

# Milli-RIO: Ego-Motion Estimation with Millimetre-Wave Radar and Inertial Measurement Unit Sensor

Yasin Almalioglu<sup>1</sup>, Mehmet Turan<sup>2</sup>, Chris Xiaoxuan Lu<sup>1</sup>, Niki Trigoni<sup>1</sup>, and Andrew Markham<sup>1</sup>

**Abstract**—With the fast-growing demand of location-based services in various indoor environments, robust indoor ego-motion estimation has attracted significant interest in the last decades. Single-chip millimeter-wave (MMWave) radar as an emerging technology provides an alternative and complementary solution for robust ego-motion estimation. This paper introduces Milli-RIO, a MMWave radar based solution making use of a fixed beam antenna and inertial measurement unit sensor to calculate 6 degree-of-freedom pose of a moving radar. Detailed quantitative and qualitative evaluations prove that the proposed method achieves precisions on the order of few centimetres for indoor localization tasks.

## I. INTRODUCTION

Robust ego-motion estimation solutions for indoor environments have a variety of real-world applications ranging from emergency evacuation to mobility aids for people with disabilities, which is still a challenging task. The millimeter-wave (MMWave) radar provides robust and reliable perceptual information of the environment especially in degraded visual conditions. MMWave radar is extensively used in several domains such as the military (air and maritime surveillance, missile guidance, etc.), the civil aviation (approach radar, surface movement radar) or the remote sensing (planetary observation) [1]. Recent advances in radar technology and material science enabled a progressive adaptation of MMWave radars to smaller platforms in terms of dimension ( $\sim 20 \times 30$ cm), weight ( $\sim 200$ gr), energy consumption ( $\sim 2$ W) and cost issues ( $\sim \$100$ ), (see Fig. 1). Their small size, low cost and fine accuracy make them suitable especially for portable low power applications [2].

Several research groups have been proposing MMWave radars as a solution for various mobile robot tasks such as navigation, localization and mapping in recent years. In obstacle detection, MMWave radar is widely studied in automotive applications to detect moving and static targets (cars, pedestrians) [3], [4]. Several studies are proposed to investigate imaging capabilities of the radars for environment representation [5], 2D/3D simultaneous localization and mapping (SLAM) [6]. MMWave radars are also fused with visual sensors for obstacle detection and map reconstruction, combining the robust depth detection ability of the radar in difficult environmental conditions with high

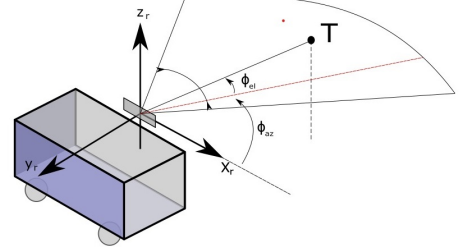


Fig. 1: **Millimetre-Wave (MMWave) radar system.** MMWave radar is a special class of radar technology that uses millimetre wavelength radio frequency (RF) signals. Our MMWave system operates at 7681 GHz spectrum, resulting in an ability to detect movements smaller than a fraction of a millimetre.

spatial resolution of the visual sensors [7]. However, ego-motion estimation methods specifically designed for single-chip, low-cost MMWave radars are needed to fully utilize the complementary features of MMWave radars for indoor location based services.

In this paper, we propose Milli-RIO, an ego motion estimation method based on single-chip MMWave radar and inertial measurement unit (IMU) sensor. The main contributions of our method are as follows:

- To the best of our knowledge, this is the first indoor ego-motion estimation approach using single-chip MMWave radar sensor.
- Unlike existing works in literature, the proposed system is based on single-chip MMWave radar scans, making it effective for indoor applications in terms of size, cost and energy consumption.
- We propose a model-free motion dynamics estimation technique for unscented Kalman filter (UKF) using Recurrent Neural Network (RNN).

As outline of the paper, Section II presents the related work. Section III introduces the proposed ego-motion estimation method based on MMWave radar. The experimental setup is described in Sec. IV. The qualitative and quantitative results are presented in Section IV. Section V concludes the study and gives future directions.

## II. RELATED WORK

Feature extraction is a fundamental task in radar motion estimation. The traditional visual localization techniques in literature such as amplitude gridmaps that transform the radar scans into grayscale images followed by SIFT and FAST

<sup>1</sup>Yasin Almalioglu, Chris Xiaoxuan Lu, Niki Trigoni, and Andrew Markham are with the Computer Science Department, The University of Oxford, UK {yasin.almalioglu, xiaoxuan.lu, niki.trigoni, andrew.markham}@cs.ox.ac.uk

<sup>2</sup>Mehmet Turan is with the Institute of Biomedical Engineering, Bogazici University, Turkey mehmet.turan@boun.edu.tr

feature extractions are investigated by [8]. The gridmaps are further studied to find continuous areas using DBSCAN, MSER, and the connected components [9]. The radar-specific solutions utilize data distortion, which is used as sources of information in order to estimate the vehicle displacement [10]. Another technique exploits spatiotemporal continuity in radar scans inferring the appearance of landmarks by estimating the radar noise characteristics [11]. In 2D radar scan processing, the accurate range information calculated with the greatest power return per azimuth eliminates the need for a filtering, which can potentially discard relevant information [12]. Finding global correspondences of feature points at different timestamps is another major task of ego-motion estimation approaches. To combine visual and radar sensors, [8] pairs radar and visual landmarks with similar feature descriptors. Vision-radar fusion approaches use radar occupancy grids to associate both sensor measurements [13]. Feature descriptors work well for images that contain complex and high-density information. However, they are unable to create useful feature descriptions from radar scans that characteristically have significant noise and sparse information. Multi-sensor fusion techniques provide an alternative to feature-based radar odometry, which uses odometry information from additional sensors to transform the incoming radar landmark pointcloud and register it to an existing landmark map. They usually make use of nearest neighbor point matching [13], and Monte Carlo methods to derive a solution from probabilistic weights [14]. The relative motion is estimated using the data association between the radar point cloud and map, which is then fused to the original odometry readings. Although existing multi-sensor fusion methods are promising, they make use of sensors that already provide highly accurate odometry results.

