Algorithms for Enhanced Indoor Positioning and Tracking based on a 60-GHz Radar Platform

F. Shamsfakhr*, M. Corrà†, A. Ferrari‡, D. Macii*, L. Palopoli§, D. Fontanelli*
§Dep. of Industrial Engineering, University of Trento, Italy, E–mail:farhad.shamsfakhr@unitn.it
†Tretec S.r.L., Trento, Italy, E–mail: michele.corra@3tec.it
‡VNG Ingegneria Trento, Italy, E–mail: alessandro.ferrari@vngingegneria.it
§Dep. of Information Engineering and Computer Science, University of Trento, Italy, E–mail: luigi.palopoli@unitn.it

***Abstract*—Wireless technologies are increasingly adopted for indoor localisation and tracking. Among them, radars based on 60 GHz phased-array transceivers are becoming a viable and low–cost solution, which delivers high accuracy with a low computational cost while preserving people privacy. In this paper, we propose two techniques for radar-based people tracking and localisation that tackle important effects such as delayed and spurious detection. We also offer a thorough and scientifically sound experimental analysis of the performance of the proposed techniques using an off-the-shelf phased-array device (i.e., the System-on-Chip (SOC) TI IWR6843).**

I. Introduction

The list of applications enabled by effective localisation and tracking of humans and goods is wide and constantly growing [1]. A variety of solutions exists depending on the considered target accuracy and the adopted technology.

Millimeter-wave (mm-wave) RF signals can be used to implement radars with a range of a few meters. Like surveillance cameras or other types of visual sensors [2], [3], mm-wave radars can be deployed in strategic locations of a given environment and used to track target that are not required to carry any device, hence they are suitable for applications on factory floors to enforce safety and optimise humans' routes. However, radars have a number of recognised advantages with respect to visual based systems, i.e. they are cheaper and need simpler signal processing algorithms; they are more resilient to obstructions; they preserve privacy since it is almost impossible to reconstruct a person's identity from a cloud of points; and finally, they are not affected by lighting conditions. Therefore, they can be used both indoor and outdoor in any weather condition.

Early ideas on low-cost radars for short-range, high accuracy applications date back to a few years ago [4]. The recent evolution of solutions based on antenna arrays resulting from the massive research on multiple-input multiple-output (MIMO) systems motivated by 5G communications [5], [6] has triggered a new interest in this type of systems. Guerra et al. [7] proposed a channel model for personal radar applications based on mm-wave antenna arrays and use it to construct a map of the environment. Guidi et al. [8] analysed the influence of different design parameters on the mapping performance of mm-wave radars. Over the last few years, mm-wave RF signals have been used to localise and to track a moving target in a number of ways, e.g., by using Angle-of-Arrival measurements, by combining path loss with Angle-of-Arrival data or through Angle Differences-of-Arrival algorithms [9], [10]. Since mm-wave radars operate at frequencies higher than 10 GHz [11], [12], arrays of small and closely spaced antennas can be easily fit into a small space and with a reduced cost. This gives us the opportunity to estimate the position of the target using the Phase Differences of Arrival (PDoA) of the backscattered signals received by different antennas [13]. This general idea has been implemented on low-cost commercial off-the-shelf Systems-on-Chip (SoCs) embedding both RF circuitry and digital signal processing algorithms for position estimation. However, the actual performance and the limitations of these platforms are often quite unclear, since data sheets are not very reliable. To fill this gap, a thorough characterisation of a commercial mm-wave radar was performed in [14]. In this paper instead, first two different algorithms for people localisation and tracking based phased-array radars are described. Both algorithms are conceived to offer a good level of performance even in the case of delayed and spurious agent detection. Then, the viability of the proposed solutions is analysed through an experimental setup.

The rest of the paper is organised as follows. Section II reports some important background information on the kinematic and measurement models adopted. Section III describes the two proposed localisation and tracking algorithms. Section IV is focused on the experimental results validating such solutions. Finally, Section V summarises our conclusions and outlines future work directions.

