Milli-RIO: Ego-Motion Estimation with Low-Cost Millimetre-Wave Radar

Yasin Almalioglu, Mehmet Turan, Chris Xiaoxuan Lu, Niki Trigoni, and Andrew Markham

***Abstract*— Robust indoor ego-motion estimation has attracted significant interest in the last decades due to the fast-growing demand for location-based services in indoor environments. Among various solutions, frequency-modulated continuous-wave (FMCW) radar sensors in millimeter-wave (MMWave) spectrum are gaining more prominence due to their intrinsic advantages such as penetration capability and high accuracy. Single-chip low-cost MMWave radar as an emerging technology provides an alternative and complementary solution for robust ego-motion estimation, making it feasible in resource-constrained platforms thanks to low-power consumption and easy system integration. In this paper, we introduce Milli-RIO, an MMWave radar-based solution making use of a single-chip low-cost radar and inertial measurement unit sensor to estimate six-degrees-of-freedom ego-motion of a moving radar. Detailed quantitative and qualitative evaluations prove that the proposed method achieves precisions on the order of few centimeters for indoor localization tasks.**

***Index Terms*— Ego-motion estimation, millimetre-wave radar, radar odometry, recurrent neural networks**

I. Introduction

ROBUST ego-motion estimation for indoor environments has a variety of real-world applications ranging from emergency evacuation to indoor robotics, and remains a challenging task. In the last decades, with the advent of integrated small and power-efficient mobile sensors, various technologies have been adapted to this domain to investigate robust solutions. A significant limitation of classical sensors such as vision or laser is that they are ineffective in visually degraded environments, e.g. glare, smoke and darkness [1].

Radars provide robust and reliable perceptual information of the environment, which are immune to visual degradation. Although radar systems were only used in the military area due to their bulky sizes and high costs in the early years, radar systems have been miniaturized and integrated onto printed circuit boards over recent decades thanks to the advance of high frequency integrated circuits. With the recent advances in the integrated circuit and packaging technologies, it is even possible to integrate a frequency-modulated continuous-wave (FMCW) radar system operating at a higher frequency millimeter-wave (MMWave) band (77 GHz) into a single chip with antenna-on-chip technologies. Despite inherent issues such as high path loss, higher operation frequency and smaller wavelength do not only improve the sensitivity and resolution of radar systems but also make radar systems further compact, extending the range of applications from military to commercial areas. Low-cost MMWave radars thus provide a viable alternative (or complementary) solution for robust indoor ego-motion estimation to overcome the shortcomings of optical sensors (see Fig. 1 for an overview). Figure 2 shows examples of the commercially available solutions used for ego-motion estimation, comparing low-cost radar in terms of cost, weight, energy consumption, field-of-view (FoV) and the detection density.

This work is funded in part by the NIST under Grant 70NANB17H185.
Yasin Almalioglu, Chris Xiaoxuan Lu, Niki Trigoni, and Andrew Markham are with the Computer Science Department, The University of Oxford, UK (e-mail: {yasin.almalioglu, xiaoxuan.lu, niki.trigoni, andrew.markham}@cs.ox.ac.uk).
Mehmet Turan is with the Institute of Biomedical Engineering, Bogazici University, Turkey (e-mail: mehmet.turan@boun.edu.tr).

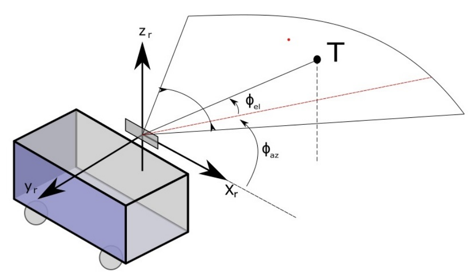


Fig. 1: **Millimetre-Wave (MMWave) radar system attached on top of a moving agent.** MMWave radar is a special class of radar technology that uses millimetre wavelength radio frequency (RF) signals. Our MMWave system operates at 7681 GHz spectrum, resulting in an ability to detect movements on the order of few centimeters.

MMWave radar is extensively used in several domains such as the military (air and maritime surveillance, missile guidance, etc.), civil aviation (approach radar, surface movement radar) or remote sensing (planetary observation) [2]. In recent years, several research groups have proposed MMWave radars as a solution for various mobile robot tasks such as navigation, localization and mapping. 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] and 2D/3D simultaneous localization and mapping (SLAM) [6]. MMWave

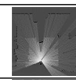
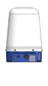
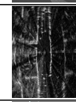
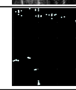
		Cost ($)	Weight (kg)	Power (W)	Field of View (º)	Point Scans
	LIDAR (Velodyne VLP-16)	8,800	0.83	8	360	
	Mechanical Radar (Navtech CTS-350)	Customized Price	6	24	360	
	Single-chip Radar (TI AWR1843)	60 (chip) 300 (dev-kit)	0.03	2	120	

Fig. 2: **Sensor comparison.** Comparison of LIDAR, mechanical radar and single-chip radar, showing the features of a major commercial model for each sensor. Notably, compared with a LIDAR and a mechanical radar used in [8], single-chip radar is much cheaper and lighter, but only provides few points within the field-of-view. White points in the scans correspond to detected objects.

radars are also fused with visual sensors for obstacle detection and map reconstruction, combining the robust depth detection ability of the radar in severe environmental conditions with a high spatial resolution of the visual sensors [7]. However, these solutions involve bulky radar systems that provide dense measurements at the cost of increased physical size, power consumption and price of the system. Thus, ego-motion estimation methods designed explicitly for single-chip, low-cost FMCW MMWave radars are needed to fully utilize the features of portable radar for indoor location-based services.

