Indoor Localization based on Short-Range Radar and Rotating Landmarks

Kolja Thormann, Simon Steuernagel, and Marcus Baum
Institute of Computer Science
University of Göttingen
Göttingen, Germany
Email: {kolja.thormann, simon.steuernagel, marcus.baum}@cs.uni-goettingen.de

***Abstract*—A novel concept for indoor self-localization based on rotating artificial landmarks with known locations using short-range radar is proposed. First, a processing pipeline for extracting range and angle measurements to the landmarks from a raw radar image is introduced, which consists of a neural network for distance estimation and a basic angle-of-arrival estimator. Second, a particle filter for tracking the pose based on the range and angle measurements is developed. Due to the ability of radar to measure range rate, i.e., the velocity in the direction of a detection, it is possible to robustly detect and localize rotating landmarks with the help of their micro-Doppler pattern. In this way, localization is possible even under difficult conditions (e.g., light changes). Experiments with a wheeled mobile robot and common office fans as landmarks demonstrate the effectiveness of the approach for indoor localization.**

***Index Terms*—Indoor localization, inside-out tracking, short-range radar, micro-Doppler**

I. Introduction

Self-localization, i.e., the sensor-based determination of the own position with respect to a global reference system is an essential task for many mobile technical systems. In particular for indoor environments, localization [1]–[6] is challenging as a high position accuracy is typically desired and Global Navigation Satellite Systems (GNSSs) such as the Global Positioning System (GPS) are not available.

A common setup for indoor localization involves environment sensors, e.g., cameras, lidar, or radar sensors, that are mounted on the mobile device (also called inside-out approach). Based on the environment sensors, either features of the surroundings can be matched with a global map [7] or relative measurements to landmarks [8]–[10] can be obtained for localization. A wide range of inside-out approaches based on cameras are available [11], in particular in many Virtual Reality (VR) systems [12]. However, the use of cameras is often a risk for the user's privacy and there is a high sensitivity to lighting conditions and occlusions. An alternative approach is conducted by SteamVR [12], which uses base stations that actively send out infrared signals. A further approach is to use special tags or beacons for radio signals [13] such as Radio Frequency Identification (RFID) tags [14], which improve the radar cross section.

Lidars [15] can provide high resolution point clouds, while radars [16], [17] on the other hand provide intensity images and can also detect range rate. Different possibilities for landmarks exist, e.g., radar corner reflectors [10] or special transponders [18] are available. A property of Frequency Modulated Continuous Wave (FMCW) radars is their ability to detect range, range rate (i.e., velocity in range direction), and angle-of-arrival. Ultra-wideband pulse radars can be faster (better update rate and lower processing time [19]), while an FMCW radar generally has better signal-to-noise ratio due to the higher power of signals, e.g., [20]. As the range resolution of FMCW radar depends on the bandwidth, ultra-wideband technology can be applied for these types of radar, see, e.g., [21].

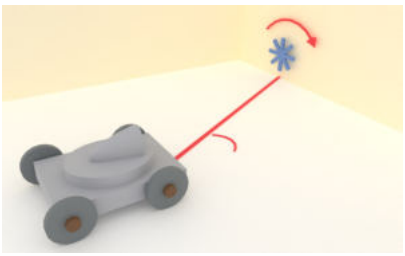


Fig. 1: The mobile robot extracts the distance and angle to the rotating landmark with its radar device by exploiting the micro-Doppler pattern of the landmark.

A. Contribution

In this work, we propose a new type of artificial landmark for indoor localization using FMCW radar, see Fig. 1, The key idea is to employ rotating parts in the landmarks that cause a micro-Doppler pattern [22], [23] in the radar image. By this means, the detection of landmarks is significantly improved, and relative position measurements can be robustly extracted.

In order to demonstrate the practicability of the idea, we present a landmark detection and self-localization pipeline for a mobile robot equipped with an FMCW radar.

A key element is a novel approach for detecting distances to micro-Doppler sources in radar images using a Convolutional Neural Network (CNN). Furthermore, the processing pipeline consists of a standard bootstrap particle filter, see, e.g., [24], [25], for pose tracking.

B. Related Work

To the best of our knowledge, there are no works using radar to detect micro-Doppler signatures of rotating artificial

landmarks to determine the distance and angle towards that landmark for self-localization.

Landmark-based approaches for localization using radar can be found in [9], [10]. Static objects are extracted as peaks in the intensity image and used as landmarks by reducing noise and multipath reflections in [9]. Only range measurements are considered for this landmark identification. The multiple points are then matched. Specific corner reflectors and the distance measurements to them are utilized in [10]. Works such as, e.g., [26], model landmarks using tags which register the radar signal and return it and can possibly encode an identifier. A different approach using distance measurements to multiple transponders, or tags, providing unique IDs is presented in [18]. Tags which return a frequency comb not suffering from multipath interference for drone localization are introduced in [27].

Simple, passive reflectors are used in [28] for fingerprinting and multilateration using a radar. Special tags are also used in an automotive context [29], but only for presence detection of vulnerable road users, not range estimation. Other materials like photonic crystals are evaluated in [30]. The Doppler shift to a spinning beacon is utilized in [31] to find the angle-of-arrival towards the beacon. In [32], a rotating RFID tag is used to increase detectability. In a similar way, a vibrating tag is used in [33]. Another work determines whether a rotating marker is present in the radar's field of view to be used as a warning signal [34], but none of these utilize the micro-Doppler pattern created by a rotating landmark.

There are works using the property of radar to detect velocity to identify the micro-Doppler signatures of drone rotors, e.g., [35]–[39]. CNNs are also applied for the classification of drones or their motion states based on micro-Doppler signatures in, e.g., [40], [41]. For detection and classification, a deep CNN is applied in [42] and a long short-term memory neural network for detection, classification, and localization is utilized in [43], in which they handle the localization by determining the angle for the received micro-Doppler pattern over time. The work [44] suggests a model which processes radar data using convolutional layers in a variational autoencoder to learn features to be used for different applications like estimating the distance to a corner reflector. In the context of 5G, the work [45] proposes a novel approach for indoor simultaneous localization and mapping via known base stations in combination with a new estimation method to determine the propagation path of the signal. Transmitter antennas are at the base station and receivers at the target to be localized.

