

Infradar-Localization: single-chip infrared- and radar-based Monte Carlo localization

Dong Wang¹, Marco Masannek¹, Stefan May² and Andreas Nüchter³

Abstract—This paper proposes a novel approach for indoor robot localization that leverages a fusion of information from single-chip infrared (Time-of-Flight) and radar sensors. The aim of our research is the development of a cost-effective and lightweight system that can achieve high-precision robot localization. Unlike traditional localization methods based on LiDARs or cameras, our proposed system uses single-chip infrared and radar sensors to overcome the limitations of high cost and bulky hardware. Specifically, we employ a Doppler radar-based velocity motion model for the estimation of the robot’s ego-motion, eliminating the need for additional sensors such as IMU or wheel encoders. Next, we describe a hybrid sensor model for single-chip infrared and radar sensors that provides robust and accurate environmental perception with dynamic outlier removal. Finally, we integrate these components into a Monte Carlo localization framework to generate accurate real-time estimation of the robot’s position and orientation. This is the first time a single-chip infrared and radar fusion-based framework has been applied to robot localization, to the best of our knowledge. Through a comprehensive experimental evaluation, we demonstrate the system’s high accuracy and efficiency, achieving an average localization error of 9 cm in diverse indoor environments. This remarkable performance, combined with the low-cost and lightweight nature of our proposed solution, positions it as a highly promising alternative for a wide range of applications, including robotics, smart homes, and autonomous vehicles. The significant advancements of this novel approach offer vast potential to revolutionize the field of localization, enabling more precise and cost-effective navigation systems.

I. INTRODUCTION

The development of autonomous robots has received considerable attention in recent years. The capacity to navigate in complex environments with accuracy and efficiency is an essential prerequisite for autonomous robots. Localization plays a critical role in achieving this goal by providing robots with information about their position and orientation relative to their surroundings. Over the years, the robotics community has witnessed a remarkable surge in the development of sensor-based localization techniques. A diverse range of

sensor-based indoor localization approaches has been proposed, including WiFi, ultra-wideband (UWB), light detection and ranging (LiDAR), cameras, inertial measurement unit (IMU), and wheel encoders, among others. These techniques have varying degrees of accuracy, cost, and complexity, making it essential to choose the most suitable method based on the specific application requirements. Traditional

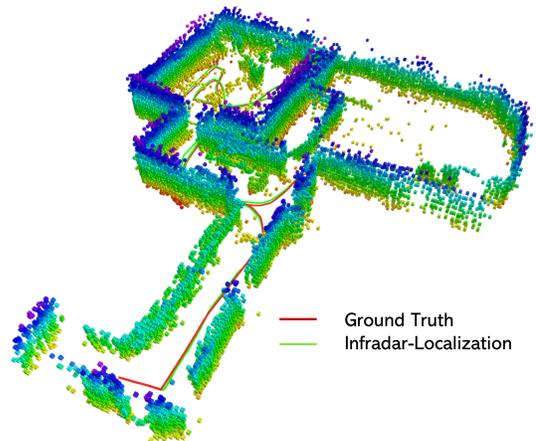


Fig. 1: Infradar-Localization in a warehouse environment with 3D occupancy map generated by low-cost Infradar sensors

localization approaches rely on external setups, which can constrain their resilience in various situations. For instance, UWB localization [1] requires the installation and calibration of numerous anchors in advance, while WiFi localization [2] depends on the availability of multiple routers to attain a high degree of accuracy. On the other hand, LiDAR- and camera-based localization offers a device/landmark-independent localization for mobile robots [3][4]. Monocular and stereo cameras are cost-effective passive sensors that can proficiently address localization issues by serving as a single source of information concerning an environment, specifically in Visual Odometry (VO) [5][6] and Visual Simultaneous Localization and Mapping (V-SLAM) [7][8]. While modern mobile robots have made great strides in their capabilities, it is undeniable that they still face limitations in terms of onboard computing power. Consequently, utilizing multiple cameras to achieve 360° coverage can present a daunting challenge for the robot’s onboard computer, as it must process an enormous amount of data in real-time to maintain optimal performance[9]. Furthermore, solutions that rely on cameras may not perform optimally when confronted with challenging lighting conditions (e.g. poor or extreme

¹Dong Wang and Marco Masannek are with the Faculty of Electrical Engineering, Precision Engineering, Information Technology, Nuremberg Institute of Technology Georg Simon Ohm, Nuremberg, 90489, Germany, and with the Department of Informatics XVII Robotics, Julius-Maximilians-University Würzburg, Würzburg, 97074, Germany dong.wang@th-nuernberg.de, marco.masannek@th-nuernberg.de

²Stefan May is with the Faculty of Electrical Engineering, Precision Engineering, Information Technology, Nuremberg Institute of Technology Georg Simon Ohm, Nuremberg, 90489, Germany stefan.may@th-nuernberg.de

³Andreas Nüchter is with the Department of Informatics XVII Robotics, Julius-Maximilians-University Würzburg, Würzburg, 97074, Germany andreas.nuechter@uni-wuerzburg.de

lighting) or when obstacles obstruct the field of view[10]. In recent years, researchers have also devoted considerable effort to developing and refining both 2D and 3D LiDAR-based indoor localization algorithms[11][12], resulting in a wealth of innovative solutions for accurately tracking the movement and position of robots in complex environments. However, these methods often have limitations, such as high cost, bulky hardware, lack of access to valuable semantic information, or sensitivity to dynamic environments.

To overcome these limitations, we propose a novel approach for indoor robot localization that combines single-chip infrared and radar (This technology will be referred to as Infradar in the following sections of this paper.) sensor fusion, Infradar ego-motion estimation, hybrid Infradar sensor model, dynamic outlier removal, and Monte Carlo localization. The key contributions of our work are as follows:

- Infradar-Localization is, to our knowledge, the first indoor robot localization framework that employs low-cost single-chip infrared and radar sensors.
- A single-chip Infradar-based ego-motion estimation method has been developed that eliminates the need for additional sensors.
- A hybrid sensor model for Infradar-Localization that incorporates dynamic outlier removal, further enhancing its reliability in detecting and mitigating errors.