In this paper, we propose Milli-RIO, an ego-motion estimation method based on single-chip low-cost MMWave radar, which is complemented by an 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 a single-chip low-cost MMWave radar sensor, making it effective for indoor applications in terms of size, cost and energy consumption.
- We propose a new point association technique to match the sparse measurements of low-cost MMWave radar.
- 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 low-cost 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

Radar typically has a wider, taller beam than the light detection and ranging (LIDAR) sensor, which makes scanning large volumes easier but results in lower bearing resolution and cluttered measurements. The longer wavelength of radars causes the radar echo to be reflected off multiple surfaces (such as the ground or walls) on its return trip to the antenna, known as the multipath effect. This effect delays the return of the signal and creates false targets further away than the real one, which is even more challenging in indoor environments due to walls, ceiling and floor reflections [9]. Thus, radar ego-motion estimation systems must be robust to clutters and false-positives, and it must demonstrate high precision despite low-resolution data.

Feature extraction is a fundamental task in radar ego-motion estimation systems. The traditional visual localization techniques such as amplitude grid maps are investigated in literature [10], [11]. [10] uses the amplitude grid maps to transform the radar scans into grayscale images and applies SIFT and FAST feature extractions. [11] studies the grid maps to find continuous areas using DBSCAN, MSER, and the connected components. The radar-specific solutions utilize data distortion, which is used as sources of information to estimate vehicle displacement [12]. Another technique exploits spatiotemporal continuity in radar scans inferring the appearance of landmarks by determining the radar noise characteristics [13]. In 2D radar scan processing, the accurate range information calculated with the highest power return per azimuth eliminates the need for a filter, which can potentially discard relevant information [14]. To combine visual and radar sensors, [10] pairs radar and visual landmarks with similar feature descriptors. Vision-radar fusion approaches use radar occupancy grids to associate both sensor measurements [15]. 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 data. Multi-sensor fusion techniques provide an alternative to feature-based radar odometry. The fusion methods use odometry information from additional sensors to transform the incoming radar landmark point cloud and register it to an existing landmark map. They usually make use of nearest neighbor point matching [15] and Monte Carlo methods to derive a solution from probabilistic weights [16]. The relative motion is estimated using the data association between the radar point cloud and map, which is then fused to the first 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 ego-motion estimation systems, the extracted features are processed in the data association step, which is frequently achieved by a scan matching algorithm that tracks shared features across consecutive radar scans. The iterative closest point (ICP) approach is typically used for scan matching to iteratively align the radar point clouds until the predefined termination criteria are met [17], which is too sensitive to outliers. In [18], the researchers developed a quantitative function describing the quality of the map created by superimposing radar point clouds according to the unknown motion parameters. 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 [12]. 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 [19]. The Doppler radar

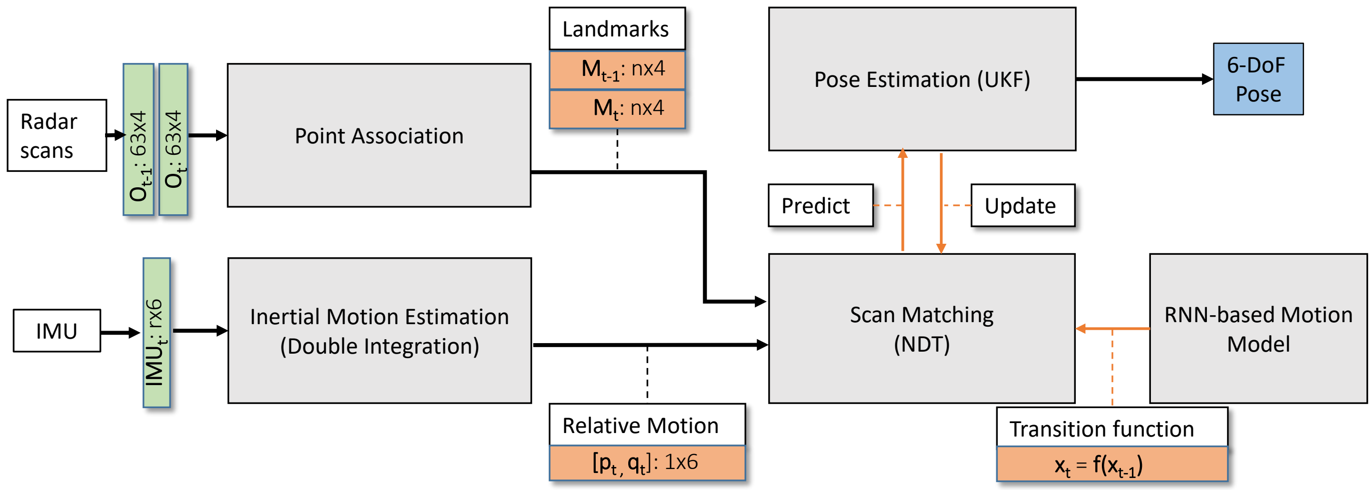


Fig. 3: **Ego-motion estimation workflow.** Raw MMWave radar point clouds are processed in the point association module to extract n landmarks, where n is the number of landmarks determined by the cost function. Scan matching module registers the landmarks using NDT scan matching algorithm, which uses a RNN-based transition model. In parallel, relative ego-motion is estimated from IMU readings using our inertial navigation system. Radar and IMU estimations are fused in the real-time pose estimation module using a UKF to regress the final 6 DoF pose values. Green, orange, gray and blue boxes represent inputs to the system, intermediate estimations, processing units, and outputs, respectively.