In radar based motion estimation systems, data association is frequently achieved by a scan matching algorithm that tracks common landmarks across consecutive radar scans. IMU as a convenient, cost-effective and highly portable sensor solution provides an additional odometry information to eliminate the effects of the radar noise floor and to improve data association performance. Iterative closest point (ICP) approach is typically used to iteratively align the radar point clouds until the pre-defined termination criteria is met [15]. In [16], the researchers developed a quantitative function describing the quality of map created by superimposing radar point clouds according to the unknown motion parameters. All of these works assume small incremental motions between the radar scans, which imposes an undesirable constraint on the algorithms and prevents them from being applied to arbitrary inputs. An innovative technique well suited for high velocities utilizes the radar scan distortions that are often a drawback of mobile radar systems to eliminate the high velocity effects using an extended Kalman filter [10]. Other scan matching algorithms operate directly on the radar outputs instead of extracting landmarks. The Fourier-Mellin transform enables efficient computation of the vehicle's rotation and translation from the entire radar output [17]. The Doppler radar returns the position and speed of the

objects around and the vehicle motion is easily computed relative to the surrounding objects given a sufficient amount of radar scans [18]. Both methods are encumbered by heavy preprocessing. However, the existing works for radar based ego-motion estimation are developed for mechanically rotating radars, which are heavy ( $\sim 10kg$ ) and expensive ( $\sim \$10k$ ) sensors, and thus, not suitable for portable indoor location based services. In this paper, we present a novel and robust motion estimation approach for indoor localization tasks making use of single-chip MMWave radar and IMU sensor to eliminate deficiencies of both sensors such as biases in IMU output, noises and sparse measurements in radar scans. Our method is a direct approach which uses entire information captured by the single-chip MMWave radar, which provides much sparser and noisy scans than mechanically rotating radars [19].

### III. MILLIMETRE-WAVE RADAR BASED EGO-MOTION ESTIMATION

MMWave radar is based on the principle of frequency modulated continuous wave (FMCW) radar, which has the ability to simultaneously measure the range and relative radial speed of a target point. Milli-RIO is an ego-motion estimation system that exploits the unique properties of single-chip MMWave radar. It transmits an RF signal and records reflection from target point that is collected in a point cloud. It then calculates ego-motion by registering the generated sparse point cloud, which uses IMU as an auxiliary sensor to improve registration performance. In this section, we present the proposed MMWave radar based data association and ego-motion estimation algorithms. Moreover, we explain details of RNN-based motion model used in the joint MMWave radar-IMU ego-motion estimation.

#### A. Data Association

Milli-RIO achieves robust point correspondences across the consecutive scans using high-level information in the radar output. The algorithm seeks to find the largest subsets of two point clouds that have a similar geometry. *A priori* knowledge of the orientations or displacements of the scans relative to one another affects the performance of point cloud registration. An IMU sensor provides a likelihood estimation for the relative transformation between two scans. Thus, the algorithm is not constrained to have a good initial estimate of the relative pose, enabling registration of point clouds captured at arbitrary times without any *a priori* map representation. In case of insufficient overlap between scans, the algorithm is able to re-initialize registration process using IMU measurements. The proposed approach performs this registration using not only individual unary landmark descriptors, but also mutual relations between landmarks. Let the landmarks  $A$ ,  $B$  and  $C$  be the vertices of a scalene triangle, and  $\mathcal{D}$  be the set of distances from each point  $A$ ,  $B$  and  $C$  to its neighbours.  $\mathcal{D}$  is unique to this set of points regardless of the overall point cloud placement, which allows the landmark to be uniquely matched to its counterpart in any other point cloud transformed by a rigid body transformation of the