C. Structure

The remainder of this work is structured as follows. The problem this paper intends to tackle is described in Section II. In Section III, the detection of rotating landmarks is presented, followed by a description of the positioning algorithm based on those landmark detections in Section IV. We discuss our approach in Section V. The paper is concluded in Section VI.

II. Problem Setting

The general problem that this work is concerned with is localization, i.e., the determination of the absolute location and orientation of a mobile robot. To this end, the pose of the robot at time k, defined as

$$\mathbf{x}_k = \begin{bmatrix} m_{k,1} & m_{k,2} & \theta_k \end{bmatrix}^{\mathrm{T}} \ , \tag{1}$$

is to be determined, based on knowledge of the global position of a set of landmarks. Here $m_{k,1}$ and $m_{k,2}$ represent the Cartesian position of the robot, and θ_k its orientation.

A set of N global landmark locations in Cartesian coordinates $\mathbf{p}^j = \begin{bmatrix} p_1^j & p_2^j \end{bmatrix}^{\mathrm{T}} \in \mathbf{P}$ with $j \in \{0, \dots, N-1\}$ is known. Note that the time index for $\mathbf{p}^j$ is omitted, since landmarks are stationary.

At each time step, by processing the radar data as discussed in Section III, we receive M observations $\mathbf{y}_k^i = \begin{bmatrix} r_k^i & \alpha_k^i \end{bmatrix}^{\mathrm{T}} \in \mathbf{Y}_k$ with $i \in \{0, \dots, M-1\}$, where r_k^i and α_k^i denote the range and bearing to the i-th landmark. For a known association (i, j), the measurement model is

$$\mathbf{y}_k^i = \begin{bmatrix} \sqrt{(p_1^j - m_{k,1})^2 + (p_2^j - m_{k,2})^2} \\ \operatorname{atan2}(p_2^j - m_{k,2}, p_1^j - m_{k,1}) - \theta_k \end{bmatrix} + \mathbf{w}_k^i \ , \tag{2}$$

with measurement noise $\mathbf{w}_k^i \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$. The objective is to determine the posterior probability distribution $p(\mathbf{x}_k | \mathbf{Y}_{1:k})$ of the state $\mathbf{x}_k$, where we can get an estimate via the posterior mean $\hat{\mathbf{x}}_k = \mathrm{E}[\mathbf{x}_k | \mathbf{Y}_{1:k}]$.

The overall localization system is illustrated in Fig. 2. It can be divided into two individual subproblems:

- The first is the extraction of the distance and angle to each landmark, given a relative estimate of the robot pose, see Section III.
- The second is the sequential (Bayesian) estimation of the overall pose given these range/angle measurements, see Section IV, which consists of alternating update and prediction steps.

We utilize a Multiple Input Multiple Output (MIMO) FMCW radar, namely the BGT60ATR24C $60\,\mathrm{GHz}$ radar from Infineon [46]. It works by sending out a chirp, i.e., a signal with increasing frequency up to a bandwidth of $4\,\mathrm{GHz}$, and receiving its reflections. It produces an intensity image for discrete range and range rates. The radar has two transmitter and four receiver antennas, arranged in a way that using virtual antennas, it is able to produce a horizontal array of up to four antennas. It can also produce up to three vertical antennas, but we only consider 2D detections for this work. The phase shift of the signal for a given frequency/range can be used to determine the detected object's angle-of-arrival.

The FMCW radar was set up to have a range resolution of $0.0187\,\mathrm{m}$ with 128 range cells, resulting in a maximum range of $2.3983\,\mathrm{m}$, and a range rate resolution of $0.0518\,\mathrm{m\,s^{-1}}$ with 64 range rate cells, resulting in a maximum range rate of $\pm 1.6567\,\mathrm{m\,s^{-1}}$. We only used one transmitting antenna, so the number of virtual antennas and thus the number of channels was four, with only the first three of them belonging to antennas arranged in a row for angle estimation.

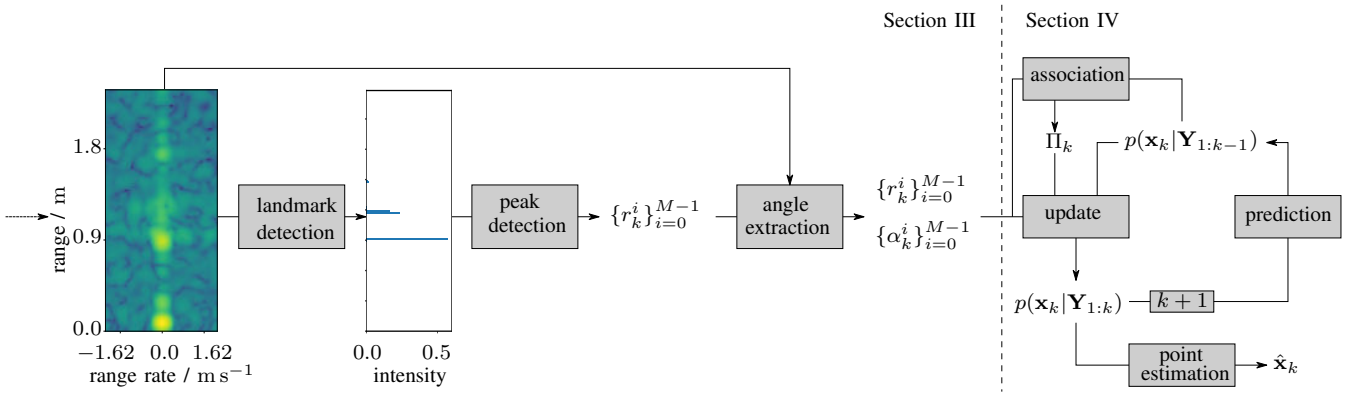


Fig. 2: Overview diagram of the proposed processing chain at time step k.