This paper is structured as follows: Section II reviews the related work. Section III presents the proposed Infradar-Localization based on low-cost single-chip infrared and radar sensors. The experimental evaluation is described in Section IV. Section V summarises the study and discusses future directions.

II. RELATED WORK

With advancements in sensor technology, researchers are turning to low-cost options for robot localization, such as radar and infrared sensors. These sensors have many benefits, including their small size, low cost, robustness to environmental factors like lighting conditions, and low power consumption. In this section, we will conduct an in-depth exploration of the latest approaches to robot localization using single-chip radar and infrared sensors.

A. Ego-motion Estimation with Single-chip Radar Sensor

Recent improvements in electronics and materials science have enabled the radar (radio detection and ranging) sensor in a smaller package compared to earlier products. A notable innovation is the antenna-in-package concept, which integrates multiple radar antennas on a small chip, also known as single-chip radar. This technology provides sparse resolution for 3D detection and radial velocity measurement using Doppler effect [13][14]. The advantage of radar over VO is that Doppler radar can directly measure the relative velocity of stationary objects within a single frame, rather than having the relative velocity derived from the changing position of the stationary object in consecutive frames. Early research, such as the work presented in [15], proposed an

instantaneous approach for 2D radar ego velocity estimation using only one radar scan with Doppler radial velocity measurements. The authors utilized the RANdom SAmple Consensus (RANSAC) algorithm to filter out the moving objects in the environment and employed the Least-Squares estimator (LSQ) to optimize the radial velocity of stationary objects relative to the radar to obtain radar ego velocity. In [16], the approach was extended to multiple radar sensors with joint optimization of spatial. To achieve higher accuracy, the fusion of radar measurements with inertial data has shown impressive results [17][18][19]. However, without the help of yaw angle the yaw drift increases with time due to changes in the yaw rate of the MEMS IMUs. To compensate for this, the authors in [20][21] presented variants of Radar Odometry based on further sensor data fusion, such as Radar Visual Inertial Odometry, Radar Thermal Inertial Odometry, and GNSS-aided Radar Inertial Odometry.

B. Localization with Single-chip Radar Sensor

Radar sensors have been studied in disaster environments due to their ability to operate effectively in visually degraded conditions. However, their accuracy and density of data are typically reduced, making localization challenging. The authors in [22] proposed a method for radar measurement registration with a dense LiDAR map previously generated for localization. They employed a point registration algorithm that is independent of density for sparse radar data. Another approach to overcome the sparsity of radar data was proposed in [23], where a conditional Generative Adversarial Network (GAN) was used, trained with LiDAR data supervision. However, the heavy use of neural networks brings considerable challenges to real-time performance on mobile robots.

C. Localization with Single-chip Infrared Sensor

Due to the limited field of view and single distance measurement capability of most infrared sensors, there have been few studies on using them alone for robot localization. Previous work [24][25] has shown that although infrared sensors have excellent measurement accuracy, but a camera sensor is still required because of their low spatial resolution. In Q2 of 2021, ST-Microelectronics released 4th generation infrared sensor. This is the first miniaturized multi-zone ranging Time-of-Flight (ToF) sensor that features an 8×8 image resolution. The sensor presented in this work addresses the issue of low spatial resolution present in prior studies [24][25], thereby opening up new potential applications such as robot localization without the need for a vision-based algorithm. After the introduction of the novel infrared sensor, a number of targeted studies have shown great promise. A lightweight method is proposed in [26] to compute collision probability from each infrared sensor and directly on board to extract features of complex environments. Authors in [27] proposed a solution that is able to detect objects in 360° with minimal blind spots. This design was implemented on educational robots using multiple infrared sensors. In [28] accurate 3D pose data

of multiple people is estimated with infrared sensors after the supervised training. In both [26] and [29], an indoor localization and navigation framework was developed using infrared sensors and an external odometry module. However, the framework is specifically designed for an online nano-UAV processor and therefore is not applicable to mobile robots.

D. Summary

Our approach significantly advances existing methods in several respects:

- Based on the physical measurement characteristics of single-chip radar, we propose the multi-strategy weighting LSQ method for ego-motion estimation without the aid of other sensors such as IMU or camera.
- Our approach doesn't rely on a high-precision LiDAR map for building 3D maps. Instead, we use Infradar sensors to build 3D maps.
- With our approach, reliable global localization can be achieved in environments with dynamic obstacles.

III. METHODOLOGY

In this paper, we employ the following conventions to represent the various mathematical and physical quantities used in our research:

- Scalars will be printed as lowercase, non-bold letters (e.g. b), and constants will be printed as uppercase, non-bold letters (e.g. B).
- Matrices will be printed as bold upper case letters, like \mathbf{B} .
- Vectors will be represented by bold lowercase letters, like \mathbf{b} .
- Subscripts and superscripts are used to denote different frames of reference. For example, a vector \mathbf{b} in frame $\{\}^r$ will be denoted as \mathbf{b}^r , and the rotation from frame $\{\}^r$ to frame $\{\}^w$ will be represented by either the matrix \mathbf{B}_r^w or the quaternion \mathbf{b}_r^w .
- The global world frame is represented by $\{\}^w$.

By using this formalism, we aim to provide a clear and consistent notation that facilitates communication and understanding of our mathematical models and results. Fig. 2 shows the overview of the proposed system consisting of four components: single-chip Infradar, Infradar motion model, hybrid Infradar sensor model, and Monte Carlo localization.