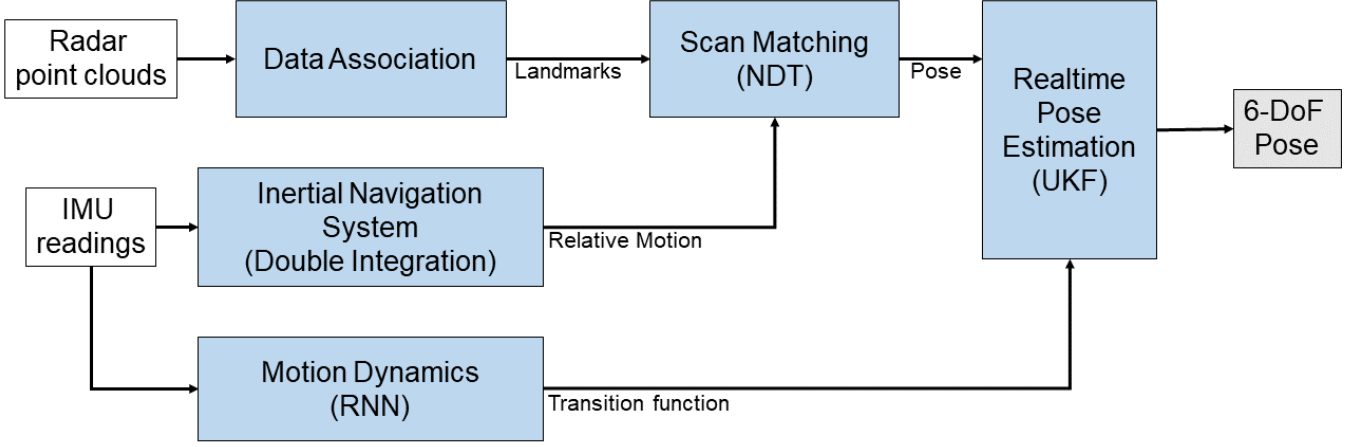


Fig. 2: **Ego-motion estimation workflow.** Raw MMWave radar point clouds are processed in the data association module to extract landmarks, which are passed to the second module which registers them using NDT scan matching algorithm. In parallel, relative ego-motion is estimated from IMU readings using our inertial navigation system. Radar and IMU estimations are fused in real-time pose estimation module using an unscented Kalman filter to regress final 6 DoF pose values, which uses recurrent neural network based transition model. White boxes, gray box and blue boxes represent inputs to the system, output, and processing units, respectively.

original triangle. This triangular metric reduces the likelihood of an individual point having the same set of pairwise distances to its neighbours as another. Moreover, the pairwise relationships within the point cloud are not dependent on the exact position and orientation of the point cloud, making large disparities between the placements and orientations of the point cloud inconsequential. These observations provide reliable matches for our large landmark sets. The landmark locations and detections are noisy, resulting in an association of non-existent points after the rigid body transformation.

Let  $\mathcal{P}_1$  and  $\mathcal{P}_2$  be the two consecutive radar scan inputs to the data association module, where  $\mathcal{P}_1$  is the original set of landmarks in Cartesian coordinates, and  $\mathcal{P}_2$  compensates for the sparse points by generating a binary Cartesian grid of resolution  $\beta$  that is interpolated from the binary polar grid of landmarks.  $\mathcal{P}_2$  is sparser and only used to sidestep the bias caused by the range-density trade-off during the layout processing of the environment. Data association is performed on  $\mathcal{P}_1$  and returns a set of matches  $\mathcal{M}$  that contains  $(i, j)$  tuples such that the landmark  $\mathcal{P}_1\{i\}$  corresponds to  $\mathcal{P}_2\{j\}$ . Data association remains accurate by operating on the landmarks and preserves a static map by interpreting the environment in Cartesian space. In the first step, for each point in  $\mathcal{P}_1$ , a potential point match in  $\mathcal{P}_2$  is calculated based on a unary comparison. For each pair of proposed matches  $g = (i, i')$  and  $h = (j, j')$ , the non-negative compatibility score is computed and assigned to the elements  $(g, h)$  and  $(h, g)$  of the  $W \times W$  symmetric and diagonally dominant matrix  $C$ . A similar relationship between  $i, j$  in  $\mathcal{P}_1$  and  $\mathcal{P}_2$  is used to calculate a compatibility score for the landmark matches between  $g$  and  $h$ . The score is defined by distances between corresponding point pairs in two scans, measuring that correctly identified landmarks have the same distance in

any two radar scans. The overall compatibility is maximum in the optimal set of matches  $\mathcal{M}$  [20]. The greedy method iteratively collects satisfactory matches in the set  $\mathcal{M}$ . On each iteration, it evaluates the remaining valid matches and returns maximum accepted reward. Matches which conflict with it are excluded from set  $\mathcal{M}$ . The algorithm terminates when the most recently selected match yields a reward score more than  $\alpha$  percent of the landmarks in either set.

### B. Relative Motion Estimation

The proposed motion estimation method performs a data association on  $\mathcal{M}$  (Section III-A), which is invariant to arbitrary rigid body transformations in terms of distance and rotation. In the proposed system, we estimate the sensor trajectory by iteratively applying the normal distributions transform (NDT) scan matching technique [21] to find the rigid body motion given two sets of corresponding points. NDT has been shown to have a better performance than other scan matching algorithms, such as iterative closest points, in terms of both reliability and processing speed [22]. We can estimate the sensor ego-motion by iteratively applying a scan matching algorithm. However, the performance of any scan matching algorithm is affected by the number of point correspondences between two sets, which might fail due to large displacements caused by rapid motions. In order to deal with this problem, we integrate angular velocity data provided by IMU sensor to the NDT scan matching algorithm using UKF [23]. The pipeline of our method is demonstrated in Fig. 2.