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 [20]. Both methods are hampered by heavy preprocessing. [8] proposes a data association method that requires power-range spectra and the extracted landmarks to match similar geometries within two radar scans. The technique requires dense radar scans that already contain visible shapes and patterns, and is not practical for sparse radar measurements.

The present works above for radar-based ego-motion estimation are developed for mechanically rotating radars, which already provide a dense and full field-of-view (FoV) of the environment. The mechanical radars are bulky, expensive and power-hungry sensors, which is incompatible with the requirements of portable indoor location-based services [21] (see Table 2). For indoor navigation and mapping, single-chip radars operating at MMWave frequencies have the following advantages over the other radars: a) reduction in size and mass of the radar allows developing small ultra-lightweight radars suitable for portable indoor systems, and b) millimeter wavelength allows detection of relatively smooth wall surfaces at oblique angles of incidence. However, challenges of single-chip radars such as sparse measurement and limited FoV require specific solutions. In this paper, we present a novel and robust motion estimation approach for indoor localization tasks based single-chip MMWave radar, complemented by the IMU sensor to eliminate deficiencies of both sensors such as biases in IMU output, noises and sparse measurements in radar scans.

III. Millimetre-Wave Radar Based Ego-Motion Estimation

Frequency modulated continuous wave (FMCW) MMWave radar 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 a 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 describe the principles of MMWave radar and present the proposed MMWave radar-based point association and ego-motion estimation algorithms. Moreover, we explain the details of the RNN-based motion model used in the joint MMWave radar-IMU ego-motion estimation.

A. Principles of MMWave Radar

FMCW MMWave radar uses a linear chirp or swept frequency transmission that is characterized by a bandwidth B, start frequency f_c and duration T_c. A mixer in the radar front-end computes the frequency difference between the transmitter and the receiver. the distance between the object and the radar is calculated from an Intermediate Frequency (IF) as:

$$d = \frac{f_{IF}c}{2S}, \tag{1}$$

where c is the speed of light, f_{IF} is the frequency of the IF signal, and $S = B/T_c$ is the frequency slope of the chirp. Each peak of the FFT result on the IF signal represents an obstacle at a corresponding distance.

The angle of arrival (AoA) is estimated using the slight difference in phase of the received signals and emitted chirp signal. For a pair of antennas, AoA is calculated as:

$$\theta = \sin^{-1}\left(\frac{\lambda\omega}{2\pi d}\right) \tag{2}$$

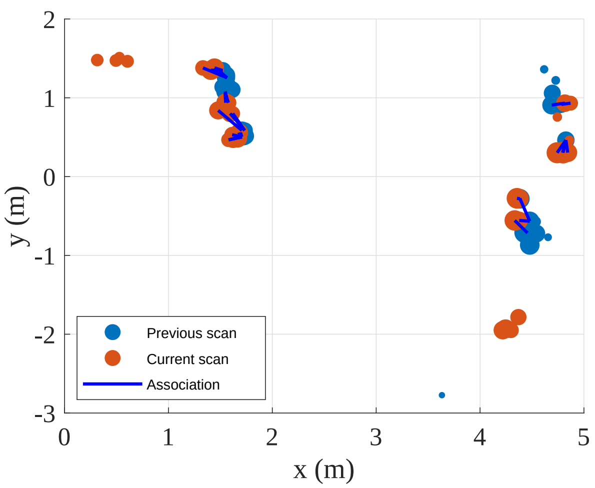


Fig. 4: **Example of point association for MMWave radar scans.** The marker sizes are proportional to the power intensity values. The proposed point association method finds the best matches, eliminating the low-intensity points and the ghost objects. It also penalizes the objects in the current scan that have no counter-part in the previous scan. The radar moves at approximately 0.50 m/s. Positions are relative to radar position.

where ω is the phase difference nad γ is the wavelength. The average result from the receiver pairs gives the final AoA, which decreases with $|\theta|$.

B. Point Association

Milli-RIO achieves robust point correspondences across consecutive scans using high-level information in the radar output. The proposed point association technique seeks to find the largest subset of points that have the same landmark descriptors.

Given two consecutive sets of detected radar points $\mathcal{O}^t = \{\mathbf{o}_1^t, \mathbf{o}_2^t, \ldots, \mathbf{o}_N^t\}$ at time t, and $\mathcal{O}^{t-1} = \{\mathbf{o}_1^{t-1}, \mathbf{o}_2^{t-1}, \ldots, \mathbf{o}_M^{t-1}\}$ at time $t-1$ in Cartesian coordinates, we aim to estimate the rigid-body transformation represented in homogeneous coordinates between the current and the previous scans. This problem of rigid transformation estimation between two sets of points is well researched, especially in the robotics and computer vision field, where it is often solved with the iterative closest point (ICP) algorithm [22]. However, points detected by the radar are not stable over time due to the noise and stronger reflectors within the FoV. Thus, a point appearing in the previous scan may not be present at the current scan. In addition, the number of detected points is low (lower than 100 points per frame), which is typically around 1000 times less than a full FoV scan of a mechanical radar or LIDAR [23], [24]. ICP performs poorly on this type of data due to these challenges. We propose a new technique in this paper, in which point association and motion estimation are performed separately using radar specific features.

