: Instead of assigning each grid cell a probability of being occupied, the TSDF grid representation assigns each cell the distance to the nearest surface, as displayed for two spatial dimensions in Figure 4a.



(a) TSDF grid visualization. [5] (b) Map update for multiple detections in range.

Fig. 4: TSDF grid and proposed update method for two different environments. Figure 4a: Numbers indicate the distance to surfaces (dotted lines). In Figure 4b multiple detections, as viewed from the robot (black dot) are displayed. Space up to the first hit (red cells) is assumed empty (white cells), second hits are colored green with unknown space colored grey.

To update the TSDF from range data, a common method is the Projective Distance Update [16] where all cells along a ray from the sensor to the observations are updated with truncated signed distance to the observation.

For lidar-based scans, it is feasible to update the cells between a hit, and the sensor, by assuming these cells are empty. This method however is not suitable for radar, since it penetrates various objects and has a certain Field Of View (FOV) in elevation, resulting in either multiple detections in range or detection of larger, elevated objects behind smaller ones. Empty space between every hit and the sensor origin can therefore not be guaranteed. Nonetheless

updating possible empty cells can improve the SLAM result by removing outliers, that otherwise would accumulate in the map.

Therefore only space up to the first detection for a fixed angle is assumed free in the following approach, outlined in Figure 4b: First, every hit is converted to the corresponding pixel in the grid and saved with the corresponding pixel polar representation as a pair in an array. Next, for every hit, a ray from the detection to the sensor origin is also converted to a pixel mask list. If one of the other hits is present in this pixel mask, no free space update is performed. For faster lookup, only the other hits within a certain angular window around the current hit are selected using a fast binary search and then compared to the pixel ray. For every other hit occurring behind a previously hit cell, only the TSDF is updated.

The approach in comparison to existing update methods is displayed in Figure 5. With this proposed method, a free space update for both lidar and radar scans is possible, without differentiating the source of the scans.

2) *Robust optimization for scan registration:* To align a new scan into the current submap, the 2D-pose $\xi = [x, y, \theta]$ is directly optimized on the TSDF by a non-linear least squares problem:

$$\arg \min_{\xi} \sum_{i=1}^N (\Phi_I(\mathbf{T}_{\xi} \mathbf{h}_i))^2, \quad (1)$$

where Φ_I is the bi-linear interpolation of the TSDF grid map, \mathbf{T}_{ξ} a 2D-transformation (translation and rotation) for the current pose and \mathbf{h}_i a 2D-scan point, as proposed in [17]. Outliers can greatly worsen the quality of this non-linear least-squares problem, especially when dealing with radar scans a sub-optimal solution is often found. Therefore, we perform robust scan matching by applying Graduated Non-Convexity (GNC) as proposed by Yang et al. [18] and extend the Ceres Solver for outlier rejection. However, the optimization requires an initial guess for the pose. Odometry is used for an initial pose, refined with correlative scan matching, and then optimized using GNC with Ceres solver, using the Geman McClure (GM) as a robust cost function. Instead of directly defining a robust cost function ρ , GNC introduces a control parameter μ that alters the cost function $\rho_{\mu}(\cdot)$, so that the original shape of the function is recovered for the limit of μ (typically $\mu \rightarrow 1$ or infinity), otherwise, the function will be convex. Starting with a certain convexity, a first solution is then computed and μ is gradually changed to increase the non-convexity until the original shape is recovered. To extend the solver, the optimization process is split up into two steps: a variable update is performed by Ceres, followed by a weight update after each Ceres iteration. Using the Black-Rangarajan duality as described in [18], the optimization problem stated in Equation 1 with a robust cost function can be written as:

$$\min_{\xi \in \mathbb{R}^3, w_i \in [0,1]} \sum_{i=1}^N [w_i r^2(\mathbf{h}_i, \xi) + \Phi_{\rho}(w_i)], \quad (2)$$

with a weight w_i for every residual $r(\mathbf{h}_i, \xi)$ and a penalty function $\Phi_{\rho}(w_i)$ depending on the robust cost function $\rho(\cdot)$, which is independent of ξ .