An example radar image from a single channel can be found on the left in Fig. 2. The closer landmark is clearly visible as a static object, but its micro-Doppler pattern as well as the second landmark are hard to see for the human eye, while the CNN, which will be introduced in the next section, is able to learn finer details.

For the experimental setup, the previously described FMCW radar was mounted to a mobile robot. We used a Turtlebot 3 [47]. Based on internal sensors and wheel odometry, the robot provides information about its own location and orientation, which was extracted for the ground truth. Furthermore, we employed small office fans (approximately $18\,\mathrm{cm}$ diameter) as the landmarks for the experiments. They were placed in front of the robot at a suitable distance facing the radar, and the angle and distance between robot and landmarks were determined for the initial robot location. Using the available robot odometry, the relative landmark position can be determined dynamically from this initial global landmark position.

III. Landmark Detection

The detection of landmarks consists of two steps. First, range estimates of possible landmarks are estimated using a CNN, described in Section III-A. As the second step, the FMCW radars multiple antennas are utilized to calculate an angle estimate for each range estimate. This is explained in Section III-B. The landmark detection is then evaluated in Section III-C.

A. Range estimation via CNN

To differentiate the landmarks from the static environment, we utilize the micro-Doppler pattern created from the rotation of the landmark. The radar provides a measurement $\mathbf{m}$ in the shape of an $h \times w \times a$ intensity image, with $a = 4$ the number of channels, i.e., antennas, w the range rate resolution, and h the range resolution. An example for a single channel can be found on the left of Fig. 2. To detect and localize the micro-Doppler pattern of a landmark, we utilize a CNN. Before $\mathbf{m}$ is passed to the CNN, it is pre-processed according to

$$\tilde{\mathbf{m}} = \frac{\log(|\mathbf{m}|) - \min(\log(|\mathbf{m}|))}{\max\Big(\log(|\mathbf{m}|) - \min(\log(|\mathbf{m}|))\Big)}\,, \tag{3}$$

i.e., the measurement is first converted to log-space, and then normalized to the interval $[0, 1]$ by subtracting the minimum and then dividing by the maximum of the resulting transformed measurement. Note also that we use the absolute values $|\cdot|$ of the radar image, which is usually complex. We ignore the phase information here because we only require the intensity to determine the micro-Doppler pattern. The phase information is used later to determine the angle of arrival in Section III-B.

The CNN gets the processed radar image $\tilde{\mathbf{m}}$, which can be interpreted as a normalized $h \times w \times a$ intensity image. The image is processed by 8 convolutional layers with Rectified Linear Unit (ReLU) functions in between. An important aspect is the 4th layer which has a kernel width equal to w, which corresponds to the range rate dimension, i.e., the range rate is flattened to consider the entire Doppler spectrum for each range cell at once. While every layer uses padding to ensure the dimensions do not change, this layer uses no padding in the range rate dimension. This network structure results in a single array of length h as output, i.e., an array with dimension equal to the range resolution of the radar, which serves as a heatmap providing high values for those range bins which are predicted to contain a landmark. The detailed structure of each convolution layer in sequential processing order is given in Table I.

For each image, the corresponding ground truth provides the true ranges of existing landmarks. During training, a heatmap is created for each input frame which has a value of 1 for each range bin containing a landmark and 0 otherwise. At the same time, a sigmoid function limits the output values of the CNN between 0 and 1. For training, we then utilize binary cross entropy loss.

Training of the CNN was conducted using a manually recorded dataset of 604 images. For the dataset collection, we used two landmarks and manually measured their distance and angle to the sensor mounted on the robot. We then

TABLE I: The number of input and output channels, kernel size, and padding used for the convolution layers of the CNN used in this work. *Same* padding means that the padding is chosen in a way that the image dimensions are preserved. There is a ReLU activation function in between each layer.

no. of input channels	no. of output channels	kernel size	padding
4	8	(3, 3)	same
8	32	(3, 3)	same
32	64	(3, 3)	same
64	128	(13, 64)	(6, 0)
128	128	(7, 1)	same
128	64	(5, 1)	same
64	32	(3, 1)	same
32	1	(3, 1)	same

collected data while moving the robot. Our setup used the robot odometry to determine the change in robot pose and update the relative landmark locations. We then saved the radar images along with the true relative landmark locations, resulting in a dataset consisting of 604 images used for training the CNN. Before conducting the training process, the images are shuffled. In the same manner, we collected 452 images to use as the validation set. To increase the data set size for training, two types of data augmentation were employed.

In addition to the original data, ten different random erasing augmentations were used, in which a rectangle covering 2%-15% of the original image was erased, with an aspect ratio of 3 to 15, i.e., vertical blocks were erased in order to ensure that the micro-Doppler pattern was not entirely removed. This first augmentation increases the data set by a factor of 11. Afterwards, the resulting data set was additionally flipped horizontally, which is possible as high intensity stationary objects are at a range rate of 0 and the micro-Doppler patterns are centered around 0, meaning flipping the data still produces valid images. In total, the composition of these two transforms yields a data set of size 13288 for training.

As the optimizer for training, AdamW [48] was used. The learning rate was 10^{-4}, and weight decay was set to 0.01. Training was conducted for merely 3 epochs, after which satisfactory results were achieved.

An example of the intensity heatmap for a given radar image can be found in Fig. 2. As the output of the CNN is a heatmap, we utilize the peak finding algorithm `find_peaks` from the scipy signal library [49] to get specific range estimates for landmarks.