The proposed technique performs point association using not only individual unary landmark descriptors but also high-level descriptors between landmarks such as signal intensity and point cloud displacement. This approach reduces the likelihood of an individual point having the same set of descriptors as another. Moreover, the signal intensity is not dependent on the exact position and orientation of the point cloud, making large disparities in pose inconsequential. In addition, the proposed point association algorithm is not constrained to have a good initial estimate of the relative pose, enabling feature extraction from point clouds captured at arbitrary times without any *a priori* map representation. These properties provide reliable matches for our sparse landmark sets.

We pose the problem of point association as a linear sum assignment problem. Given two point sets $\mathcal{O}^t$ and $\mathcal{O}^{t-1}$, the objective is to find a complete assignment while minimizing a cost:

$$\min \sum_i \sum_j D_{i,j} A_{i,j} \tag{2}$$

where $D_{i,j}$ is the score of matching point i in $\mathcal{O}^t$ with point j in $\mathcal{O}^{t-1}$ and A is a binary matrix such that $A_{i,j} = 1$ if point $\mathbf{o}_i^t$ is matched with $\mathbf{o}_j^{t-1}$. We find the optimal alignment using the Munkres algorithm [25].

We construct the similarity score matrix $\mathbf{D}$ as follows:

$$D_{i,j} = \begin{cases} \frac{1}{1+dst(\mathbf{o_i^t}, \mathbf{o_j^{t-1}})} & \text{if } \varphi(\mathbf{o_i^t}, \mathbf{o_j^{t-1}}) == True \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where the unary landmark descriptor $dst(\mathbf{o_i^c}, \mathbf{o_j^P})$ is the squared Euclidean distance between the points $\mathbf{o_i^t}$ and $\mathbf{o_j^{t-1}}$, measuring that correctly identified landmarks have the same spatial and dynamic features in any two radar scans. $\mathbf{D}$ is diagonally dominant for the optimal set of matches, making the overall pair-wise similarity maximum. $\varphi(\mathbf{o_i^t}, \mathbf{o_j^{t-1}})$ is a policy function that defines the high-level landmark descriptors to eliminate improbable associations caused by ghost reflections and noise. $\varphi(\mathbf{o_i^t}, \mathbf{o_j^{t-1}})$ returns $False$ if two points cannot be associated. Given the assumptions that (a) the radar moves forward with a given maximum speed, (b) the lateral translation is low, and (c) the landmarks have minimum signal intensity value, we define $\varphi(\mathbf{o_i^t}, \mathbf{o_j^{t-1}})$ as follows:

$$\varphi(\mathbf{o_i^t}, \mathbf{o_j^{t-1}}) = \begin{cases} False & \text{if } dst(\mathbf{o_i^t}, \mathbf{o_j^{t-1}}) > MaxValue \\ False & \text{if } o_j^t|_x - o_i^{t-1}|_x < 0 \\ False & \text{if } (o_j^t|_y - o_i^{t-1}|_y)^2 > MaxLateral \\ False & \text{if } I(o_j^t) < MinIntensity \\ True & \text{otherwise} \end{cases} \tag{3}$$

where $o_j^p|_x$ and $o_j^p|_y$ respectively denote the longitudinal component (along the x axis) and the lateral component (along the y axis) of $\mathbf{o_j^P}$. $MaxValue$ denotes the maximum distance, and $MaxLateral$ the maximum lateral displacement of the ego vehicle between two iterations. $I(o_j^t)$ denotes the signal intensity of point o_j^t, which is conditioned on the required minimum signal intensity $MinIntensity$. Note that $\varphi(\mathbf{o_i^t}, \mathbf{o_j^{t-1}})$ could also be defined with additional information, for example using the previous results of the motion estimation or using information from the odometers.

The greedy method iteratively collects satisfactory matches into the set $\mathcal{M}$. On each iteration, it evaluates the remaining valid matches and calculates the score values. The remaining point pairs (i, j) with a score $D_{i,j} > d_\tau$ are collected in a set of matches $\mathcal{M}$ for a given score threshold d_τ. An example of point association given in Fig. 4 shows that the point association is coherent despite the noisy detections, where the ego speed is around 0.50 m/s.

C. Relative Motion Estimation

In the relative motion estimation module of the proposed system, we estimate the sensor trajectory by iteratively applying the normal distributions transform (NDT) scan matching technique [26] to find the rigid body motion given the set of corresponding points $\mathcal{M}$. NDT is shown to have a better performance than other scan matching algorithms, such as ICP, in terms of both reliability and processing speed [27]. 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. To deal with this problem, we integrate angular velocity data provided by the IMU sensor to the NDT scan matching algorithm using UKF [28]. The pipeline of our method is demonstrated in Fig. 3.