For the variable update every iteration t , this can be written as

$$\xi^{(t)} = \arg \min_{\xi \in \mathbb{R}^3} \sum_{i=1}^N [w_i^{(t-1)} r^2(\mathbf{h}_i, \xi)], \quad (3)$$

which is solved using the Levenberg-Marquardt method with automatic differentiation. The weight update for every iteration t for the GM-function is then solved in closed-form:

$$w_i^{(t)} = \left(\frac{\mu_t \bar{c}^2}{\mu_t \bar{c}^2 + r(\mathbf{h}_i, \xi^{(t)})} \right)^2, \quad (4)$$

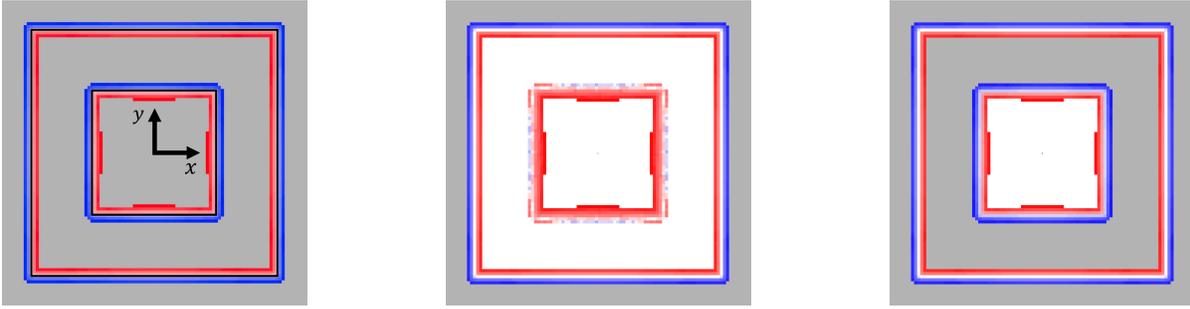
where the given parameter \bar{c} determines the shape of the GM function. Finally, after every iteration t , the convexity of the GM function is decreased by updating $\mu_t \leftarrow \mu_{t-1}/k_{\mu}$ with a given parameter $k_{\mu} > 1$.

IV. EVALUATION

To evaluate the robustness of the system in degraded visual conditions and in different environments, our detailed evaluation is presented in the following section. We benchmark our approach with artificially created data in a controlled environment to demonstrate the robustness of GNC-based TSDF registration method. Afterward, we investigate the performance of the presented method in both good visual conditions and under smoke with poor visibility. Finally, the results are compared to another publicly available dataset by commonly used metrics. All processing regarding the evaluation was performed on a laptop with an Intel i7-9750H CPU at 2.6 GHz and 16 GB RAM, without GPU acceleration.

A. Robust scan matching on artificial data

To assess the effectiveness of the modified robust scan matching in presence of outliers, we use simulated pointclouds as seen in Figure 5. For every iteration, the TSDF grid is filled with training data consisting of $n_{inlier} = 100$ scan points, simulating an empty, rectangular room. The data is subject to random zero-mean Gaussian noise, with a standard deviation of $\sigma = 0.01$ and a sensor resolution of 0.025 m. To simulate the outliers present in radar scans, randomly generated points $n_{outlier}$ twice the side length of the rectangular room are added to the training set, resulting in the outlier rate $o = n_{outlier}/n_{inlier}$. A matching pointcloud test set is distorted with the same number of outliers, transformed by applying a random rotation and translation and matched against the grid using both the unmodified TSDF cost function and the extended one by GNC. The translation and rotation errors for the test set were empirically set to 0.1 m and 0.1 rad for each pointcloud since correlative scan matching, which is applied beforehand, usually provided a first estimate from the reference in that range. The results per Ceres iteration for 100 matched pointclouds for both the normal TSDF and robust scan matching with 33 % outliers are shown in Figure 6. All simulations were solved using



(a) No free space update, update only TSDF. (b) Update free space up to every detection. (c) Proposed method: update only free space up to first detection.

Fig. 5: Simulated free space update methods. Black lines represent detected surfaces, cells are colored as follows; white: free space, red-blue gradient: TSDF, Grey: unknown cells. In Figure 5a, no cells except the hits will be updated. When performing a free space update for every detection, the gradient of first detections is getting overwritten, as seen in Figure 5b. For the proposed approach (Figure 5c), the TSDF gradient for multiple detections in range will not get overwritten.