A. Infradar Sensor Measurements

This section provides a brief summary of the Infradar sensors used in this work, together with their main characteristics. Since the Infradar sensor is a combination of a single-chip infrared sensor and a single-chip radar sensor, the two sensors will be introduced separately in the following sections.

1) *Single-chip Radar Measurements*: The single-chip millimeter wave (mmWave) radar uses frequency-modulated continuous wave (FMCW) to transmit (TX) signals that get reflected by objects in its path. By detecting the reflected signal, the range, velocity, and angle of the objects can be determined. When the radar front-end receives the signal (RX) reflected from an obstacle, an onboard signal mixer combines the received and transmitted signals to generate an intermediate frequency (IF) signal. The distance d from object to radar can be derived as:

$$d = \frac{f_{IF}C}{2s} \quad (1)$$

where C denotes the speed of light $3 \times 10^8 m/s$, f_{IF} represents the frequency of the IF signal, and s is the frequency slope of the chirp. Fast Fourier Transform (FFT) is applied to the IF signal when there are multiple obstacles present at different distances, resulting in peaks that correspond to one or more obstacles at specific distances. To estimate the angle of obstacles, a linear receiver antenna array is employed in mmWave radar. The process involves transmitting chirps with identical initial phases and simultaneously sampling signals from multiple receiver antennas. By analyzing the phase differences of received signals, the Angle of Arrival (AoA) α of the reflected signal can be mathematically calculated as [30]:

$$\alpha = \sin^{-1}\left(\frac{\lambda \cdot \Delta\phi}{2\pi d}\right) \quad (2)$$

where $\Delta\phi$ is the phase change of the FFT peak, d represents the distance between consecutive antennas and λ is the wavelength. It should be noted that $\Delta\phi$ depends on $\sin(\alpha)$, which exhibits a non-linear dependency. The approximation of $\sin(\alpha)$ as a linear function is only valid when α is small in magnitude:

$$\sin(\alpha) \sim \alpha \quad (3)$$

Therefore, the accuracy of the estimation depends on the value of AoA, and better accuracy can be achieved when α is small. To measure velocity, an FMCW radar emits two chirps separated by a time interval of t_c . The phase difference obtained by processing each reflected chirp through the range-FFT is used to calculate the velocity of the object according to the following equation:

$$v = \frac{\lambda \cdot \Delta\varphi}{4\pi t_c} \quad (4)$$

where $\Delta\varphi$ denotes the phase difference in range-FFT and λ represents the wavelength. However, if several moving objects at varying speeds are at the same distance from the radar at the time of measurement, the two-chirp velocity method will not work, because they generate reflective chirps with identical IF frequencies. In this case, the radar system must transmit a set of n equally separated chirps. These chirps are then processed by range-FFT to produce a set of n peaks at the same location. However, each peak has a unique phase that incorporates the phase input of both objects. Therefore, a second FFT, known as Doppler-FFT, is conducted on the n phasors.

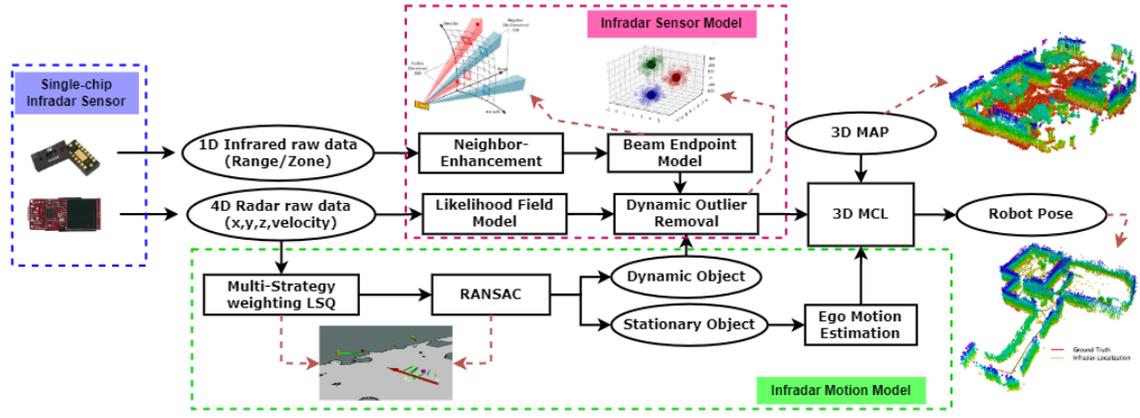


Fig. 2: Pipeline of Infradar-Localization

2) *Single-chip Infrared Measurements*: The single-chip infrared sensor, presented in this paper, is the 4th generation infrared sensor from STMicroelectronics. It is developed as a compact and lightweight optical multi-zone time-of-flight sensor for indoor sensing purposes. One of the key features of the infrared sensor is its multi-zone capability, which allows for up to 8×8 zones with a broad 65° diagonal field of view (FoV). Distance estimation is achieved by counting the number of photons returned from each zone in each time range and fitting the data with histogram algorithms to detect peaks (see Fig. 3). In the case of ToF interference at 940 nm wavelength from other optical sources, an error flag is returned to filter out noise and errors. In order to convert the distance information of each zone into 3D coordinates in the sensor coordinate system $\{ \}^s$, we apply the following transformation based on the index and distance of each zone:

$$x = d \cdot \frac{\sin\beta}{\cos\beta} \quad y = d \cdot \cos(\alpha) \quad z = d \cdot \sin(\alpha) \quad (5)$$

where the distance of the zone is represented by d , while α and β represent the angles around the x-axis and y-axis respectively, as shown in Fig. 3, which are dependent on the zone index.