We define the sensor state to be:

$$\mathbf{x}_t = [\mathbf{p}_t, \mathbf{q}_t, \mathbf{v}_t, \mathbf{b}_t^a]^T, \quad (1)$$

where,  $\mathbf{p}_t$  is the position,  $\mathbf{q}_t$  is the rotation quaternion,  $\mathbf{v}_t$  is the velocity,  $\mathbf{b}_t^a$  is the bias of the angular velocity of the

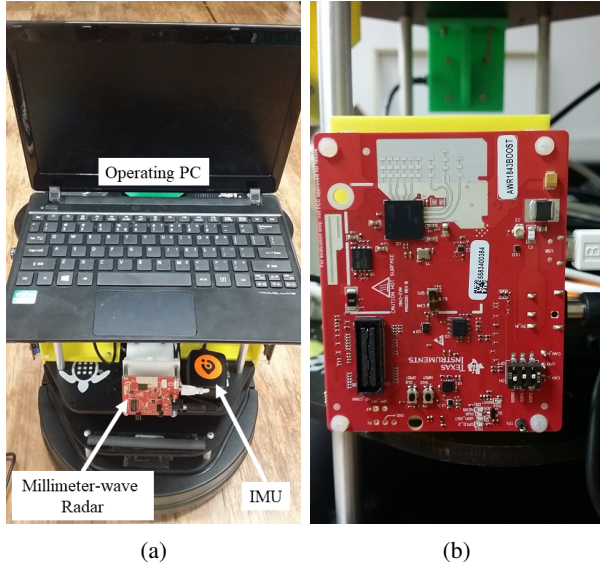


Fig. 3: **Experimental setup.** a) Turtlebot2 data collection platform. b) Short-range, TI AWR1843BOOST model, Millimetre-Wave radar employed in the experiments.

sensor at time  $t$ . Assuming a transition function  $f(\cdot)$  for the sensor motion model and constant bias for the angular velocity sensor, the system equation for predicting the state is defined as:

$$\mathbf{x}_t = [\mathbf{p}_{t-1} + f(\mathbf{x}_{t-1}), \mathbf{q}_{t-1} \cdot \Delta \mathbf{q}_t, \mathbf{v}_{t-1}, \mathbf{b}_{t-1}^a]^T, \quad (2)$$

where  $\Delta \mathbf{q}_t$  is the rotation during  $t-1$  and  $t$  caused by  $y$  the bias-compensated angular velocity  $\mathbf{a}_t' = \mathbf{a}_t - \mathbf{b}_{t-1}^a$ :

$$\Delta \mathbf{q}_t = \left[ 1, \frac{\Delta t}{2} \mathbf{a}_t^{x'}, \frac{\Delta t}{2} \mathbf{a}_t^{y'}, \frac{\Delta t}{2} \mathbf{a}_t^{z'} \right]^T. \quad (3)$$

Using Eq. 2 and UKF, the system predicts the sensor pose, and applies NDT to register the observed point cloud into the global map, the estimated  $\mathbf{x}_t$  and  $\mathbf{q}_t$  being the initial guess of the sensor pose. Then, the system corrects the sensor state using sensor pose estimated by the scan matching  $\mathbf{z}_t = [\mathbf{p}_t', \mathbf{q}_t']^T$ . The observation equation is defined as:

$$\mathbf{z}_t = [\mathbf{p}_t, \mathbf{q}_t]^T. \quad (4)$$

We normalize  $\mathbf{q}_t$  in the state vector after each prediction and correction step of UKF to avoid norm changes due to unscented transform and accumulated calculation error. It is worth mentioning that we also implemented pose prediction which takes acceleration into account, as well, which is omitted in the proposed approach. However, the estimation performance deteriorates due to strong acceleration noise and constant bias.

### C. RNN-based Motion Model

Existing data fusion methods based on traditional filters have limitations for nonlinear dynamic systems. UKF accommodates a wide variety of dynamic models, allowing for highly complex dynamics in the state variables given an accurate motion model.

In the last decade, deep learning (DL) techniques have exceeded the performance of traditional methods in various domains such as computer vision, speech recognition and natural language processing. Contrary to these high-level tasks, data fusion problem is mainly concerned with motion dynamics and the temporal relations across pose sequences coming from different ego-motion algorithms, which can be formulated as a sequential learning problem. Unlike traditional feed-forward DL networks, RNNs are very suitable to model the dependencies across time sequences and to create a temporal motion model. RNNs represent current hidden state as a function of arbitrary sequences of inputs by having a memory of hidden states over time and directed cycles among hidden units. Thus, the real-time pose estimation of the current time step benefits from information encapsulated in previous time steps and is suitable to formulate the state transition function  $f$  in Eq. 2 using RNN [24]. UKF tracks the 6-DoF pose of moving radar using the transition function modelled by the Long Short-Term Memory (LSTM) network. To train the LSTM, the inputs are accelerometer and gyroscope readings (states) at time step  $t-1$ , and output labels are 6-DoF poses at time  $t$ . In that way, the LSTM learns the motion model of the mobile radar.