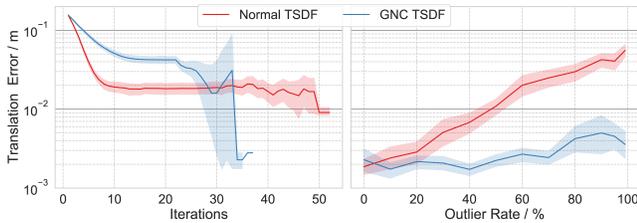


Fig. 6: Mean translation error for GNC scan matching (Blue) compared to unmodified approach (Red) for 100 matched pointclouds. The left plot shows the translation error per iteration for 33 % outlier, the right plot shows the translation error for increasing outliers. Shaded region displays the errors within two std. deviations.

the GM function with a shape of $\bar{c} = 7$ and $k_\mu = 1.2$. For both methods, the same convergence criteria were used. After approx. 10 iterations, the unmodified approach converges and does not find a significantly better solution. However, the proposed optimization finds a significantly better solution after approx. 25 iterations, once changing the convexity of the cost function sufficiently for outlier rejection.

The same test was performed with an increasing outlier ratio from 0 to 100 % with a step size of 10 %, again with 100 simulated pointclouds per outlier step. The results are displayed in Figure 6.

With the proposed method, both the error for translation and rotation were on average 3 and 2 times smaller, while the additional time and the number of iterations needed for solving the problem did only increase by 42 % and 14 % respectively. For outlier rates greater than 40 %, which usually occur for dense radar scans, the number of iterations for GNC scan matching was always lower. With the higher accuracy in presence of outliers as well as the relatively low additional computational effort due to smaller radar pointcloud sizes, the proposed GNC scan matching proved to be better performing for radar scans, which for our sensor contained outlier rates in the range of 80 % to 90 %.

B. Evaluation approach radar-SLAM

For evaluating the quality of the radar-based SLAM, lidar is used as a reference because an accurate ground truth system was not available for all different test scenarios. The quality of the SLAM is measured based on the created trajectories and not on the created maps since the environment perceived by radar varies greatly in comparison to the environment captured by lidar and is heavily dependent on the radar sensor used, e.g. the number of available antennas. Two metrics, the mean absolute errors (MAEs) of both the Relative Pose Errors (RPEs) and Absolute Trajectory Errors (ATEs) are calculated to evaluate both the local respectively global consistency of the SLAM. To estimate the RPE, for each pose \mathbf{x}_t created by the radar SLAM at time t , one reference pose \mathbf{x}_t^* nearest in time t' is considered. By only taking one nearest reference pose into account, local mismatches are weighted more. The relative errors for the translation and rotation are evaluated separately. By defining the relative transformation $\delta_{i,j} = \mathbf{x}_j \ominus \mathbf{x}_i$ between two poses \mathbf{x}_i and \mathbf{x}_j , the RPEs in translation and rotation can be written as [19]:

$$d_i(\delta) = d_i = \left\| \mathbf{trans} \left(\delta_{t,t+1} \ominus \delta_{t', (t+1)'}^* \right) \right\|_2 \quad (5)$$

$$r_i(\delta) = r_i = \mathbf{rot}(\delta_{t,t+1} \ominus \delta_{t', (t+1)'}^*). \quad (6)$$

The ATE is determined by calculating the displacement for the last radar pose to the last lidar SLAM pose. Since the pose update rate for lidar-based SLAM with 4 Hz usually is faster than for radar-based SLAM with around 1 Hz, we interpolate two nearest lidar poses in time assuming linear motion, before calculating the RPEs.

All data sets were recorded on a compact tracked rescue robot called Asterix, which was developed and built at the "Simulation, System Optimization and Robotics" group of Technical University of Darmstadt (TUDA).