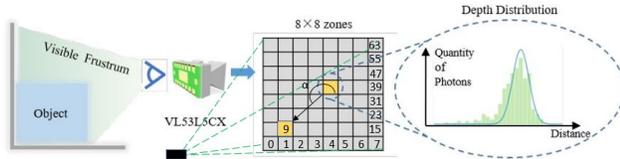


Fig. 3: Infrared sensor sensing principle and 8×8 zone mapping

B. Infradar Ego Motion Estimation

As mentioned in the previous chapter III-A.1, a single-chip mmWave radar measurement consists of a set of targets, their three-dimensional (3D) positions \mathbf{p}^r , corresponding Doppler radial velocities v_d^r and the signal-to-noise ratio (SNR) s of each target. The velocity of radar at this moment is \mathbf{v}^r . The Doppler velocity v_d^r is determined by taking the magnitude

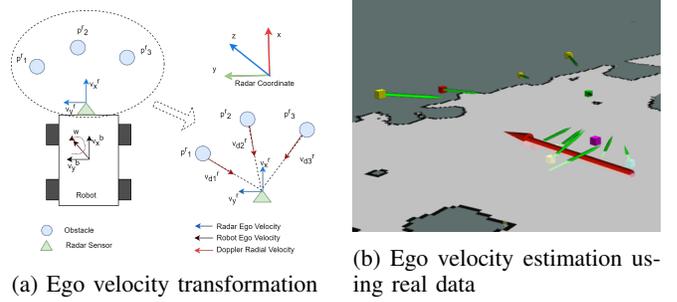


Fig. 4: Radar Ego Velocity Estimation. In (b) the green arrow (size and orientation) represents the Doppler radial velocity from the object, the red arrow (size and orientation) shows the radar ego velocity, and squares are objects.

of the projection of the relative velocity vector between the target and radar onto the ray connecting the target and the radar, as illustrated in Fig. 4. This calculation is achieved through the dot product of the target's velocity v_d^r in the radar frame $\{ \}^r$ and the unit vector pointing from the radar to the target:

$$-v_d^r = \frac{\mathbf{p}^r}{\|\mathbf{p}^r\|} \cdot \mathbf{v}^r = \mathbf{r}^r \cdot \mathbf{v}^r = r_x^r v_x^r + r_y^r v_y^r + r_z^r v_z^r \quad (6)$$

If we assume that the targets within the scene are motionless and only the sensor platform is moving, each target can serve as a constraint on the estimated velocity of the radar. If we get a set of N detections in a radar measurement and write 6 in matrix notation, we get equation 7, and the residual \mathbf{e} can be derived as equation 8.

$$\underbrace{\begin{bmatrix} -v_{d,1}^r \\ -v_{d,2}^r \\ \vdots \\ -v_{d,N}^r \end{bmatrix}}_{-\mathbf{v}_d^r = \mathbf{H}\mathbf{v}^r} = \underbrace{\begin{bmatrix} r_{x,1}^r & r_{y,1}^r & r_{z,1}^r \\ r_{x,2}^r & r_{y,2}^r & r_{z,2}^r \\ \vdots & \vdots & \vdots \\ r_{x,N}^r & r_{y,N}^r & r_{z,N}^r \end{bmatrix}}_{\mathbf{H}} \begin{bmatrix} v_x^r \\ v_y^r \\ v_z^r \end{bmatrix} + \mathbf{v}_d^r \quad (7)$$

$$\mathbf{e} = \mathbf{H}\mathbf{v}^r + \mathbf{v}_d^r \quad (8)$$

If the measurement errors are only present in the variable \mathbf{v}_d^r , we can solve Equation 7 using the linear Least

Squares (LSQ) method. However, as we discussed in III-A.1, according to the measurement principles of single-chip mmWave radar, these errors that appear in the equation 7 are not only velocity-related but also position-related, especially in terms of angle estimation of objects. Moreover, single-chip radar is highly susceptible to noise and outliers caused by factors such as multi-path and ghost points. Therefore maintaining a high signal-to-noise ratio is crucial to minimizing errors. Based on the above error analysis, a multi-strategy weighting LSQ optimization that combines residuals, azimuth contribution, and SNR consistency is proposed. We define the weight w_i for each correspondence as the product of $w_i^{residual}$, $w_i^{azimuth}$ and w_i^{snr} , where each multiplier has the following significance. First, to address the issue of outliers, we propose a residual weighting function derived from a robust kernel function [31], which belongs to a family of M-estimators and provides improved resistance to outliers.

$$w_i^{residual} = \begin{cases} 1 & \text{if } \kappa = 2 \\ \frac{2\epsilon_i}{\epsilon_i^2 + 2} & \text{if } \kappa = 0 \\ \epsilon_i \left(\frac{\epsilon_i^2}{|\kappa - 2|} + 1 \right)^{\frac{\kappa}{2} - 1} & \text{otherwise} \end{cases} \quad (9)$$

where κ represents the coefficient for the kernel's shape, $\epsilon_i = \frac{v_{d,i}^r}{\delta}$ denotes the normalized residual and δ represents the inlier noise threshold. We fix $\kappa = 1$, a pseudo-Huber kernel function, in our work to improve computational efficiency. Furthermore, the contribution of correspondences may not always be evenly distributed in azimuth. The following weighting function equation considering single-chip radar azimuth estimation is proposed to reduce the impact of systematic errors in angular measurements on radar ego-motion estimation.

$$w_i^{azimuth} = \frac{\cos(\theta_i)}{\sum_{i=1}^N \cos(\theta_i)} \quad (10)$$

Third, since the SNR channel provides additional information for measurement confidence, equation 11 is designed to penalize measurements with low SNR.

$$w_i^{snr} = e^{-\frac{|s_i|}{s_{max}}} \quad (11)$$

Applying the residual definition of equation 8 yields the residual with multi-strategy weighting LSQ optimization:

$$\mathbf{e}^m = (w_i^{residual} \cdot w_i^{azimuth} \cdot w_i^{snr})(\mathbf{H}\mathbf{v}^r + \mathbf{v}_d^r) \quad (12)$$

To account for the dynamic environment, a three-point RANdom SAMple Consensus (RANSAC) [20] approach is employed to identify the set of inliers from static objects for equation 12. This method is highly efficient as only three points are required to evaluate a hypothesis.