II. Background and Problem Formulation

Some recent results [15], [16] reveal that, despite the kinematic flexibility of the human body, most of time people tend to move following nonholonomic unicycle-like kinematic model. In this section, we will rapidly review this model and show how to set up a radar–based tracking scenario. Hence, consider a unicycle-like agent moving according to the dynamic model

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} v\cos\theta \\ v\sin\theta \\ \omega \end{bmatrix}, \tag{1}$$

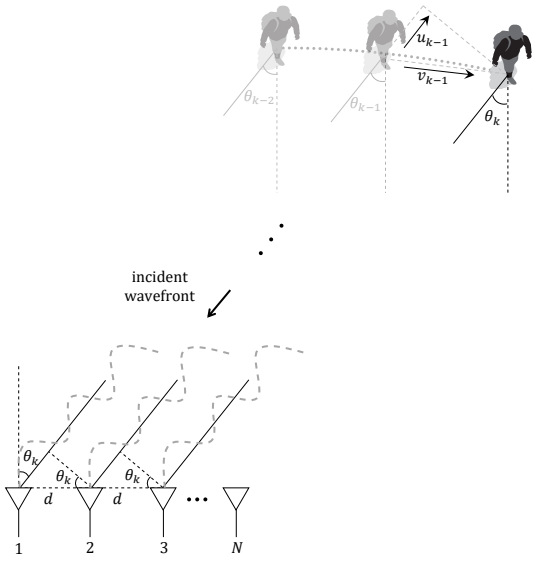


Fig. 1. Tracking scenario based on a stationary MIMO-FMCW radar and a moving target.

where $[p^T, \theta]^T$ is the state of the system (i.e., agent pose with Cartesian position $p = [x, y]^T$), while v and ω are agent forward and angular velocities, respectively. Assuming that the underlying motion inputs (namely the velocities) are sampled with a period T_s and held constant in each sampling period, we can write $v(t) = v(kT_s) = v_k \ \forall t \in [kT_s, (k+1)T_s)$. Similar assumptions and the same notation can be used for the angular velocity $\omega_k = \omega(kT_s)$. Therefore, the discrete–time equivalent model becomes

$$\boldsymbol{s}_{k+1} = \begin{bmatrix} x_{k+1} \\ y_{k+1} \\ \theta_{k+1} \end{bmatrix} = \begin{bmatrix} x_k + T_s \cos(\theta_k)\, v_k \\ y_k + T_s \sin(\theta_k)\, v_k \\ \theta_k + T_s \omega_k \end{bmatrix} = f(\boldsymbol{s}_k, v_k, \omega_k). \quad (2)$$

If the input velocities v_k and ω_k at a given time kT_s are not measured directly, but they are rather reconstructed from past data, it follows that the estimated values $\overline{v}_k = v_k + \varepsilon_k$ and $\overline{\omega}_k = \omega_k + \eta_k$, where ε_k and η_k are the respective uncertainty contributions, which are both assumed to be white, with zero mean and covariance matrix $\mathbf{Q}_k$.

A. Measurement Model

The principle of operation of a 60-GHz Frequency–Modulated Continuous-Wave (FMCW) radar platform relies on a sequence of n linear chirp signals that are first broadcasted and then received by a phased array after being reflected by one or more targets [17].

Consider for instance a target that at time step kT_s is located at a distance ρ_k and it is moving with a relative instantaneous velocity of u_k (see Figure 1). The round-trip delay of the signal backscattered by the target is a linear function of the instantaneous velocity u_k, of the distance ρ_k and of the speed of light in vacuum. Additionally, the angle ϕ_k of the target with respect to the radar may also affect the backscattered radar signal [18].

In addition, the frequency evolution depends on the Doppler frequency shift due to the relative speed u_k between the target and radar. This Doppler shift can be obtained by calculating the spectral change of the Intermediate Frequency (IF) signal over n consecutive chirps. To summarise, the measurements that can be collected from the radar at time kT_s are

$$\boldsymbol{z}_k = \begin{bmatrix} \overline{p}_k \\ \overline{v}_k, \\ \overline{\omega}_k \end{bmatrix} = h(\boldsymbol{s}_k, v_k, \omega_k) + \gamma_k, \quad (3)$$

where $\overline{p}_k = \rho_k[\cos(\phi_k), \sin(\phi_k)]^T$ and γ_k is the vector of white and zero-mean measurement uncertainty contributions with covariance matrix $\mathbf{R}_k$.