C. TUDA Lab with lidar reference

For a first evaluation regarding the accuracy of pure radar-based SLAM, a total of 4 test runs were performed

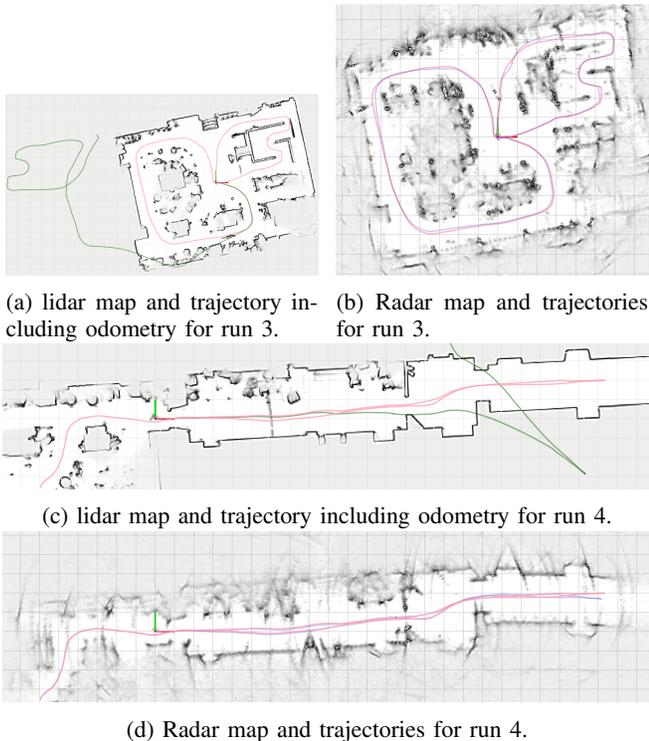


Fig. 7: Maps created by lidar- and radar-SLAM including trajectories: Pink: lidar trajectory, Green: odometry, Blue: radar trajectory. One square corresponds to 1 m².

under normal visual conditions without smoke. The first experiments took place in an approx. 12 by 15 m cluttered laboratory with almost no straight, empty walls present and people walking by. One half of the room consists of a work area with tables and chairs, on the other half a small arena with uneven terrain is present, inducing motion in pitch and roll. Additionally, the last data set was captured while driving through a long, small corridor.

In Figure 7, both maps and trajectories of the reference lidar and radar SLAM for the last two runs are shown. In run 3, the robot starts in the middle of the map, drives down to the right, one lap around the working area back to the starting position, and another lap through the arena. For the last run, the robot starts on the left side of the map, drives down a small ramp, and approx. 22 m to the right, turns on the spot, returns to the starting position, and drives into the laboratory, back to the starting position of run 3. For this run, a small angular drift was present when visiting a separate section of the corridor. This area was previously unobserved due to the small angle of incident of radar detections. Therefore, when only detecting two straight walls opposite of each other, without constraining the scan matching by observing other, more distinct features, the scan matching becomes more error-prone in rotation due to noise and inaccurate odometry, which in turn is subject to the same angular drift.

The average scores for all 4 runs are listed in Table II. Even for long runs, the ATE for the finale pose could be

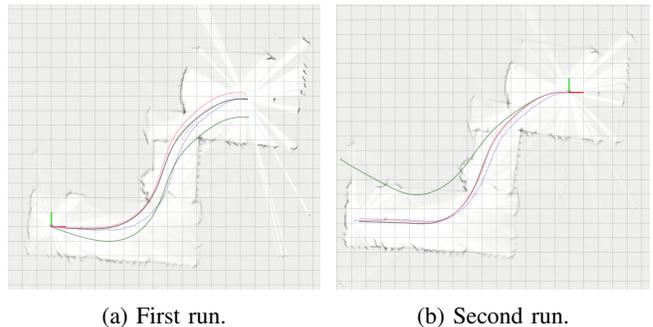


Fig. 8: Map and trajectories created by the proposed system from the radar perspective for the first run provided by [3]: Black: Ground Truth, Green: Odometry, Blue: Radar trajectory, Pink: lidar trajectory. One square corresponds to 1 m².

kept small and stayed under 7 cm and 2.75 deg. Both the radar and lidar maps contain small details of the environment, preserving walls and even corners. Also, the observed area along the radar trajectory is correctly marked as free space. Only indirectly or briefly observed areas are marked as unknown, as can be seen in the left section of Figure 7b.