So far we have computed the linear velocities in x and y directions in the radar frame $\{\}^r$. Since the angular velocity $\boldsymbol{\omega}_m^b$ of the robot base frame $\{\}^b$ leads to an additional velocity in the radar frame $\{\}^r$ as shown in equation 13, we apply rigid body motion to determine the $\boldsymbol{\omega}_m^b$ as equation 13.

$$\mathbf{v}^r = \boldsymbol{\omega}_m^b \times \mathbf{l}_{br}^b + \mathbf{v}^b \quad (13)$$

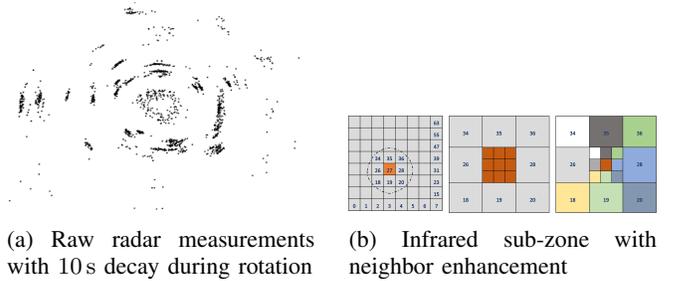


Fig. 5: Hybrid Infradar Sensor Model

C. Infradar Hybrid Sensor Model

The hybrid Infradar sensor model presented in this section contains three main components, namely the radar likelihood field model [32], the infrared endpoint beam model, and the neighbor-enhancement model. As the visualization of raw single-chip radar measurements presented in Fig. 5a, it is obvious that the *Beam Range* model is inadequate for the single-chip radar due to signal sparsity, multi-path reflections, and the presence of ghost points. To address these challenges, we have utilized the *Likelihood Field* model for radar observations. The noise in single-chip radar measurement is typically modeled using Gaussian distributions. In x-y-z-space, this requires finding the nearest obstacle in the map to the measurement coordinates. We can then compute the Euclidean distance, denoted by d , between the measurement coordinates and the nearest object in the map $\{\}^m$. The probability of a sensor measurement is modeled by a zero-centered Gaussian distribution that captures the characteristics of the sensor noise:

$$\phi(d, \sigma) = \exp\left(-\frac{d^2 s_i}{2\sigma^2 s_{max}}\right) \quad (14)$$

here, σ is the standard deviation of the radar noise, s_i and s_{max} are the SNR of the radar detections. The likelihood field is then the product of probability from each measured point to the closest point in the map.

In order to integrate single-chip infrared sensor distance measurements we use the endpoint beam model proposed by [32]. Each of the 8×8 zones is approximated as a conical beam from the sensor origin to the obstacle, with each measurement returning the nearest obstacle in that beam zone, as shown in Fig. 5b. Here, the likelihood of a single zone measurement $l_{t,id}$ is determined by the distance d_{id} between the corresponding hypothetical zone beam endpoint and the nearest obstacle represented in the map:

$$\phi(d_{id}, \sigma) = \exp\left(-\frac{d_{id}^2}{2\sigma^2}\right) \quad (15)$$

where σ represents the standard deviation of the single-chip infrared sensor noise and d_{id} is the distance of id_{th} zone. The integration of a full infrared scan, which consists of 64 zone beams, is calculated by multiplying the likelihoods of each individual beam:

$$p(\mathbf{1}_t | \mathbf{x}_t) = \prod_{id=1}^{64} \phi(d_{id}, \sigma) \quad (16)$$

Although the multi-zone single-chip infrared sensor has achieved a high resolution compared to conventional single-chip infrared sensors, it is still insufficient for precise indoor localization. We, therefore, propose a novel neighbor-enhancement approach based on the endpoint beam measurement model to improve the resolution of the sensor. First, each zone (square) is divided equally into nine sub-zones (squares). The original zone distance is assigned to the sub-zone in the center, as shown in Fig. 5b, and the remaining eight sub-zones are filled with no measurement. The next step is to check the neighboring zone of each sub-zone. If the difference between their detection distance and the distance of the zone where the sub-zone is located falls below the threshold γ , the sub-zone is interpolated with the average of two distances. The threshold value is defined as $\gamma = \zeta d_{id}$, where ζ represents the enhancement-factor. On the other hand, if the difference is above the threshold, the sub-zone remains empty. In this way, we increase the resolution from 8×8 to $8 \times 8 \times 8$. The whole neighbor-enhancement approach is demonstrated in Fig. 5b.

D. Infradar Monte Carlo Localization

This section first outlines the MCL algorithm and then describes the adaptations we have made for Infradar-Localization.

The three-step approach of the conventional MCL method is adopted in this paper, which includes a prediction phase using robot motion model, a correction phase using a sensor measurement model, and a resampling phase [33]. The particle filter utilized in this paper is factorized based on a prior map M built also by Infradar, as expressed below:

$$\hat{p}(x_t | Z_t, M) = \eta p(z_t | x_t, M) \hat{p}(x_t | Z_{t-1}, M) \quad (17)$$

where η is a normalizer and Z_t represents the measurements. During the prediction phase, the state of the robot is estimated through the use of samples:

$$\hat{p}(x_t | Z_{t-1}, M) = \sum_i p(x_t | s_{t-1}^i, u_{t-1}) \quad (18)$$

here, s^i is the i_{th} particle in the estimator and u is the robot motion, which is predicted based on the Infradar motion model. In the correction phase, the weight of each particle is calculated from the Infradar hybrid sensor model given observations z_t in map M . During the resampling phase, the particles with higher weights are retained and the robot's position is calculated by combining the states and weights of these particles. The particle filter can be initialized in two ways: by using a distribution samples with equal weights around the initial pose estimate, which is called *tracking*, or by using a uniform distribution over all possible hypotheses, which is referred to as *global localization*.