We define the sensor state to be:

$$\mathbf{x}_t = [\mathbf{p}_t, \mathbf{q}_t, \mathbf{v}_t, \mathbf{b}_t^a]^T, \tag{4}$$

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 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:

$$\mathbf{x}_t = [\mathbf{p}_{t-1} + f(\mathbf{x}_{t-1}), \mathbf{q}_{t-1}.\Delta\mathbf{q}_t, \mathbf{v}_{t-1}, \mathbf{b}_{t-1}^a]^T, \tag{5}$$

where $\Delta\mathbf{q}_t$ is the rotation during $t-1$ and t. The rotation is given by:

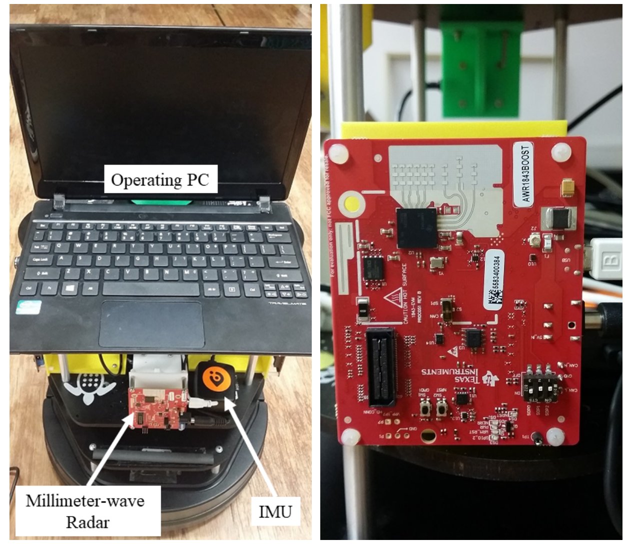
$$\Delta\mathbf{q}_t = \left[1, \frac{\Delta t}{2}\mathbf{a}_t^{x'}, \frac{\Delta t}{2}\mathbf{a}_t^{y'}, \frac{\Delta t}{2}\mathbf{a}_t^{z'}\right]^T, \tag{6}$$

where $\mathbf{a}_t^{'} = \mathbf{a}_t - \mathbf{b}_{t-1}^a$ is the bias-compensated angular velocity.

The motion estimation module uses Eq. 5 and UKF to predict the sensor pose, the estimated $\mathbf{x}_t$ and $\mathbf{q}_t$ being the initial guess of the sensor pose. We iteratively apply NDT on the set $\mathcal{M}$ to register the observed point cloud into the global map. 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 of UKF is defined as:

$$\mathbf{z}_t = [\mathbf{p}_t, \mathbf{q}_t]^T. \tag{7}$$

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, but the estimation performance deteriorates due to strong acceleration noise and constant bias. Thus, we omit the acceleration update step from Milli-RIO.



(a) (b)

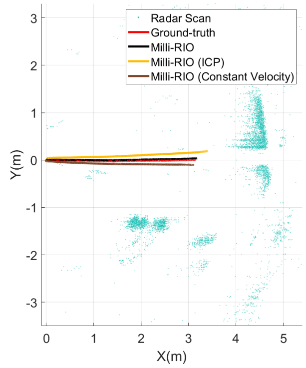
Fig. 5: **Experimental setup.** a) Turtlebot2 data collection platform. b) TI AWR1843BOOST model, short-range millimetre-wave radar employed in the experiments.

D. RNN-based Motion Model

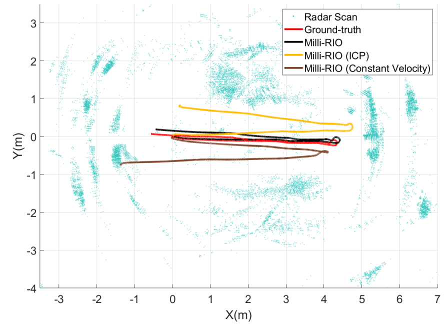
Existing motion estimation methods based on traditional filters have limitations for nonlinear dynamic systems. UKF accommodates a wide variety of dynamic models, allowing for highly complex dynamics in the state variables given an accurate motion model.

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, the motion estimation problem is mainly concerned with the 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 the 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. LSTM is a specific implementation of RNN to avoid the vanishing gradient problem, enabling the 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.

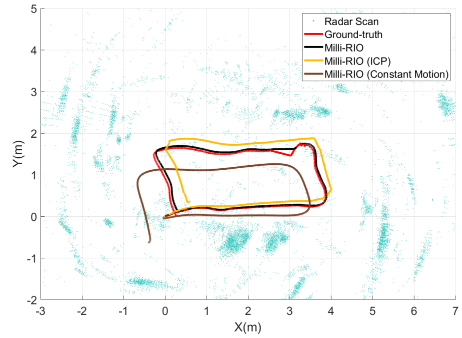
The pose estimation of the current time step benefits from information encapsulated in previous time steps. Thus, LSTM is a suitable RNN implementation to formulate the state transition function f in Eq. 5 [29]. Our implementation is based on bi-directional LSTM (bi-LSTM) that has a memory function not only for the forward sequences but also for the