LSTM is a specific implementation of RNN to avoid the vanishing gradient problem, enabling exploitation of temporal position information for a long time. Thus, LSTM has a higher capacity of learning long-term relations among the pose sequences by introducing memory gates such as input, forget and output gates, and hidden units of several blocks. The input gate controls the amount of new information fed into the current state, the forget gate balances the information accumulated in the memory, and the output gate triggers the activations. Let the the input vector be  $x_k$  at time  $k$ , the output vector  $h_{k-1}$  and the cell state vector  $c_{k-1}$  of the previous LSTM unit. The LSTM updates at time step  $k$  according to the following equations:

$$\begin{aligned} f_k &= \sigma(W_f \cdot [x_k, h_{k-1}] + b_f) \\ i_k &= \sigma(W_i \cdot [x_k, h_{k-1}] + b_i) \\ g_k &= \tanh(W_g \cdot [x_k, h_{k-1}] + b_g) \\ c_k &= f_k \odot c_{k-1} + i_k \odot g_k \\ o_k &= \sigma(W_o \cdot [x_k, h_{k-1}] + b_o) \\ h_k &= o_k \odot \tanh(c_k) \end{aligned}$$

where  $\sigma$  is sigmoid non-linearity function,  $\tanh$  is hyperbolic tangent non-linearity function,  $W$  terms denote corresponding weight matrices of LSTM,  $b$  terms denote bias vectors,  $i_k$ ,  $f_k$ ,  $g_k$ ,  $c_k$  and  $o_k$  are input gate, forget gate, input modulation gate, the cell state and output gate at time  $k$ , respectively, and  $\odot$  is the Hadamard product [25].

## IV. EXPERIMENTS AND RESULTS

In this section, we present MMWave radar and spatial and temporal sensor calibration approaches employed in our experiments. Moreover, details of dataset creation procedure and evaluation results with quantitative and qualitative metrics are given.



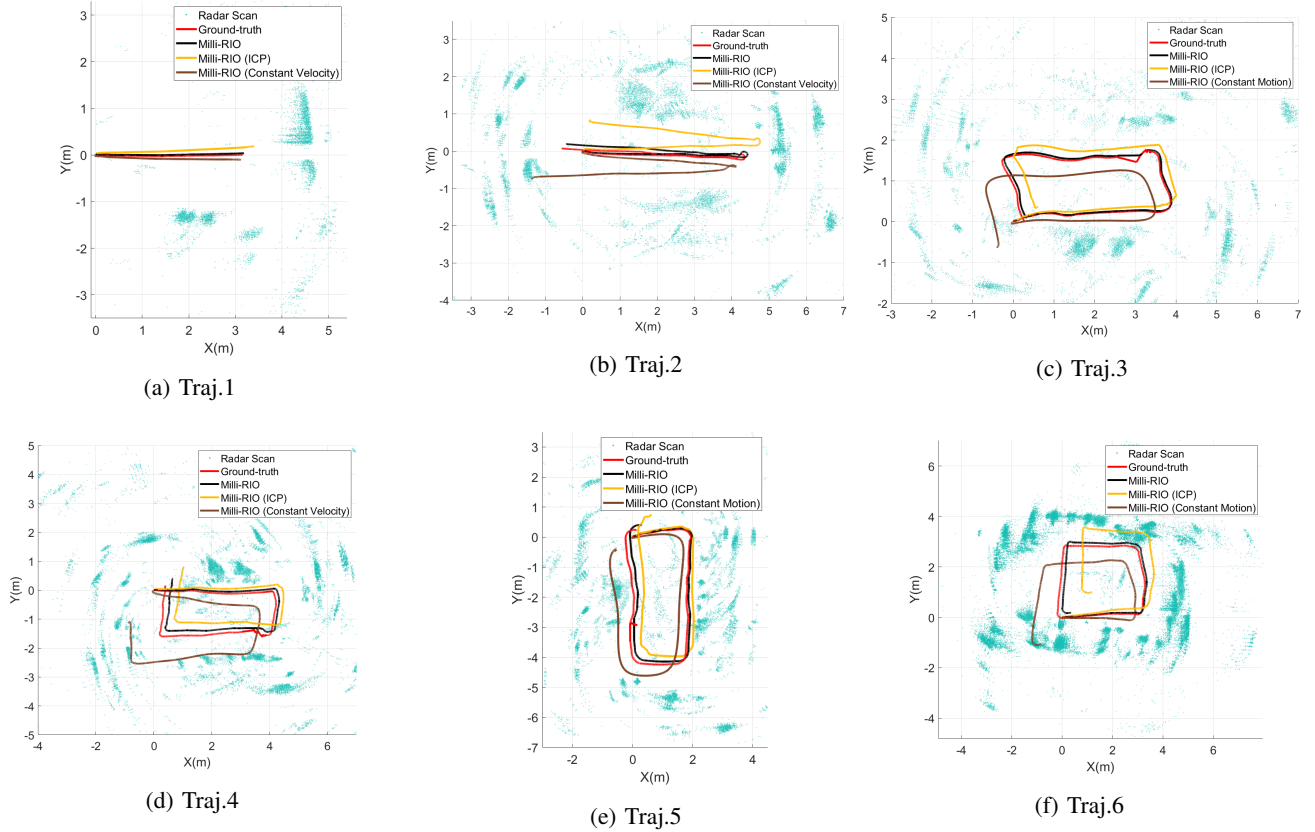


Fig. 4: **Trajectory estimation results.** Sample trajectories of the moving radar and the corresponding trajectories estimated by the proposed method. The trajectories include various type of motions such as linear and circular motions, sharp turns, and smooth transitions etc.