However, a key limitation of MCL arises from the static world assumption, or Markov assumption [34]. To address this challenge, we propose a novel Dynamic Outlier Removal method that leverages the unique properties of single-chip radar. In section III-B we present a three-point RANSAC to distinguish between dynamic and static objects. The static

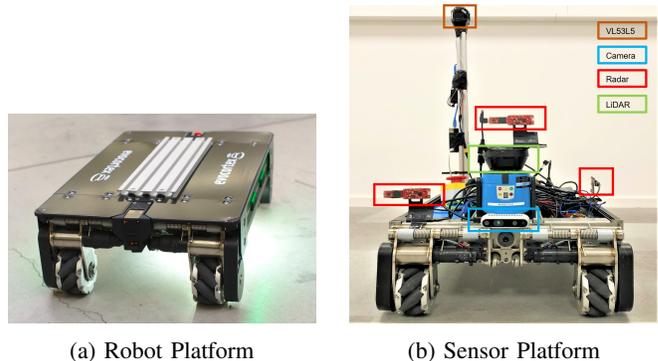


Fig. 6: Experimental Setup

objects are then used as inliers to estimate the radar ego velocity. The dynamic objects, which are identified as outliers by RANSAC, will also be utilized to exclude measurements of infrared sensors that are within a certain distance threshold ρ .

IV. EXPERIMENT EVALUATION

This section presents our experiments to demonstrate the effective capabilities of our Infradar-Localization framework.

A. Experimental Setup

The test robot platform integrated with an omnidirectional mecanum drive (as shown Fig. 6 left), was used in the experiment. Experimental data was acquired using the sensor platform shown in Fig. 6 on the right, which featured three single-chip Infradar sensors, each containing a single-chip mmWave Radar sensor (TI IWR6843AOP with 120° FoV), and a custom single-chip infrared sensor (VL53L5CX). All Infradar sensors were processed on-chip. In addition, the robot was also equipped with a Sick LiDAR sensor and an Intel Realsense D435 camera to provide baseline results. The proposed algorithm was run on NVIDIA Jetson TX2 NX Module.

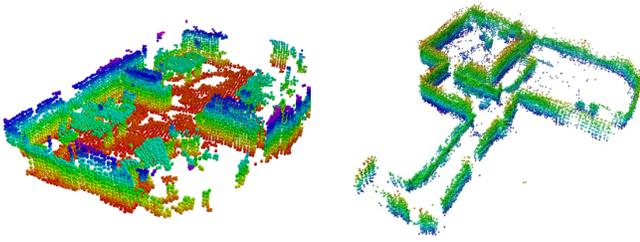
The calibration method described in [35] was used to calibrate the radar sensor's extrinsic parameters, while the approach presented in [26] was used for calibrating the infrared sensor. The extrinsic transformation between the coordinate frames of the infrared and radar sensors was measured manually.

Our implementation is built upon the open-source framework *mcl_3dl*¹, which provides pointcloud-based 3D Monte Carlo localization. We expanded this framework by introducing our Infradar motion model, integrating our hybrid Infradar sensor model, and optimizing the processing chain for real-time data processing. The algorithm is programmed in C++ and is implemented on Ubuntu 18.04 and ROS Melodic.

B. Evaluation

To evaluate the localization results, our method is compared with the ground truth, which is provided by a LiDAR-Camera-based localization system. The robot platform was

¹https://github.com/at-wat/mcl_3dl



(a) 3D map of an office (b) 3D map of a warehouse

Fig. 7: 3D maps generated by Infradar sensors

manually controlled to move around different indoor scenarios, typically traveling between 0.4 m/s and 0.6 m/s. When turning, the platform moved at a maximum of 0.5 rad/s. Initially, we created a 3D occupancy map using Infradar sensors with the assistance of ground truth. We then employed only the Infradar sensors to carry out localization within this map, as shown in Fig. 7. This is also where our approach differs from traditional methods, which relied on the use of high-precision LiDAR to build maps. The performance of the proposed method was evaluated in terms of *tracking* and *global localization*.

1) *Tracking*: During the pose tracking experiment, we provided the algorithm with the robot’s initial position and orientation and then manually controlled the robot to traverse varying scenarios. The result is shown in Fig. 8 and 9, where the trajectories of ground truth and Infradar-Localization are drawn in red and green lines, respectively. LiDAR maps are only used for reference. Translation error is defined as the root square error of each pose Euclidean distance with respect to the ground truth pose. It is obvious that despite relying on low-cost Infradar sensors alone, the proposed approach is still able to accurately track the robot’s pose in different complex and dynamic environments and under varying robot motions. In brief, our method is capable of achieving a mean translation error of 9 cm in various environments while maintaining an average computing time of 21 ms per frame. These tracking results are comparable to high-precision 3D LIDAR used for indoor localization, as reported in a comparative study by [11].

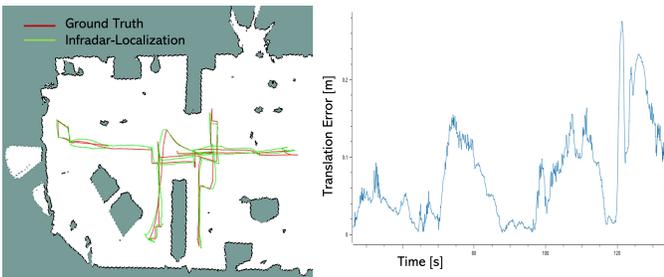


Fig. 8: Example of Infradar-Localization in a 7 m × 10 m office environment. Mean translation error 0.07 m, maximum translation error 0.28 m.