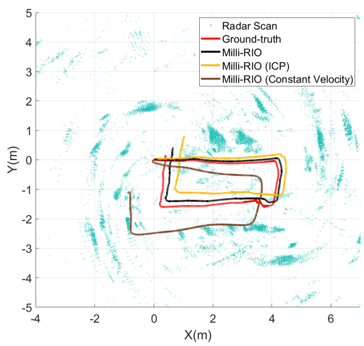
(a) Traj.1



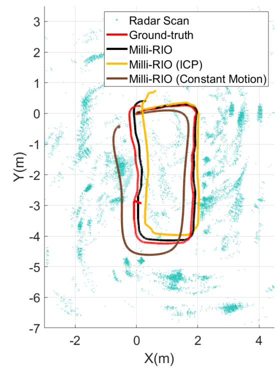
(b) Traj.2



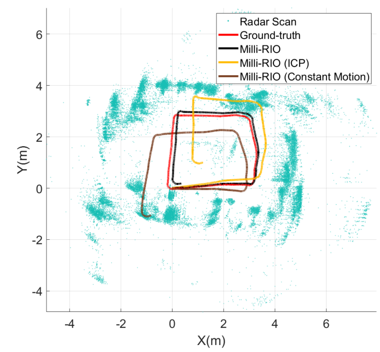
(c) Traj.3



(d) Traj.4



(e) Traj.5



(f) Traj.6

Fig. 6: **Trajectory estimation results.** Sample trajectories of the moving radar estimated by the proposed method are compared against ground-truth and traditional approaches. Milli-RIO (ICP) replaces the normal distributions transform (NDT) scan matching module of Milli-RIO with the iterative closest point (ICP) approach. Milli-RIO (Constant Velocity) replaces the recurrent neural network based motion model of Milli-RIO with a constant velocity model. The proposed Milli-RIO architecture outperforms the traditional approaches. The trajectories include various type of motions such as linear and circular motions, sharp turns, and smooth transitions. Point clouds in blue are the registered radar scans.

backward sequences to fully consider the mutual relationship between the sequences. UKF tracks the 6-DoF pose of moving radar using the transition function modeled by a bi-LSTM network that consists of 256 LSTM nodes in each direction. We recorded various trajectories containing rotational and translational motions to train our bi-LSTM network. We use different trajectories for training and testing to ensure the network captures the motion dynamics and avoids overfitting to the training dataset. We also use batch normalization and dropout layers with a rate of $\alpha = 0.25$ to prevent overfitting. The inputs to the bi-LSTM 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 bi-LSTM learns the non-linear motion model of the mobile radar.

IV. Experiments and Results

In this section, we describe our experimental setup and single-chip MMWave radar configurations. We present the spatial and temporal sensor calibration approach employed in our experiment and the details of the dataset creation procedure. Moreover, we show and discuss the evaluation results with quantitative and qualitative metrics.

A. Low-Cost Millimetre-Wave Scanning Radar

Single-chip MMWave radar is a promising solution for low-power, self-monitored, ultra-accurate radar systems. MMWave radar has several advantages such as it provides accurate range measurements, gathers readings at close range, and operates at low peak power. Sidelobe 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. Other issues causing noise in the data include phase jitter, saturation, and atmospheric attenuation.

We employ a Texas Instruments AWR1843BOOST model, short-range MMWave scanning radar, which is shown in Fig. 5. This radar is attached to a mobile agent, and it continuously transmits and receives frequency modulated radio waves within the maximum angular FoV of 120°. The power peaks 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 two transmit (TX), four 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 RF signal processing unit to generate a point cloud of 63 points per frame. The device includes an ARM R4F-based processor subsystem, which is responsible for front-end configuration, control and calibration. The starting frequency f_c of the device is 76 GHz with 12.5dBm TX power and 15dB RX noise figure. The radar is configured with a bandwidth of 4 GHz, the ramp slope S of 70 MHz/μs, resulting in a range resolution of 4.3cm, maximum unambiguous range of 22.55m, and radial velocity of 2.28m/s with 0.29m/s resolution. The number of samples per chirp is 128 and the frame rate is 20 frames-per-second (fps). The radar is placed on the roof of a mobile platform with the axis of the antenna perpendicular to the motion plane (see Fig. 5). 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-millimeter accuracy.

B. Spatial and Temporal Sensor Calibration

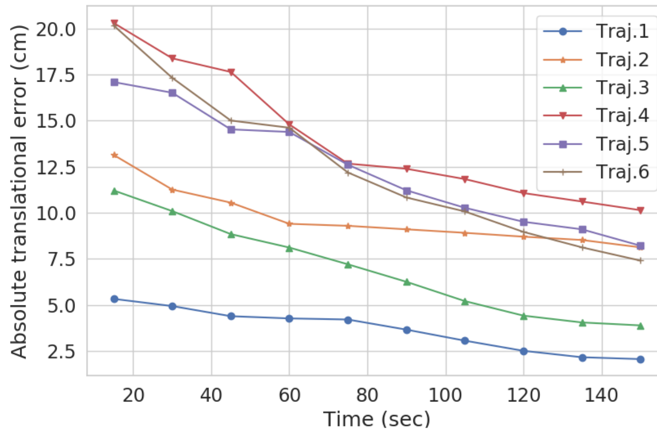
To calibrate the IMU sensor and MMWave radar with respect to the VICON motion tracking system, we first recorded a sequence with an '∞'-loop. Then, we registered radar scans using the NDT algorithm. To obtain accurate point cloud registration, we placed strongly reflective markers in the environment, which are only used during the calibration session and removed during the experiments. 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 the standard approach explained in [31].