#### A. Millimetre-Wave Scanning Radar

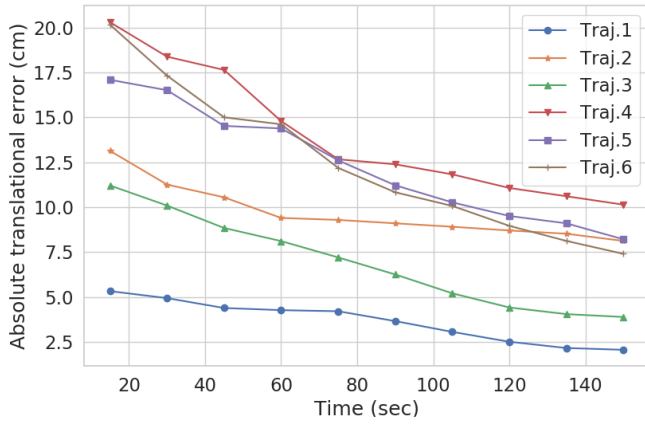
We employ a Texas Instruments AWR1843BOOST model, short range MM-Wave scanning radar, which is shown in Fig. 3. This radar is attached to a mobile agent and it continuously transmits and receives frequency modulated radio waves within the maximum angular field of view. The power received by the antenna corresponds to a position in the environment indicating the reflectivity, size, and orientation of an object at that position. The device is an integrated single-chip MMWave sensor based on FMCW radar technology capable of operation in the 76 to 81 GHz band with up to 4 GHz continuous chirp. The AWR1843 includes a monolithic implementation of a 2 transmit (TX), 4 receive (RX) radio frequency (RF) components system with built-in PLL and A2D converters. The device also integrates a DSP subsystem, which contains TI C674x DSP for the Radio Frequency (RF) signal processing. The device includes an ARM R4F-based processor subsystem, which is responsible for front-end configuration, control, and calibration. Single-chip MMWave radar is a promising solution for low-power, self-monitored, ultra-accurate radar systems. The range accuracy of the MM-Wave radar is high and the measurement results are stable. The advantages of MMWave radar provides accurate range-measurement, gathers readings at close range, and operates

at low peak power. Sidelobes radiation sent in unintended directions and multipath reflections that occur when a wave encounters additional reflection points before returning to the receiver antenna cause noise and non-existing object locations in the scan data. The relative motion calculation by the Doppler effect and 3D information compression introduce errors in range measurement. Other issues causing noise in the data include phase jitter, saturation, and atmospheric attenuation.

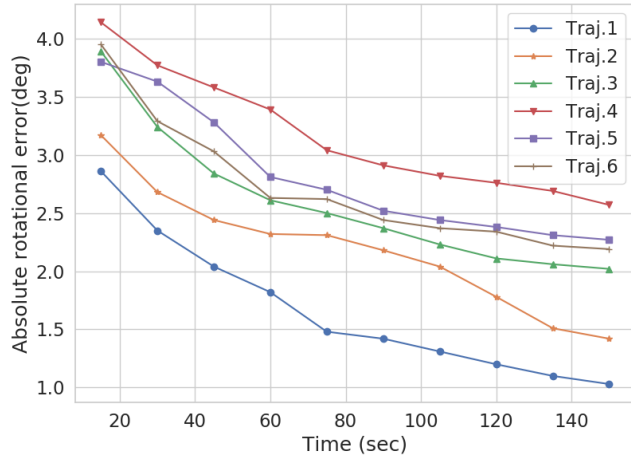
The tuning range of the device is 76 – 81 GHz with a minimum range resolution of 4 cm and angular accuracy of  $1^\circ$ . The radar is placed on the roof of a mobile platform with axis of antenna perpendicular to the motion plane (see Fig. 3). The platform is typically moved between 0.40 and 0.60 m/s; when turning, up to 0.40 rad/s. The robot is driven through a typical lab environment where it is tracked with a VICON tracking system that provides ground-truth with sub-millimetre accuracy.

#### B. Spatial and Temporal Sensor Calibration

To calibrate IMU sensor and MMWave radar with respect to VICON motion tracking system, we first recorded a sequence with an '∞'-loop. Then, we registered radar scans using the NDT algorithm. To obtain an accurate point cloud registration, we placed strongly reflective markers in the



(a) Absolute translational error.



(b) Absolute rotational error.

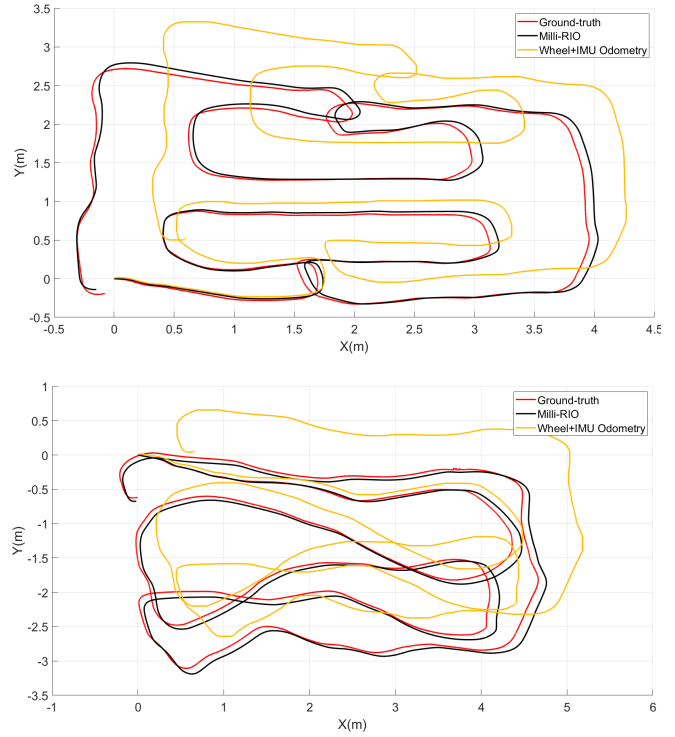
**Fig. 5: The change of error in time for trajectories in Fig. 4.** Absolute trajectory errors decrease over time because Milli-RIO registers the current point cloud to the accumulated point cloud of the environment, proving the effectiveness of the proposed global alignment approach.

environment. Given pairs of IMU-VICON and radar-VICON trajectories, this problem corresponds to the well-known hand-eye calibration. We performed hand-eye calibration using standard approach explained in [27].