D. Comparison on Mielle data-set

To the authors' best knowledge, no standardized dataset containing comparable radar data was available. To assess the performance of the proposed approach, the data provided by Mielle et al. [3] was used for comparison, since they offer comparable data with a detailed evaluation utilizing the same metrics. Similar to this work, Mielle et al. also collect lidar measurements, odometry data, and measurements from a fast rotating, but custom-built radar sensor (called Mechanically Pivotal Radar (MPR)) with a robotic platform to perform SLAM. Additionally, they also provide accurate localization by a positioning system used as a ground-truth reference. In their evaluation, they use two different SLAM frameworks to perform radar SLAM, gmapping [20], and normal distribution transform occupancy maps (NDT-OM) fuser [21], to calculate the RPE and the ATE of the last poses in comparison to ground truth.

To estimate the accuracy of our proposed system, the radar and odometry data provided by Mielle et al. were used to perform SLAM using the extended Cartographer framework. The resulting trajectory is then compared to the ground-truth reference and the error scores are calculated as described above. Unfortunately, the third run in the data set was incomplete, so only the first two runs were considered.

The created maps including trajectories are displayed in Figure 8, the results for radar SLAM in comparison to the averaged scores provided by Mielle are listed in Table I. With the proposed method, the scores for the RPE were lowered approx. by a third for translation and halved for rotation with a significantly smaller standard deviation. Especially the global consistency could be improved, resulting in an ATE approx. 6 and 3 times smaller when compared to NDT-OM respectively gmapping.

TABLE I: Average scores over run 1 and 2 with MPR, in comparison to scores provided by [3]. Best results are written in bold.

Metric	related to ground truth				related to lidar trajectory	
	proposed method	NDT-OM [3]		gmapping [3]		proposed method
\bar{d} / m	0.018 \pm 0.008	0.027	\pm 0.030	0.028	\pm 0.026	0.061 \pm 0.058
\bar{r} / mrad	33.02 \pm 18.13	75.50	\pm 105.5	68.50	\pm 145.0	36.04 \pm 21.38
D / m	0.289	1.962		1.097		0.363
R / mrad	77.81 (4.458 deg)	456.3	(26.16 deg)	215.5	(12.35 deg)	69.10 (3.959 deg)

Since an accurate ground truth system was not available in the previous evaluations and to get error scores for this data set that are comparable to our evaluation approach, the same evaluation method as previously was used by creating both the lidar and radar trajectories with our proposed approach and use lidar as a reference, instead of the positioning system. Since the scores of the lidar trajectory created by our approach compared to the provided positioning system were all lower than the scores by Mielle et al., it is valid to compare the created radar to the lidar trajectory instead of ground truth. The scores are also listed in Table I and are in the same order of magnitude as the results of the previous runs, but with a higher ATE in translation.

E. Deutsches Rettungsrobotik Zentrum (DRZ) Challenge with smoke

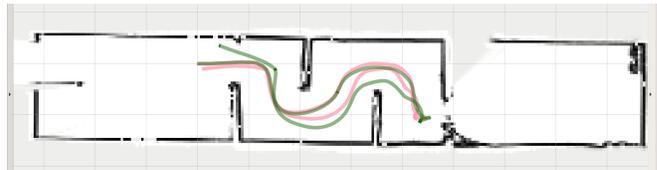
To validate the system in degraded visual conditions, a total of 8 runs, with and without smoke were recorded. The recorded environment can be completely filled with smoke and consists of a small S-shaped corridor with a total size of 4 m by 2 m formed by thin plasterboard walls. These walls are easily penetrated by the radar, resulting in multiple detections in range direction. The smoke is created by evaporating distilled water containing polyethylene glycol, using a conventional disco fog machine.

The error scores for the first 3 runs without smoke and using lidar as reference are listed in Table II. For the last runs, the area was completely filled with dense smoke, resulting in zero visibility rendering lidar-based SLAM approaches ineffective.

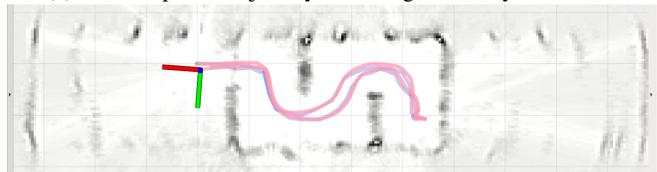
The radar detections however were less or not at all effected by smoke, as can be seen when comparing the radar map in Figure 9b to the radar map created under smoke, displayed in Figure 9c. For the latter, the lidar pointcloud for one frame, as viewed from the smaller axes, is also pictured, showing almost only false detections near the robot caused by smoke. The radar pointcloud also shows some outliers, which accumulate in unobserved areas, due to not updating the free space there.