In order to synchronize the sensors, we used the timestamps of MMWave radar which has a lower fps rate 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 [32] system clock.

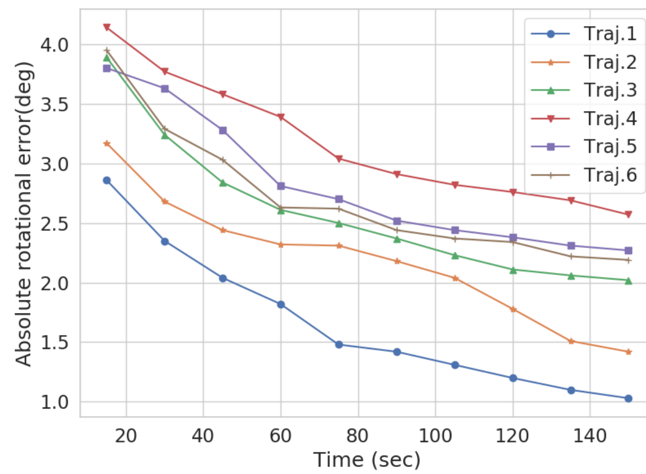
C. Assessment of Odometry Performance

The dataset is collected in typical office environments, 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 of the trajectories in Fig. 6 is 61.38m with average trajectory length of 10.23m, which is recorded in a total time of 913sec. In addition, we record longer trajectories to evaluate the robustness of the proposed method against motion drifts. The trajectories in Fig. 8 have longer average length of 26.90m with a total path length of 53.81m and duration of 628sec. The trajectories contain both sharp and smooth motions to evaluate the robustness of the proposed approach in indoor environments.

Figure 6 illustrates sample trajectories of the moving radar and the corresponding estimated trajectories by the proposed



(a) Absolute translational error.



(b) Absolute rotational error.

Fig. 7: **The change of error in time for trajectories in Fig. 6.** 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.

radar-based odometry system. Figure 6 also depicts the overall point cloud registration performed by the proposed approach. To demonstrate the effectiveness of Milli-RIO architecture, we perform ablation studies, as shown in Fig. 6. We replace the normal distributions transform (NDT) scan matching module of Milli-RIO with the standard ICP approach, called Milli-RIO (ICP). Although Milli-RIO (ICP) estimates the full trajectories without tracking failure, it is prone to motion drift and deviates from the ground-truth. Furthermore, we replace the RNN based motion model of Milli-RIO with the standard constant velocity model in UKF, called Milli-RIO (Constant Velocity). However, Milli-RIO (Constant Velocity) suffers from a significant drift over time. Although it has a slightly better performance on trajectories following a line, it performs poorly on trajectories that contain sharp turns and rotations. The proposed Milli-RIO architecture outperforms traditional approaches. In addition, we removed the radar odometry from the motion estimation pipeline to test the contribution of the

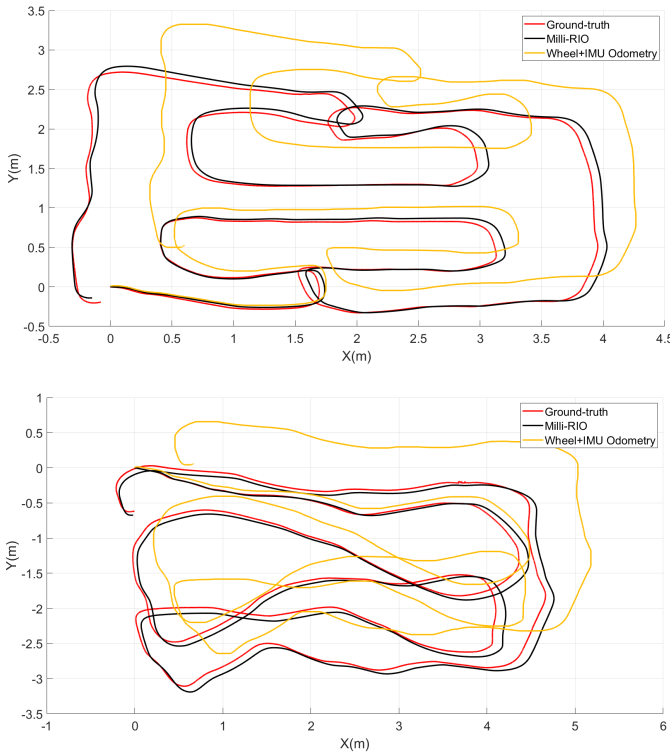


Fig. 8: **Comparative odometry estimation performance.** Milli-RIO is resistant to accumulating drift, whereas Turtlebot odometry based on the gyro and motor encoders [30] rapidly deviates from the ground-truth.

IMU sensor to the overall motion estimation performance. However, the filter rapidly diverges and the motion estimation fails after few iterations due to the accumulated quadratic error in the double integration of unconstrained IMU bias and noise, proving the effectiveness of radar odometry.

In Fig. 7, we display 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).

We further compare Milli-RIO with an off-the-shelf odometry method, fusing wheel odometry with the inertial navigation system. Figure 8 shows the odometry results of Milli-RIO on longer trajectories, comparing with the ground-truth and Turtlebot odometry based on the gyro and motor encoders [30]. The evaluations prove that the proposed method is resistant to accumulating drift even on long indoor trajectories that contain complex motions. A video demonstration of Milli-RIO is available online[1].