In order to synchronize the sensors, we used the timestamps of MMWave radar that has a lower frame-per-second rate (30 FPS) as a reference. We collected the information with the closest IMU timestamp to the radar timestamp for a particular frame, resulting in a worst-case time difference of 5 ms between IMU and radar data package. All timestamps were recorded on our host computer using ROS [28] system clock.

### C. Assessment of Odometry Performance

The dataset is collected in an office environment, including various types of translational and rotational motions. Such a detailed dataset enables us to evaluate if the proposed method is biased towards certain motion types. The total path length



**Fig. 6: Comparative odometry estimation performance.** Milli-RIO is prone to accumulating drift, whereas Turtlebot odometry based on the gyro and motor encoders [26] rapidly deviates from the ground-truth.

of the trajectories in Fig. 4 is 61.38m, which is recorded in a total time of 913sec. The trajectories in Fig. 6 have longer lengths with a total path length of 53.81m and duration of 628sec. The trajectories contain both sharp and smooth transitions to evaluate robustness of the proposed approach. Figure 4 illustrates sample trajectories of the mobile agent and the corresponding estimated trajectories by the proposed radar based odometry system. Figure 4 shows the overall point cloud registration performed by the proposed approach. Figure 5 displays both the translational and rotational ATE (absolute trajectory error) in *cm* and *deg*, respectively. The translational and rotational error decreases over time because Milli-RIO registers the current point cloud to the accumulated point cloud of the environment. Such a global alignment approach is more effective than local alignment due to better data association. Table I and II quantitatively shows ATE results in terms of mean, median, standard deviation and root mean square error (RMSE). Figure 6 compares the odometry performance of Milli-RIO with Turtlebot odometry based on the gyro and motor encoders [26], which shows the proposed method is prone to accumulating drift even on complex trajectories that contain sharp rotations and long translations.

As shown in Fig.4 and Table I, II, high concentration in the clustered regions of point clouds results in lower trajectory error. Similarly, scattered point clouds reduce the performance of point cloud registration and, thus, cause higher trajectory errors. One can see the successful pose

Error (cm)	Traj.1	Traj.2	Traj.3	Traj.4	Traj.5	Traj.6
Mean	2.57	9.06	4.81	12.39	10.96	10.28
Median	2.54	9.09	4.67	12.27	9.06	10.29
Std.	1.49	5.28	2.93	7.59	5.51	8.26
RMSE	2.97	10.48	5.63	10.22	10.98	13.50

TABLE I: Translational ATE (absolute trajectory error) results for MILLI-RIO.

Error (deg)	Traj.1	Traj.2	Traj.3	Traj.4	Traj.5	Traj.6
Mean	1.38	1.93	2.43	2.87	2.60	2.37
Median	1.25	1.76	2.27	2.72	2.54	2.69
Std.	1.01	1.33	1.86	2.17	1.79	2.15
RMSE	1.49	1.60	2.09	2.35	2.20	2.55

TABLE II: Rotational ATE (absolute trajectory error) results for MILLI-RIO.

estimation of the user-defined trajectories with minimal deviations on the order of centimetre scale in both Fig. 4 and Table I. Deviations from the desired trajectories are caused by unstable points in radar scans due to the signal attenuation caused by radiation through a different medium. A video demonstration is available online<sup>1</sup>.

## V. CONCLUSION

In this paper, we introduced an accurate and robust radar-IMU motion estimation system that achieves centimetre accuracy and demonstrates the effectiveness of MMWave radars for indoor localization. As an onboard low-cost radar sensor, the successful implementation of MMWave radar odometry improves the reliability and versatility of mobile systems. Our method stands out because it is not only dependable and accurate, but also straightforward and intuitive without a need for hand-engineered motion model. In future, we plan to incorporate a robust 3D map reconstruction module into the pipeline.