However, the longer those areas are observed, the confidence of each cell being occupied or free increases. This effect can be seen when comparing Figure 9b and Figure 9c. The velocity of the platform was lowered in run 6 from approx. 1 m/s to 0.5 m/s due to the smoke, resulting in longer observations, making the lower map less fuzzy and the walls darker shaded.

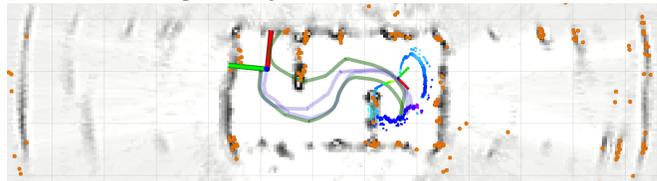
The small angular resolution of about 15° and the FOV in elevation affects the map in the left section of Figure 9b and



(a) lidar map and trajectory including odometry for run 3.



(b) Radar map and trajectories for run 3, without smoke.



(c) Radar map and trajectories for run 6, with smoke and overlapped pointclouds.

Fig. 9: Maps and trajectories created by lidar and radar SLAM. Green: Odometry, Blue: Radar trajectory, Pink: lidar trajectory. One square corresponds to 1 m². Figure 9c shows the pointclouds for lidar and radar for one scan, Rainbow: lidar pointcloud, Orange: radar pointcloud.

Figure 9c. On the lidar map, an open door can be seen, which is not recognizable on the radar map, most likely because the top of the door frame is detected from further away and the whole area is only observed for a short period of time.

Due to optimized odometry and a simpler environment compared to the first runs, all scores listed in Table II except the absolute rotation error were lowered. The RPE stayed under 3.3 cm and 1.2 deg, the ATE D under 3.3 cm. The increased final ATE for rotation R is caused by ambiguous scan matching at the end for the second run (as seen in Figure 9b), which resulted in a high absolute error of 17 deg. Additionally, the average ATE for translation along the trajectory, based on corresponding timestamps, was also calculated and was as low as 3.8 cm.

For the last runs, no lidar reference data could be created due to smoke. However, since the radar was less or not at all effected by smoke, similar quality and accuracy in terms of the error scores can be expected for the last runs. This shows the suitability for using the proposed radar-based SLAM method in situations with poor visibility.

TABLE II: Averaged scores over all runs for TUDA Lab and DRZ challenge, with lidar as reference.

\bar{d} and \bar{r} : MAE of the RPE for translation respectively rotation with averaged std. dev., D and R : Absolute translation respectively rotation error for final pose.

Metric	TUDA Lab (4 runs)	DRZ challenge (3 runs)
\bar{d} / m	0.0563 ± 0.0483	0.0323 ± 0.0301
\bar{r} / mrad	26.534 ± 25.06	17.433 ± 20.173
D / m	0.0348	0.0199
R / mrad	21.186 (1.635 deg)	137.75 (7.892 deg)

V. CONCLUSION AND FUTURE WORK

In this paper, we presented and evaluated specific methods to perform accurate real-time radar SLAM using TSDF maps in degraded visual conditions, where conventional optical sensors fail to operate.

Special radar processing methods were investigated and fine-tuned for both accuracy and number of detections. By taking the radar sensor model into account, the advanced SLAM framework Cartographer was extended by radar specific map update rules, as well as a robust scan matching approach in the presence of outliers, both leading to a more detailed map and higher accuracy for pose estimation.

The extensive evaluation presented demonstrates the robustness of the system in varying environments, as well as under zero visibility while creating a rich map and accomplishing localization accuracy with an average RPE of 4.6 cm and 1.3 deg comparable to lidar SLAM.

Evaluation on other data sets showed that the proposed approach can also be used for different radar sensors while achieving higher accuracy when compared to other SLAM frameworks.