III. Motion Tracking Solutions

In this section, we will present our method for tracking a moving target using an FMCW-MIMO radar in cluttered conditions. The most common solution reported in the literature relies on applying the three-dimensional Fast Fourier Transform (3D-FFT) algorithm on the mixed transmitted and backscattered signals followed by a sensible thresholding of the resulting range-velocity-azimuth spectrum to extract the corresponding parameters [19]. However, spectrum thresholding will result in undesirable effects such as multiple detections of the same target or even loss of the target [14]. To account for this data association problem, two different approaches are proposed. The first one is based on a Track-Oriented Multi-Hypothesis Tracking (TOMHT) method [20], [21] combined with an Extended Kalman Filter (EKF), which relies on the assumption that the motion model of the target is known. In the second approach instead, an agnostic approach for robust tracking based on Expectation Maximisation (EM) is adopted both to estimate the model parameters and to track the target [22], [23].

A. TOMHT-EKF

The basic idea of TOMHT is to create alternative track hypotheses from measurement data in order to add them to a tree structure and, finally, to update the pool of valid track on a probabilistic basis by associating each measurement data to one of the available hypotheses (i.e., an existing track, a new track or an outlier). Each edge of the track tree is a track hypothesis which may trigger several new tracks anytime new measurement data are collected. The track tree will be updated upon the reception of new measurements: those hypotheses having no measurements associated in a given time horizon will be eliminated from the tree, thus limiting the number of global hypotheses on the trajectories of the moving target. Notice that even in the presence of only one moving target, multiple global hypotheses may exist due to possible static obstacles in the environment or outliers, which may create ghost targets as a result of reflections and interference phenomena. The position of the various global hypotheses are then tracked using different Extended Kalman Filters (EKFs), whose prediction step ensures target tracking even in the case of missing observations.

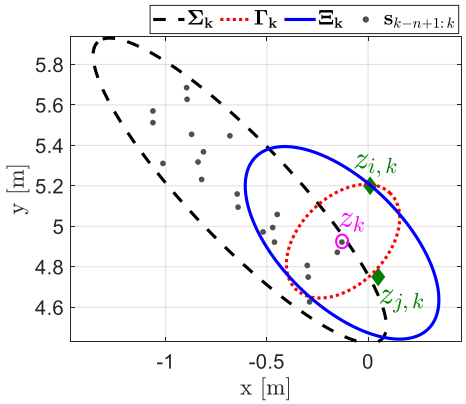


Fig. 2. Different uncertainty ellipsoids used in the calculation of Mahalanobis distance for measurement data association underlying the the Nearest-Neighbour (NN) method.

1) EKF: By using the dynamic model in (2), the state predicted by a standard EKF is given by $\hat{\boldsymbol{s}}_{k+1}^{-} = f(\hat{\boldsymbol{s}}_k, \overline{v}_k, \overline{\omega}_k)$, and the covariance matrix of such a predicted state is approximately given by

$$\mathbf{P}_{k+1}^{-} = \mathbf{A}_k \mathbf{P}_k \mathbf{A}_k^T + \mathbf{B}_k \mathbf{Q}_k \mathbf{B}_k^T, \tag{4}$$

where $\mathbf{A}_k = \frac{\partial f(\boldsymbol{s}_k, v_k, \omega_k)}{\partial \boldsymbol{s}_k}$ and $\mathbf{B}_k = \frac{\partial f(\boldsymbol{s}_k, v_k, \omega_k)}{\partial [v_k, \omega_k]}$ are the Jacobian matrices of function $f(\cdot,\cdot,\cdot)$ with respect to state and input variables, respectively, the $\hat{\cdot}$ symbol denotes the estimated quantities while superscript $\cdot^-$ represents the predicted ones. It then follows that the update equations are