B. Angle extraction via local averaging

Due to the FMCW radar using multiple antennas, a detection coming from an angle will result in a phase shift. Therefore, for each range and range rate cell of the radar image, we can use the phase shift between the corresponding antennas along with the intensity of the received signals to calculate weights for different possible angles. This is also called digital beamforming [50]. We consider 181 virtual beams in between $\pm 45°$. We then use the cells' intensity to create the norm of the weights for each range-angle pair over the range rate dimension. As we know that movement stems from the rotor blades of the landmark, but static objects can be anywhere in the environment, we decided to ignore the range rate cells in a band of $\pm 20\%$ of the maximum range rate around 0 range rate. As a visual example, we can say that we remove the area of the central intensity column in the image on the left of Fig. 2. Finally, we normalize those final weights to calculate a weighted average of the angles for each range cell.

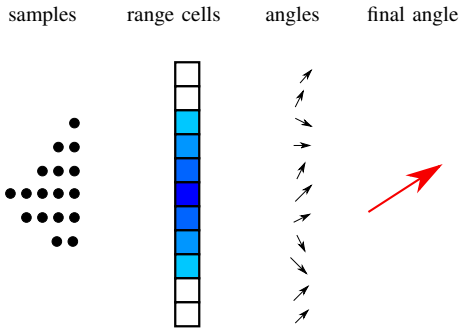


Fig. 3: Schematic example of the angle estimation process. The colored squares are a simplified representation of the range cells and their intensity represents a Gaussian distribution with the mean at the range estimate. On the right, the angle estimates calculated from the different channels of the corresponding range cell are visualized as arrows. An example of randomly drawn particles from the Gaussian is displayed on the left, each contributing with their corresponding angle to the overall angle estimate in red.

We described the range estimation in Section III-A. However, the range estimate is uncertain. Due to the discrete nature, even small deviations of the range can lead to large errors in the angle estimate if the wrong range cell is chosen. To handle this, we also consider the range cells in the vicinity of the estimate. Using a Gaussian distribution with fixed variance and the range estimate as mean, we generate some, e.g., 100 range values. Due to the nature of the Gaussian distribution, more range values will be drawn closer to the original estimate. For each range value, the corresponding range cell is determined and its angle estimate is calculated. The final angle estimate is then taken as the mean of the individual estimates. An appropriate uncertainty for the Gaussian is determined experimentally. We found a standard deviation of $5\,\mathrm{cm}$, or approximately three range cells, to yield satisfactory results. The process is sketched in Fig. 3.

C. Experimental Evaluation of the Detections

For the evaluation, we use the CNN to calculate a range estimate for each frame of the validation set. We apply the Hungarian assignment algorithm, implemented in the scipy [49] library, to associate the estimates to the landmarks based on their ranges, removing all associations with an error higher than $30\,\mathrm{cm}$. For each such estimate, we use the method described above to estimate the angle. This results in a set of range-angle estimates and associated landmarks per frame. Then, the mean and median of the range and angle errors are calculated. Furthermore, we additionally consider missed landmarks and false estimates to determine the precision and recall over the entire validation set. The precision is the

TABLE II: The mean and median of the range and angle error (low is good) as well as precision and recall (high is good) of the CNN (ours) and a baseline method detecting the peaks of the sums of range rows (sum), both using the same angle estimator on the validation set.

Error measure	Ours	Baseline
Range mean(median) error in cm	3.67(1.2)	7.17(5.33)
Angle mean(median) error in degree	9.53(8.94)	10.14(10.30)
Precision in $[0, 1]$	0.87	0.61
Recall in $[0, 1]$	0.73	0.54

percentage of produced detections matching a true landmark, whereas the recall is the percentage of true landmarks which were detected.

For our method, we consider only peaks with a minimum height of 0.01 in the heatmap. As a naive baseline, we utilize that the micro-Doppler pattern of a landmark results in higher intensities at the corresponding cells, so we create an alternative heatmap via the sum of the range rate cells per range cell. As this heatmap has much higher values and much more peaks, we consider a minimum peak height of 30 and a minimum peak width of 5 cells. During our evaluation, this resulted in the best trade-off between precision and recall. Results of this evaluation are presented in Table II.

Given a maximum detection range of approximately $2.4\,\mathrm{m}$, the range error is small in relation to it, with a median error of merely $1.2\,\mathrm{cm}$ or 0.5% of the maximum range. The error is also reduced to approximately half of that of the baseline. Furthermore, it is important to note that due to the spatial extent of the landmark, a detection on the outer parts of the fan will introduce an angular error, even though the landmark is correctly detected. The precision demonstrates low false detections relative to the overall detections, while the recall shows that the CNN more often misses a landmark instead of detecting false ones. In both categories, the CNN improves the naive approach. The occasional missed and false detections can be handled by fusing measurements over time in a filtering algorithm, which is described in the following section.

IV. Robot Localization

Due to the non-linearity of the estimation problem, we employ a bootstrap Particle Filter (PF) [25]. The transition density is assumed to be

$$p(\mathbf{x}_k|\mathbf{x}_{k-1}) \approx \mathcal{N}(\mathbf{x}_k; \mathbf{x}_{k-1}, \Delta_t \cdot \mathbf{Q}) \ , \tag{4}$$

with Δ_t being the time difference between the current and last scan in seconds and process noise covariance matrix $\mathbf{Q}$.

The overall distribution of the state is represented by a weighted set of n_p particles. Given the particles $\mathbf{x}_{k-1}^l$ and weights w_{k-1}^l for time $k-1$, and the newly received measurements $\mathbf{Y}_k$ from time k, the updated particles $\mathbf{x}_k^l$ and weights w_k^l $(l = 0, \ldots, n_\mathrm{p} - 1)$ are obtained by

1) sampling the particle state via the state transition density

$$\mathbf{x}_k^l \sim \mathcal{N}(\mathbf{x}_k^l; \mathbf{x}_{k-1}^l, \Delta_t \cdot \mathbf{Q}) \ , \tag{5}$$

2) computing the new particle weights as

$$w_k^l \propto p(\mathbf{Y}_k|\mathbf{x}_k^l, \mathbf{P}) \ , \tag{6}$$

3) normalizing the particle weights to sum to 1,
4) performing resampling, setting the weights to be uniform.