Future work already in progress is aiming to replace the rotating radar assembly with two statically mounted sensors, to maintain data density while reducing mechanical components and space requirements. Finally, the accuracy of the new system may be further evaluated by a more accurate ground truth system.

REFERENCES

- [1] T. Shan et al. “LIO-SAM: Tightly-coupled Lidar Inertial Odometry via Smoothing and Mapping”. In: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 5135–5142.
- [2] P. Geneva et al. “OpenVINS: A Research Platform for Visual-Inertial Estimation”. In: *Proc. of the IEEE International Conference on Robotics and Automation*. Paris, France, 2020.
- [3] M. Mielle et al. “A comparative analysis of radar and lidar sensing for localization and mapping”. In: *2019 European Conference on Mobile Robots (ECMR)*. 2019, pp. 1–6.
- [4] M. Holder et al. “Real-Time Pose Graph SLAM based on Radar”. In: *2019 IEEE Intelligent Vehicles Symposium (IV)*. 2019, pp. 1145–1151.
- [5] K. Daun et al. “Large Scale 2D Laser SLAM using Truncated Signed Distance Functions”. In: *2019 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. 2019, pp. 222–228.
- [6] W. Hess et al. “Real-Time Loop Closure in 2D LIDAR SLAM”. In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*. 2016, pp. 1271–1278.
- [7] J.-H. Kim et al. “Firefighting Robot Stereo Infrared Vision and Radar Sensor Fusion for Imaging through Smoke”. In: *Fire Technology* 51.4 (June 2015), pp. 823–845.
- [8] M. O’Toole et al. “Homogeneous codes for energy-efficient illumination and imaging”. In: *ACM Transactions on Graphics (ToG)* 34.4 (2015), pp. 1–13.
- [9] N. Mandischer et al. “Radar SLAM for Autonomous Indoor Grinding”. In: *2020 IEEE Radar Conference (RadarConf20)*. IEEE, 2020, pp. 1–6.
- [10] C. X. Lu et al. *See Through Smoke: Robust Indoor Mapping with Low-cost mmWave Radar*. 2020.
- [11] A. Kramer et al. *ColoRadar: The Direct 3D Millimeter Wave Radar Dataset*. 2021.
- [12] T. D. B. Keenan Burnett Angela P. Schoellig. “Do We Need to Compensate for Motion Distortion and Doppler Effects in Spinning Radar Navigation?” In: *IEEE Robotics and Automation Letters* 6.2 (2021), pp. 771–778.
- [13] F. Hofele. “An innovative CFAR algorithm”. In: *2001 CIE International Conference on Radar Proceedings (Cat No.01TH8559)*. 2001, pp. 329–333.
- [14] J. Capon. “High-resolution frequency-wavenumber spectrum analysis”. In: *Proceedings of the IEEE* 57.8 (1969), pp. 1408–1418.
- [15] M. Ester et al. “A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise”. In: *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. Portland, Oregon: AAAI Press, 1996, pp. 226–231.
- [16] R. A. Newcombe et al. “KinectFusion: Real-time dense surface mapping and tracking”. In: *Mixed and augmented reality (ISMAR), 2011 10th IEEE international symposium on*. IEEE, 2011, pp. 127–136.
- [17] E. Bylow et al. “Real-time camera tracking and 3D reconstruction using signed distance functions.” In: *Robotics: Science and Systems*. Vol. 2. 2013.
- [18] H. Yang et al. “Graduated Non-Convexity for Robust Spatial Perception: From Non-Minimal Solvers to Global Outlier Rejection”. In: *IEEE Robotics and Automation Letters* 5.2 (2020), pp. 1127–1134.
- [19] R. Kümmerle et al. “On measuring the accuracy of SLAM algorithms”. In: *Autonomous Robots* 27 (2009), pp. 387–407.
- [20] G. Grisetti et al. “Improved Techniques for Grid Mapping With Rao-Blackwellized Particle Filters”. In: *IEEE Transactions on Robotics* 23.1 (2007), pp. 34–46.
- [21] T. Stoyanov et al. “Normal Distributions Transform Occupancy Map fusion: Simultaneous mapping and tracking in large scale dynamic environments”. In: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. 2013, pp. 4702–4708.