## REFERENCES

- [1] R. McMillan, "Terahertz imaging, millimeter-wave radar," in *Advances in sensing with security applications*. Springer, 2006, pp. 243–268.
- [2] G. L. Charvat, *Small and short-range radar systems*. CRC Press, 2014.
- [3] J. Hasch, E. Topak, R. Schnabel, T. Zwick, R. Weigel, and C. Waldschmidt, "Millimeter-wave technology for automotive radar sensors in the 77 ghz frequency band," *IEEE Transactions on Microwave Theory and Techniques*, vol. 60, no. 3, pp. 845–860, 2012.
- [4] P. Zhao, C. X. Lu, J. Wang, C. Chen, W. Wang, N. Trigoni, and A. Markham, "mid: Tracking and identifying people with millimeter wave radar," in *2019 15th International Conference on Distributed Computing in Sensor Systems (DCOSS)*. IEEE, 2019, pp. 33–40.
- [5] G. Brooker, D. Johnson, J. Underwood, J. Martinez, and L. Xuan, "Using the polarization of millimeter-wave radar as a navigation aid," *Journal of Field Robotics*, vol. 32, no. 1, pp. 3–19, 2015.
- [6] M. Jaud, R. Rouveure, L. M. Arvis, P. Faure, and M. Monod, "Boat borne radar mapping versus aerial photogrammetry and mobile laser scanning applied to river gorge monitoring," *Open Journal of Remote Sensing and Positioning*, vol. 1, no. 1, pp. 48–63, 2014.
- [7] T. Wang, N. Zheng, J. Xin, and Z. Ma, "Integrating millimeter wave radar with a monocular vision sensor for on-road obstacle detection applications," *Sensors*, vol. 11, no. 9, pp. 8992–9008, 2011.
- [8] F. Schuster, C. G. Keller, M. Rapp, M. Haueis, and C. Curio, "Landmark based radar slam using graph optimization," in *Intelligent Transportation Systems (ITSC), 2016 IEEE 19th International Conference on*. IEEE, 2016, pp. 2559–2564.
- [9] K. Werber, J. Klappstein, J. Dickmann, and C. Waldschmidt, "Interesting areas in radar gridmaps for vehicle self-localization," in *Microwaves for Intelligent Mobility (ICMIM), 2016 IEEE MTT-S International Conference on*. IEEE, 2016, pp. 1–4.
- [10] D. Vivet, P. Checchin, and R. Chapuis, "Localization and mapping using only a rotating fmcw radar sensor," *Sensors*, vol. 13, no. 4, pp. 4527–4552, 2013.
- [11] E. Jose and M. D. Adams, "An augmented state slam formulation for multiple line-of-sight features with millimetre wave radar," in *Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on*. IEEE, 2005, pp. 3087–3092.
- [12] J. W. Marck, A. Mohamoud, E. vd Houwen, and R. van Heijster, "Indoor radar slam a radar application for vision and gps denied environments," in *Radar Conference (EuRAD), 2013 European*. IEEE, 2013, pp. 471–474.
- [13] F. Schuster, M. Wörner, C. G. Keller, M. Haueis, and C. Curio, "Robust localization based on radar signal clustering," in *Intelligent Vehicles Symposium (IV), 2016 IEEE*. IEEE, 2016, pp. 839–844.
- [14] T. Deissler and J. Thielecke, "Uwb slam with rao-blackwellized monte carlo data association," in *Indoor Positioning and Indoor Navigation (IPIN), 2010 International Conference on*. IEEE, 2010, pp. 1–5.
- [15] E. Ward and J. Folkesson, "Vehicle localization with low cost radar sensors," in *Intelligent Vehicles Symposium (IV), 2016 IEEE*. Institute of Electrical and Electronics Engineers (IEEE), 2016.
- [16] M. Chandran and P. Newman, "Motion estimation from map quality with millimeter wave radar," in *Intelligent Robots and Systems, 2006 IEEE/RSJ International Conference on*. IEEE, 2006, pp. 808–813.
- [17] P. Checchin, F. Gérossier, C. Blanc, R. Chapuis, and L. Trassoudaine, "Radar scan matching slam using the fourier-mellin transform," in *Field and Service Robotics*. Springer, 2010, pp. 151–161.
- [18] D. Kellner, M. Barjenbruch, J. Klappstein, J. Dickmann, and K. Dietmayer, "Instantaneous ego-motion estimation using multiple doppler radars," in *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, 2014, pp. 1592–1597.
- [19] Z. Peng, L. Ran, and C. Li, "A k-band portable fmcw radar with beamforming array for short-range localization and vital-doppler targets discrimination," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 9, pp. 3443–3452, 2017.
- [20] 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)*. IEEE, 2018, pp. 1–8.
- [21] M. Magnusson, A. Lilienthal, and T. Duckett, "Scan registration for autonomous mining vehicles using 3d-ndt," *Journal of Field Robotics*, vol. 24, no. 10, pp. 803–827, 2007.
- [22] M. Magnusson, A. Nuchter, C. Lorken, A. J. Lilienthal, and J. Hertzberg, "Evaluation of 3d registration reliability and speed-a comparison of icp and ndt," in *2009 IEEE International Conference on Robotics and Automation*. IEEE, 2009, pp. 3907–3912.
- [23] E. A. Wan and R. Van Der Merwe, "The unscented kalman filter for nonlinear estimation," in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*. IEEE, 2000, pp. 153–158.
- [24] M. Turan, Y. Almalioglu, H. Gilbert, H. Araujo, T. Cemgil, and M. Sitti, "Endosensorfusion: Particle filtering-based multi-sensory data fusion with switching state-space model for endoscopic capsule robots," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 1–8.
- [25] F. A. Gers, J. Schmidhuber, and F. Cummins, "Learning to forget: Continual prediction with lstm," 1999.
- [26] D. Stonier. Ros kobuki node. [Accessed 9-September-2019]. [Online]. Available: <https://wiki.ros.org/kobukinode>
- [27] R. Horaud and F. Dornaika, "Hand-eye calibration," *The international journal of robotics research*, vol. 14, no. 3, pp. 195–210, 1995.
- [28] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "Ros: an open-source robot operating system," in *ICRA workshop on open source software*, vol. 3, no. 3.2. Kobe, Japan, 2009, p. 5.

<sup>1</sup><https://youtu.be/VoPKrC8st8I>