A. Particle Likelihood

This leaves only the particle likelihood $p(\mathbf{Y}_k|\mathbf{x}_k^l, \mathbf{P})$ to be determined. Due to the unknown associations between detections $\mathbf{Y}_k$ and landmark positions $\mathbf{P}$, an approximation needs to be found. To this end, we compute a matrix $\mathbf{L}_k \in \mathbb{R}^{M \times N}$, where an entry $\mathbf{L}_k^{i,j}$ represents the log-likelihood[1] of observing the detection $\mathbf{y}_k^i$ for the true landmark $\mathbf{p}^j$, given the current particle state $\mathbf{x}_k^l$. For brevity, we omit the particle index l when describing $\mathbf{L}_k^{i,j}$. An individual entry $\mathbf{L}_k^{i,j}$ is computed as

$$\mathbf{L}_k^{i,j} = \log\left(\mathcal{N}\left(\mathbf{y}_k^i;\ \hat{\mathbf{y}}_k^{p^j}, \mathbf{R}\right)\right) \ , \tag{7}$$

where $\mathbf{R}$ is the (polar) measurement noise covariance matrix and $\hat{\mathbf{y}}_k^{p^j}$ is the expected observation of landmark p^j. Using the particle state $\mathbf{x}_k^l = \begin{bmatrix} m_1^l & m_2^l & \theta^l \end{bmatrix}^\mathrm{T}$, it is computed as

$$\hat{\mathbf{y}}_k^{p^j} = \mathrm{cartToPol}\left(\mathrm{rot}(-\theta^l) \cdot (\mathbf{p}^j - \begin{bmatrix} m_1^l & m_2^l \end{bmatrix}^\mathrm{T})\right) \ , \tag{8}$$

using $\mathrm{cartToPol}(\cdot)$ to convert a vector from a Cartesian to a polar representation, and $\mathrm{rot}(\cdot)$ to construct a rotation matrix.

Furthermore, we apply gating as a standard approach to handle clutter, by setting all values in $\mathbf{L}_k$ corresponding to an association between a detection and landmark with (Cartesian) distance above a threshold γ to $\log(10^{-100})$. From the matrix $\mathbf{L}_k$, the estimated associations $\Pi_k \in (\{0, \ldots, M-1\} \times \{0, \ldots, N-1\})$ can be found by solving the linear assignment problem for the negative log-likelihood matrix $-\mathbf{L}_k$. To this end, we again employed the Hungarian algorithm. Given Π_k, the overall likelihood is then approximated as

$$p(\mathbf{Y}_k|\mathbf{x}_k^l, \mathbf{P}) \approx \exp\left(\sum_{(i,j)\in\Pi_k} \mathbf{L}_k^{i,j}\right) \ , \tag{9}$$

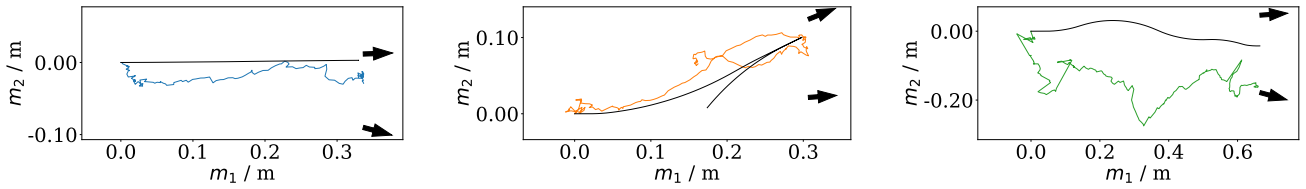
i.e, the overall likelihood is computed as the product of the individual likelihood of associated detection-landmark pairs. Using (9), the unnormalized particle weight in (6) can be directly computed.

B. Implementation Details

For the practical implementation of the filter, it is important to note that the orientation as part of the state space is bounded in $[0, 2\pi)$ and periodic. Therefore, care must be taken when computing an overall consolidated estimate given the individual particle states. One way to deal with this issue is to compute the average in Cartesian space by converting each θ^l to $\begin{bmatrix} \cos(\theta^l) & \sin(\theta^l) \end{bmatrix}^\mathrm{T}$, and then computing the angle of the result to the x-axis.

Furthermore, it is necessary to choose the free parameters in the previously defined PF algorithm. We tuned all parameters empirically for the evaluated setting of indoor

[1]The log-space representation is merely used for computational stability of the numerical calculation.



(a) First trajectory. (b) Second trajectory. (c) Third trajectory.

Fig. 4: Bird's eye view of the three trajectories. Black arrows point towards landmarks, placed at a distance between $0.85\,\mathrm{m}$ and $1.53\,\mathrm{m}$ from the initial robot position. The black line corresponds to the true robot trajectory, which always starts at $\mathbf{x}_0 = \begin{bmatrix}0\,\mathrm{m} & 0\,\mathrm{m} & 0\,\mathrm{rad}\end{bmatrix}^\mathrm{T}$, whereas the colored lines represent the estimate.

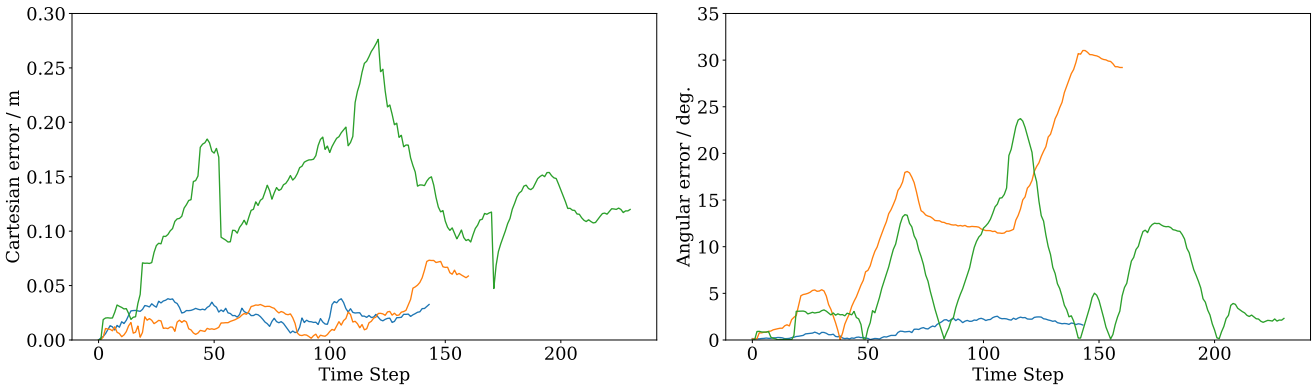


Fig. 5: Absolute translational and angular errors throughout the trajectories over time. The colors correspond to the colors of the individual trajectories in Fig. 4. Note that due to the different lengths of the trajectories, the number of time steps varies between them.