[1] https://youtu.be/VoPKrC8st8I

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.

V. Conclusion

In this paper, we introduced an accurate and robust radar-IMU motion estimation system that achieves centimeter 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 a hand-engineered motion model. In the future, we plan to incorporate a robust 3D map reconstruction module into the pipeline.

References

[1] P. Fritsche, B. Zeise, P. Hemme, and B. Wagner, "Fusion of radar, lidar and thermal information for hazard detection in low visibility environments," in *2017 IEEE International Symposium on Safety, Security and Rescue Robotics (SSRR)*. IEEE, 2017, pp. 96–101.

[2] R. McMillan, "Terahertz imaging, millimeter-wave radar," in *Advances in sensing with security applications*. Springer, 2006, pp. 243–268.

[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] 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.

[9] M. Adams and E. Jose, *Robotic navigation and mapping with radar*. Artech House, 2012.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] 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.

[15] 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.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] R. F. Brena, J. P. García-Vázquez, C. E. Galván-Tejada, D. Muñoz-Rodriguez, C. Vargas-Rosales, and J. Fangmeyer, "Evolution of indoor positioning technologies: A survey," *Journal of Sensors*, vol. 2017, 2017.

[22] P. J. Besl, N. D. McKay *et al.*, "A method for registration of 3-d shapes," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 14, no. 2, pp. 239–256, 1992.

[23] P. Fritsche, S. Kueppers, G. Briese, and B. Wagner, "Radar and lidar sensorfusion in low visibility environments." in *ICINCO (2)*, 2016, pp. 30–36.

[24] D. Barnes, M. Gadd, P. Murcutt, P. Newman, and I. Posner, "The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset," *arXiv preprint arXiv:1909.01300*, 2019.

[25] J. Munkres, "Algorithms for the assignment and transportation problems," *Journal of the society for industrial and applied mathematics*, vol. 5, no. 1, pp. 32–38, 1957.

[26] 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.

[27] 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.

[28] 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.

[29] 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.

[30] D. Stonier. Ros kobuki node. [Accessed 9-September-2019]. [Online]. Available: https://wiki.ros.org/kobukinode

[31] R. Horaud and F. Dornaika, "Hand-eye calibration," *The international journal of robotics research*, vol. 14, no. 3, pp. 195–210, 1995.

[32] 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.

Yasin Almalioglu is currently enrolled as a DPhil (PhD) student in computer science at the University of Oxford since January 2018. He received the BSc degree with honors in computer engineering from Bogazici University, Istanbul, Turkey, in 2015. He was a research intern at CERN Geneva, Switzerland and Astroparticle and Neutrino Physics Group at ETH Zurich, Switzerland, in 2013 and 2014, respectively. He was awarded the Engin Arik Fellowship in 2013. He received the MSc degree with high honors in computer engineering from Bogazici University, Istanbul, Turkey, in 2017.

Mehmet Turan received his diploma degree from RWTH Aachen University, Germany in 2012 and his PhD degree from ETH Zurich, Switzerland in 2018. Between 2013-2014, he was a research scientist at UCLA (University of California Los Angeles). Between, 2014-2018 he was a research scientist at Max Planck Institute for Intelligent Systems and between 2018-2019, he was a postdoctoral researcher at the Max Planck Institute for Intelligent Systems. He received the DAAD (German Academic Exchange Service) fellowship between years 20052011 and Max Planck fellowship between 20142019. He has also received Max Planck-ETH Center for Learning Systems fellowship between 20162019. Currently, he is a faculty at the Institute of Biomedical Engineering, Bogazici University, Turkey.

Chris Xiaoxuan Lu is a post-doctoral fellow at University of Oxford, working on EPSRC Programme Grant Project Mobile Robotics: Enabling a Pervasive Technology of the Future. He obtained his Ph.D. degree in Computer Science at Oxford University. His Ph.D. study was generously funded by Google DeepMind. Prior to joining Oxford, he received his MEng degree in 2015 from Nanyang Technological University (NTU).

Niki Trigoni is a Professor at the Oxford University Department of Computer Science and a fellow of Kellogg College. She obtained her DPhil at the University of Cambridge (2001), became a postdoctoral researcher at Cornell University (2002-2004), and a Lecturer at Birkbeck College (2004-2007). At Oxford, she is currently Director of the EPSRC Centre for Doctoral Training on Autonomous Intelligent Machines and Systems, a program that combines machine learning, robotics, sensor systems and verification/control. She also leads the Cyber-Physical Systems Group, which is focusing on intelligent and autonomous sensor systems with applications in positioning, healthcare, environmental monitoring and smart cities.

Andrew Markham is an Associate Professor and he works on sensing systems, with applications from wildlife tracking to indoor robotics to checking that bridges are safe. He works in the cyber-physical systems group. He designs novel sensors, investigates new algorithms (increasingly deep and reinforcement learning-based), and applies these innovations to solving new problems. Previously he was an EPSRC Postdoctoral Research Fellow, working on the UnderTracker project. He obtained his Ph.D. from the University of Cape Town, South Africa, in 2008, researching the design and implementation of a wildlife tracking system using heterogeneous wireless sensor networks.