2) *Global Localization*: To further demonstrate the localization performance, the proposed approach is also evaluated

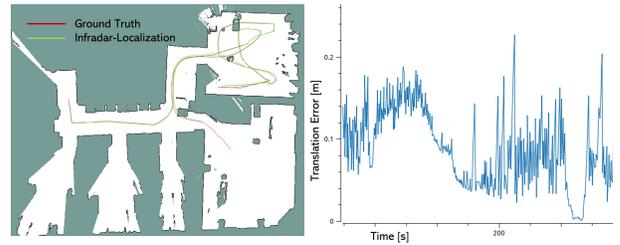


Fig. 9: Example of Infradar-Localization in a 13 m × 17 m warehouse environment with dynamic obstacles (moving people and robots). Mean translation error 0.09 m, maximum translation error 0.24 m.

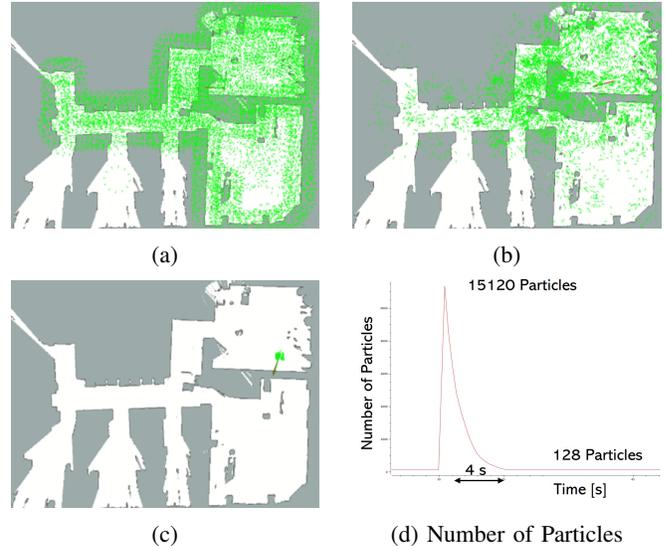


Fig. 10: Global Localization with Infradar: (a), (b), and (c) represent the state of particles (green arrow) while red arrow denotes the ground truth.

with *global localization*. This means that there is no robot’s initial pose for the algorithm. In this paper, the number of particles in the initial distribution is determined based on the size of the 3D map, which is 15120 particles for the 13 m × 17 m warehouse map. As shown in 10, in the initial state the particles are evenly distributed over all spaces, and then with Infradar measurements, these particles converge to the correct pose. Through repeated experiments at different locations, the proposed method is capable of converging the particles to the correct location in less than 5 s on average. It can be concluded that the proposed method offers reliable and efficient global localization solutions in indoor environments.

V. CONCLUSIONS

This paper introduces Infradar-Localization, the first low-cost indoor Monte Carlo localization framework based on single-chip infrared and radar sensors. The framework mainly consists of Infradar ego-motion estimation, hybrid Infradar sensor model, dynamic outlier removal, and lightweight Monte Carlo localization. The proposed approach

is capable of supporting real-time localization on an embedded PC with low-cost Infradar sensors. Experiments in different environments have been performed to evaluate the proposed method. The results demonstrate the reliability and accuracy of the proposed framework, which is comparable to high-precision LiDAR-based localization methods. This remarkable performance, combined with the low-cost and lightweight nature of our proposed solution enables more precise and cost-effective navigation systems for most mobile platforms such as UAVs and autonomous mobile robots. Future works will extend the proposed system to a full simultaneous localization and mapping (SLAM) framework.