localization. The gating distance was set to $\gamma = \frac{1}{3}\mathrm{m}$. The process noise covariance used for importance sampling was set to $\mathbf{Q} = \mathrm{diag}(0.05\,\mathrm{m}, 0.05\,\mathrm{m}, \frac{\pi}{64}\,\mathrm{rad})^2$ and the measurement noise covariance to $\mathbf{R} = \mathrm{diag}(0.04\,\mathrm{m}, 0.175\,\mathrm{rad})^2$. The values for the measurement noise covariance were chosen in accordance with the results obtained in Section III. The number of particles was $n_\mathrm{p} = 1000$. Since the initial position was by definition known to be at $\begin{bmatrix}0\,\mathrm{m} & 0\,\mathrm{m} & 0\,\mathrm{m}\end{bmatrix}^\mathrm{T}$, all particles were initialized at that pose. Note that before the first weight update, an importance sampling step will be carried out.

C. Experimental Analysis

For the evaluation, the presented filtering system was employed together with the CNN-based detector. Three trajectories were recorded in the same manner as described in Section II and Section III-C using two fans as rotating landmarks facing the robot. Results of the localization are shown in Fig. 4. In all cases, the robot started on the left facing the landmarks and moves forward to the right (and in the second trajectory, it moves backwards at the end) while keeping at least one landmark in its field of view. The corresponding average Cartesian position and angular errors are given in Fig. 5. In the first trajectory, the robot moved along a straight line towards the two landmarks. On this trajectory, the localization is highly accurate, with an average error of $2.29\,\mathrm{cm}$ and no noteworthy outliers. The second trajectory yielded similarly good results. However, towards the end of the trajectory, the orientation estimation begins to drift off, which in turn causes a slight increase in error to at most $0.1\,\mathrm{m}$. This is also visible in the trajectory estimation in Fig. 4b, where towards the end, the robot is moving away from the landmarks in a slightly incorrect angle. Nevertheless, the overall trajectory is closely followed, and the mean Cartesian error is still $2.31\,\mathrm{cm}$. On the third trajectory, however, the filter exhibits problematic behaviour earlier in the sequences. An in-depth analysis of the processing chain revealed that throughout these time steps, only a single landmark is successfully detected. Hence, the overall problem is underdetermined, and the robot location starts drifting off. Still, the general direction of the robot is followed, i.e., the filter does not diverge. As more landmarks are visible towards the end of the trajectory, the pose estimate begins to correct. In total, the mean Cartesian position error in this trajectory is $13.03\,\mathrm{cm}$. Overall, the estimation of the angular component of the robot pose appears to be more challenging for the system than the Cartesian localization.

V. Discussion

In Sections III and IV, results of both the landmark detection as well as the entire localization results were presented. Overall, the pipeline was found to yield accurate results, in particular regarding ranging and positioning, with a mean range detection of less than $4\,\mathrm{cm}$. Throughout the relative

localization of a mobile robot, positioning errors of below $5\,\mathrm{cm}$ were observed, which did, however, increase in individual situations. This can be contributed to several factors.

Landmark visibility and hence detection recall were found to strongly impact errors, as visible in Fig. 4c and 5. Furthermore, the angular estimation is more challenging than the range detection. While the larger error of the angle in comparison to the distance shown in Table II can partially be attributed to the physical extent of the landmark, the results in Section IV have shown that incorrect angle estimates can cause drift-like behaviour of the positioning system. Also, due to our angle extraction framework, multiple landmarks occupying the same range cell can not be distinguished. Handling such cases would be possible with an adapted framework.

The current filtering approximates the measurement likelihood as a Gaussian. However, the difference between mean and median detection quality indicates that this model assumption is inaccurate. For the ranging estimate, the mean is three times as large as the median, indicating that the error distribution is not Gaussian, but instead skewed by outliers. These outliers are likely caused by misdetections of the CNN, as also visible in the precision of the network. Neither the precision nor the recall of the network is explicitly modeled in the current exemplary filter, potentially hindering performance. A more elaborate filtering system taking these error sources into account could, hence, prove noticeably more robust against outliers, improving performance in the challenging situations causing the current filter to drift from the true trajectory. With better knowledge of the robot motion model, an improved sampling procedure for the PF may drastically reduce the required number of particles.

VI. Conclusion

In this work, we introduced a novel concept for the localization of a mobile robot, specifically with application to indoor positioning. In particular, we presented rotating fans as a novel type of landmark for radar sensors. A prototypical processing pipeline was proposed, consisting of pre-processing of the radar, a CNN-based landmark detector, and finally a PF-based positioning system. In experiments using real-world data collected using an FMCW radar and small office fans, we found the system to yield centimetre-level accuracy. However, in individual situations, erroneous detections caused the positioning results to drift further from the true robot location. Conceptually, compared to a camera-based system, the radar-based processing chains means the system is not reliant on ambient lighting, and offers safety regarding privacy concerns. The landmarks employed here are merely required to be rotating and stationary, but do not need to actively engage in the localization process.

The presented experimental setup was designed using simple components with the aim to validate the potential of employing landmarks with rotating parts at all. Naturally, the positioning system can be improved. One possible direction for such future work can be an improved Bayesian filtering system, taking the object dynamics into account. Additionally, the precision and recall of the detector could be incorporated into the filter, avoiding issues in cases where not enough landmarks are detected to fully determine the robot location. Furthermore, exploring the influence of prior pose uncertainty on the filtering accuracy can be of interest. It is also worth investigating whether the phase information can be used directly in the CNN. Moreover, different technical setups, e.g., with more or different landmarks could be considered in order to further improve the accuracy of the localization system.

References

[1] T. Yang, A. Cabani, and H. Chafouk, "A survey of recent indoor localization scenarios and methodologies," *Sensors*, vol. 21, no. 23, Dec. 2021.

[2] F. Zafari, A. Gkelias, and K. K. Leung, "A survey of indoor localization systems and technologies," *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568–2599, Jul.-Sep. 2019.

[3] G. M. Mendoza-Silva, J. Torres-Sospedra, and J. Huerta, "A meta-review of indoor positioning systems," *Sensors*, vol. 19, no. 20, 2019.

[4] R. F. Brena, J.-P. García-Vázquez, C. E. Galván-Tejada, D. M. Rodríguez, C. V. Rosales, and J. Fangmeyer, "Evolution of indoor positioning technologies: A survey," *J. Sensors*, vol. 2017, pp. 2 630 413:1–2 630 413:21, 2017.

[5] A. Nessa, B. Adhikari, F. Hussain, and X. N. Fernando, "A survey of machine learning for indoor positioning," *IEEE Access*, vol. 8, pp. 214 945–214 965, 2020.

[6] T. Savić, X. Vilajosana, and T. Watteyne, "Constrained localization: A survey," *IEEE Access*, vol. 10, pp. 49 297–49 321, 2022.

[7] Z. Yang and Z. Zhu, "An ego-motion estimation method using millimeter-wave radar in 3D scene reconstruction," in *2022 14th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, 2022, pp. 18–21.

[8] B. Jang, H. Kim, and J. wook Kim, "Survey of landmark-based indoor positioning technologies," *Information Fusion*, vol. 89, pp. 166–188, 2023.

[9] S. H. Cen and P. Newman, "Precise ego-motion estimation with millimeter-wave radar under diverse and challenging conditions," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 2018, pp. 6045–6052.

[10] J. Michalczyk, C. Schöffmann, A. Fornasier, J. Steinbrener, and S. Weiss, "Radar-inertial state-estimation for UAV motion in highly agile manoeuvres," in *2022 International Conference on Unmanned Aircraft Systems (ICUAS)*, 2022, pp. 583–589.

[11] M. S. Alam, F. B. Mohamed, A. Selamat, and A. B. Hossain, "A review of recurrent neural network based camera localization for indoor environments," *IEEE Access*, vol. 11, pp. 43 985–44 009, 2023.

[12] V. Holzwarth, J. Gisler, C. Hirt, and A. Kunz, "Comparing the accuracy and precision of SteamVR tracking 2.0 and Oculus Quest 2 in a room scale setup," in *2021 the 5th International Conference on Virtual and Augmented Reality Simulations*, 2021, pp. 42–46.

[13] T. Kim Geok, K. Zar Aung, M. Sandar Aung, M. Thu Soe, A. Abdaziz, C. Pao Liew, F. Hossain, C. P. Tso, and W. H. Yong, "Review of indoor positioning: Radio wave technology," *Applied Sciences*, vol. 11, no. 1, 2021.

[14] M. El-Absi, A. A.-h. Abbas, A. Abuelhaija, K. Solbach, and T. Kaiser, "Chipless RFID infrastructure based self-localization: Testbed evaluation," *IEEE Transactions on Vehicular Technology*, vol. 69, no. 7, pp. 7751–7761, 2020.

[15] L. Xie, J. Hu, N. Du, M. Zhou, Y. Wang, and W. Nie, "Reflective markers assisted indoor LiDAR self-positioning algorithm based on joint optimization with FIM and GDOP," *IEEE Sensors Journal*, pp. 1–1, 2023.

[16] K. Harlow, H. Jang, T. D. Barfoot, A. Kim, and C. Heckman, "A new wave in robotics: Survey on recent mmwave radar applications in robotics," 2024.

[17] Z. Hong, Y. Petillot, and S. Wang, "RadarSLAM: Radar based large-scale SLAM in all weathers," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2020, pp. 5164–5170.

[18] S. Böller, T. Grenter, and A. Grabmaier, "FMCW based positioning using multiple SHF RFID transponders," in *2023 13th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2023, pp. 1–5.

[19] A. Figueroa, B. Al-Qudsi, N. Joram, and F. Ellinger, "Comparison of two-way ranging with FMCW and UWB radar systems," in *2016 13th Workshop on Positioning, Navigation and Communications (WPNC)*, 2016, pp. 1–6.

[20] D. Wang, S. Yoo, and S. H. Cho, "Experimental comparison of IR-UWB radar and FMCW radar for vital signs," *Sensors*, vol. 20, no. 22, 2020.

[21] B. Welp, G. Briese, and N. Pohl, "Ultra-wideband FMCW radar with over 40 GHz bandwidth below 60 GHz for high spatial resolution in SiGe BiCMOS," in *2020 IEEE/MTT-S International Microwave Symposium (IMS)*, 2020, pp. 1255–1258.

[22] F. Fioranelli, H. Griffiths, M. Ritchie, and A. Balleri, Eds., *Micro-Doppler Radar and its Applications*. The Institute of Engineering and Technology (IET), 2020.

[23] S. Peter and V. V. Reddy, "Extraction and analysis of micro-Doppler signature in FMCW radar," in *2021 IEEE Radar Conference (RadarConf21)*, 2021, pp. 1–6.

[24] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, pp. 174–188, 2002.

[25] S. Särkkä, *Bayesian filtering and smoothing*. Cambridge university press, 2013, vol. 17.