REFERENCES

- [1] T.-M. Nguyen, Z. Qiu, T. H. Nguyen, M. Cao, and L. Xie, "Persistently excited adaptive relative localization and time-varying formation of robot swarms," *IEEE Transactions on Robotics*, vol. 36, no. 2, pp. 553–560, 2019.
- [2] H. Zou, C.-L. Chen, M. Li, J. Yang, Y. Zhou, L. Xie, and C. J. Spanos, "Adversarial learning-enabled automatic wifi indoor radio map construction and adaptation with mobile robot," *IEEE Internet of Things Journal*, vol. 7, no. 8, pp. 6946–6954, 2020.
- [3] H. Hong and B. H. Lee, "Probabilistic normal distributions transform representation for accurate 3d point cloud registration," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 3333–3338.
- [4] C. Debeunne and D. Vivet, "A review of visual-lidar fusion based simultaneous localization and mapping," *Sensors*, vol. 20, no. 7, p. 2068, 2020.
- [5] D. Nistér, O. Naroditsky, and J. Bergen, "Visual odometry," in *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, vol. 1. Ieee, 2004, pp. I–I.
- [6] M. He, C. Zhu, Q. Huang, B. Ren, and J. Liu, "A review of monocular visual odometry," *The Visual Computer*, vol. 36, no. 5, pp. 1053–1065, 2020.
- [7] T. Taketomi, H. Uchiyama, and S. Ikeda, "Visual slam algorithms: A survey from 2010 to 2016," *IPSJ Transactions on Computer Vision and Applications*, vol. 9, no. 1, pp. 1–11, 2017.
- [8] I. A. Kazerouni, L. Fitzgerald, G. Dooly, and D. Toal, "A survey of state-of-the-art on visual slam," *Expert Systems with Applications*, p. 117734, 2022.
- [9] P. Karpyshev, E. Kruzhkov, E. Yudin, A. Savinykh, A. Potapov, M. Kurenkov, A. Kolomeytshev, I. Kalinov, and D. Tsetsrukou, "Mucaslam: Cnn-based frame quality assessment for mobile robot with omnidirectional visual slam," in *2022 IEEE 18th International Conference on Automation Science and Engineering (CASE)*. IEEE, 2022, pp. 368–373.
- [10] Z. Wei, F. Zhang, S. Chang, Y. Liu, H. Wu, and Z. Feng, "Mmwave radar and vision fusion based object detection for autonomous driving: A survey," *arXiv e-prints*, pp. arXiv–2108, 2021.
- [11] Q. Zou, Q. Sun, L. Chen, B. Nie, and Q. Li, "A comparative analysis of lidar slam-based indoor navigation for autonomous vehicles," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 6907–6921, 2021.
- [12] I. Ullah, Y. Shen, X. Su, C. Esposito, and C. Choi, "A localization based on unscented kalman filter and particle filter localization algorithms," *IEEE Access*, vol. 8, pp. 2233–2246, 2019.
- [13] A. Fischer, Z. Tong, A. Hamidipour, L. Maurer, and A. Stelzer, "77-ghz multi-channel radar transceiver with antenna in package," *IEEE Transactions on Antennas and Propagation*, vol. 62, no. 3, pp. 1386–1394, 2013.
- [14] C.-Y. Ho, S.-C. Hsieh, M.-F. Jhong, C.-C. Wang, and C.-Y. Ting, "A 77ghz antenna-in-package with low-cost solution for automotive radar applications," in *2018 IEEE 68th Electronic Components and Technology Conference (ECTC)*. IEEE, 2018, pp. 191–196.
- [15] D. Kellner, M. Barjenbruch, J. Klappstein, J. Dickmann, and K. Dietmayer, "Instantaneous ego-motion estimation using doppler radar," in *16th International IEEE Conference on Intelligent Transportation Systems (ITSC 2013)*. IEEE, 2013, pp. 869–874.
- [16] —, "Instantaneous ego-motion estimation using multiple doppler radars," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 1592–1597.
- [17] A. Kramer, C. Stahoviak, A. Santamaria-Navarro, A.-A. Agha-Mohammadi, and C. Heckman, "Radar-inertial ego-velocity estimation for visually degraded environments," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2020, pp. 5739–5746.
- [18] C. Doer and G. F. Trommer, "An ekf based approach to radar inertial odometry," in *2020 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*. IEEE, 2020, pp. 152–159.
- [19] Y. S. Park, Y.-S. Shin, J. Kim, and A. Kim, "3d ego-motion estimation using low-cost mmwave radars via radar velocity factor for pose-graph slam," *IEEE Robotics and Automation Letters*, vol. 6, no. 4, pp. 7691–7698, 2021.
- [20] C. Doer and G. F. Trommer, "Radar visual inertial odometry and radar thermal inertial odometry: Robust navigation even in challenging visual conditions," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2021, pp. 331–338.
- [21] C. Doer, J. Atman, and G. F. Trommer, "Gnss aided radar inertial odometry for uas flights in challenging conditions," in *2022 IEEE Aerospace Conference (AERO)*. IEEE, 2022, pp. 1–10.
- [22] Y. S. Park, J. Kim, and A. Kim, "Radar localization and mapping for indoor disaster environments via multi-modal registration to prior lidar map," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2019, pp. 1307–1314.
- [23] C. Xiaoxuan Lu, S. Rosa, P. Zhao, B. Wang, C. Chen, J. A. Stankovic, N. Trigoni, and A. Markham, "See through smoke: Robust indoor mapping with low-cost mmwave radar," *arXiv e-prints*, pp. arXiv–1911, 2019.
- [24] K. McGuire, C. De Wagter, K. Tuyls, H. Kappen, and G. C. de Croon, "Minimal navigation solution for a swarm of tiny flying robots to explore an unknown environment," *Science Robotics*, p. eaaw9710, 2019.
- [25] K. N. McGuire, G. C. de Croon, and K. Tuyls, "A comparative study of bug algorithms for robot navigation," *Robotics and Autonomous Systems*, 2019.
- [26] V. Niculescu, H. Müller, I. Ostovar, T. Polonelli, M. Magno, and L. Benini, "Towards a multi-pixel time-of-flight indoor navigation system for nano-drone applications," in *2022 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. IEEE, 2022, pp. 1–6.
- [27] A. Kongpecth, N. Kwankeo, and V. Manthamkarn, "360 degrees object detection using multiple tof sensors for educational robot," in *2022 19th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*. IEEE, 2022, pp. 1–4.
- [28] A. Ruget, M. Tyler, G. M. Martín, S. Scholes, F. Zhu, I. Gyongy, B. Hearn, S. McLaughlin, A. Halimi, and J. Leach, "Real-time, low-cost multi-person 3d pose estimation," *arXiv preprint arXiv:2110.11414*, 2021.
- [29] H. Müller, N. Zimmerman, T. Polonelli, M. Magno, J. Behley, C. Stachniss, and L. Benini, "Fully on-board low-power localization with multizone time-of-flight sensors on nano-uavs," *arXiv preprint arXiv:2212.00710*, 2022.
- [30] T. Instruments, "The fundamentals of millimeter wave radar sensors," 2020.
- [31] J. T. Barron, "A general and adaptive robust loss function," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4331–4339.
- [32] S. Thrun, "A probabilistic online mapping algorithm for teams of mobile robots," *The International Journal of Robotics Research*, vol. 20, no. 5, pp. 335–363, 2001.
- [33] F. J. Perez-Grau, F. Caballero, A. Viguria, and A. Ollero, "Multi-sensor three-dimensional monte carlo localization for long-term aerial robot navigation," *International Journal of Advanced Robotic Systems*, vol. 14, no. 5, p. 1729881417732757, 2017.
- [34] S. Thrun, "Probabilistic robotics," *Communications of the ACM*, vol. 45, no. 3, pp. 52–57, 2002.
- [35] C. Doer and G. F. Trommer, "Radar inertial odometry with online calibration," in *2020 European Navigation Conference (ENC)*. IEEE, 2020, pp. 1–10.