[26] R. Sorrentino, E. Sbarra, L. Urbani, S. Montori, R. V. Gatti, and L. Marcaccioli, "Accurate FMCW radar-based indoor localization system," in *2012 IEEE International Conference on RFID-Technologies and Applications (RFID-TA)*, 2012, pp. 362–368.

[27] H. M. Hussein, S. Kim, M. Rinaldi, A. Alù, and C. Cassella, "Passive frequency comb generation at radiofrequency for ranging applications," *Nature Communications*, vol. 15, no. 1, p. 2844, 2024.

[28] S. Hinderer, P. Schlachter, Z. Yu, X. Wu, and B. Yang, "Indoor positioning based on active radar sensing and passive reflectors: Reflector placement optimization," in *2023 13th International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, 2023, pp. 1–7.

[29] A. Lazaro, R. Villarino, M. Lazaro, and D. Girbau, "Spread spectrum modulated mm-wave tag for V2X applications," in *2023 IEEE 13th International Conference on RFID Technology and Applications (RFID-TA)*, 2023, pp. 181–184.

[30] J. Sánchez-Pastor, L. Piotrowsky, A. Jiménez-Sáez, M. Schüßler, N. Pohl, and R. Jakoby, "Evaluation of chipless RFID indoor landmarks at 80 GHz and 240 GHz using FMCW radars," in *2022 16th European Conference on Antennas and Propagation (EuCAP)*, 2022, pp. 1–4.

[31] H.-l. Chang, J.-b. Tian, T.-T. Lai, H.-H. Chu, and P. Huang, "Spinning beacons for precise indoor localization," in *Proceedings of the 6th ACM Conference on Embedded Network Sensor Systems*, ser. SenSys '08. New York, NY, USA: Association for Computing Machinery, 2008, p. 127–140.

[32] A. Azarfar, N. Barbot, and E. Perret, "Chipless RFID based on micro-Doppler effect," *IEEE Transactions on Microwave Theory and Techniques*, vol. 70, no. 1, pp. 766–778, 2022.

[33] ——, "Vibration sensing using Doppler-modulated chipless RFID tags," in *2022 IEEE/MTT-S International Microwave Symposium - IMS 2022*, 2022, pp. 129–132.

[34] J. Verdú, E. Guerrero, J. Parrón, A. Lázaro, and P. de Paco, "Star-shaped wheel for mechanical micro-Doppler modulation," *IEEE Antennas and Wireless Propagation Letters*, vol. 20, no. 12, pp. 2452–2456, Dec. 2021.

[35] S. Zulkifli and A. Balleri, "Design and development of K-band FMCW radar for nano-drone detection," in *2020 IEEE Radar Conference (RadarConf20)*, 2020, pp. 1–5.

[36] J. Gong, J. Yan, D. Li, R. Chen, F. Tian, and Z. Yan, "Theoretical and experimental analysis of radar micro-Doppler signature modulated by rotating blades of drones," *IEEE Antennas and Wireless Propagation Letters*, vol. 19, no. 10, pp. 1659–1663, Oct. 2020.

[37] K.-B. Kang, J.-H. Choi, B.-L. Cho, J.-S. Lee, and K.-T. Kim, "Analysis of micro-Doppler signatures of small UAVs based on Doppler spectrum," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 5, pp. 3252–3267, Oct. 2021.

[38] J. Gong, J. Yan, D. Li, and D. Kong, "Detection of micro-Doppler signals of drones using radar systems with different radar dwell times," *Drones*, vol. 6, no. 9, p. 262, Sep. 2022.

[39] J. Park, D.-H. Jung, K.-B. Bae, and S.-O. Park, "Range-Doppler map improvement in FMCW radar for small moving drone detection using the stationary point concentration technique," *IEEE Transactions on Microwave Theory and Techniques*, vol. 68, no. 5, pp. 1858–1871, May 2020.

[40] B. K. Kim, H.-S. Kang, and S.-O. Park, "Drone classification using convolutional neural networks with merged Doppler images," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 1, pp. 38–42, 2017.

[41] W. Wang, Z. Tang, X. Xiong, Y. Chen, Y. Zhang, Y. Sun, Z. Zhu, and C. Zhou, "Motion states classification of rotor target based on micro-Doppler features using CNN," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019, pp. 1390–1393.

[42] H. C. Kumawat, M. Chakraborty, and A. A. B. Raj, "DIAT-RadSATNet—a novel lightweight DCNN architecture for micro-Doppler-based small unmanned aerial vehicle (SUAV) targets' detection and classification," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–11, 2022.

[43] Y. Sun, S. Abeywickrama, L. Jayasinghe, C. Yuen, J. Chen, and M. Zhang, "Micro-Doppler signature-based detection, classification, and localization of small UAV with long short-term memory neural network," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 8, pp. 6285–6300, 2021.

[44] A. Safa, T. Verbelen, O. Catal, T. V. de Maele, M. Hartmann, B. Dhoedt, and A. Bourdoux, "FMCW radar sensing for indoor drones using learned representations," 2023, arXiv:2301.02451.

[45] E. Rastorgueva-Foi, O. Kaltiokallio, Y. Ge, M. Turunen, J. Talvitie, B. Tan, M. F. Keskin, H. Wymeersch, and M. Valkama, "Millimeter-wave radio SLAM: End-to-end processing methods and experimental validation," 2023, arXiv:2312.13741.

[46] Infineon, "BGT60ATR24C," https://www.infineon.com/cms/en/product/sensor/radar-sensors/radar-sensors-for-automotive/60ghz-radar/bgt60atr24c/_prediction_topic, (accessed May 5, 2023).

[47] ROBOTIS, "Turtlebot 3," http://web.archive.org/web/20240226111346/https://emanual.robotis.com/docs/en/platform/turtlebot3/overview/, (accessed Feb 26, 2024).

[48] I. Loshchilov and F. Hutter, "Decoupled weight decay regularization," 2017, arXiv:1711.05101.

[49] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, vol. 17, pp. 261–272, 2020.

[50] J. Gamba, *Radar signal processing for autonomous driving*. Springer, 2020, vol. 1456.