$$\begin{aligned} \mathbf{K}_k &= \mathbf{P}_k^{-} \mathbf{H}_k^T (\mathbf{H}_k \mathbf{P}_k^{-} \mathbf{H}_k^T + \mathbf{R}_k)^{-1} = \mathbf{P}_k^{-} \mathbf{H}_k^T \boldsymbol{\Gamma}_k^{-1}, \\ \boldsymbol{s}_k &= \boldsymbol{s}_k^{-} + \mathbf{K}_k \widetilde{\boldsymbol{z}}_k, \\ \mathbf{P}_k &= (\mathbf{I} - \mathbf{K}_k \mathbf{H}_k) \mathbf{P}_k^{-}, \end{aligned} \tag{5}$$

where, considering (3), $\mathbf{K}_k$ is the Kalman gain, $\mathbf{H}_k = \frac{\partial h(\boldsymbol{s}_k, v_k, \omega_k)}{\partial \boldsymbol{s}_k}$, and $\widetilde{\boldsymbol{z}}_k = \boldsymbol{z}_k - h(\boldsymbol{s}_k^{-}, \overline{v}_k, \overline{\omega}_k)$ is the innovation term with covariance matrix $\boldsymbol{\Gamma}_k$.

2) Data Association: Due to the presence of obstacles and reflections, we need to associate the radar measurement data related to a given target, accepting those that are in close proximity, while rejecting those the distant ones. The Nearest-Neighbour (NN) method based on the Mahalanobis distance

$$d_{i,k} = \sqrt{\widetilde{\boldsymbol{z}}_{i,k}^T \boldsymbol{\Gamma}_k^{-1} \widetilde{\boldsymbol{z}}_{i,k}}, \tag{6}$$

was used to this purpose [24]. In (6) subscripts i, k stand for the i-th measurement at time kT_s. Indeed, multiple measurement data possibly associated to the same actual target, static object or outlier may exist, as previously discussed. Notice that the weighting matrix $\boldsymbol{\Gamma}_k$ accounts only for the impact of $\mathbf{P}_k^{-}$ and $\mathbf{R}_k$ covariance matrices, which are not sufficient for a robust association. This problem is shown for example in Figure 2: the values of $d_{i,\,k}^2 = 5.55$ and $d_{j,\,k}^2 = 7.78$ given by (6) result in the rejection of measurement $\boldsymbol{z}_{j,\,k}$ for the existing track $\boldsymbol{s}_{k-n+1:\,k}$, even if this measurement should be actually associated to the target considered. To account for this issue, we adopt $\boldsymbol{\Xi}_k = w\boldsymbol{\Sigma}_k + (1-w)\boldsymbol{\Gamma}_k$ as weighting matrix, i.e.,

$$d_{i,\,k} = \sqrt{\widetilde{\boldsymbol{z}}_{i,k}^T \boldsymbol{\Xi}_k^{-1} \widetilde{\boldsymbol{z}}_{i,k}} \tag{7}$$

where $w \in [0, 1)$ is a mixing factor and $\boldsymbol{\Sigma}_k$ represents the sample covariance matrix of a normal random vector modelling the distribution of $\boldsymbol{s}_{k-n+1:\,k}$, i.e., the past n samples associated to the corresponding track. The sample population size depends on the radar sampling frequency and how fast a target can change its direction. For a generic pedestrian walking, we empirically found that setting $n = 15$ (corresponding to approximately 1.5 s of sampling window with a sampling frequency of 10 frame per second) provides reasonably accurate results. With reference to Figure 2 and using (7) with $w = 0.3$, we have $d_{i,\,k}^2 = 5.40$ and $d_{j,\,k}^2 = 1.08$, thus correctly associating $\boldsymbol{z}_{j,\,k+1}$ to the track $\boldsymbol{s}_{k-n+1:\,k}$.

3) Velocity estimation: In order to use the EKF to account for missing measurement data, velocity estimation is needed. If human targets have to be tracked, both the forward $\overline{v}_k$ and angular $\overline{\omega}_k$ velocities of the target can be estimated from the past motion history along trajectory $\boldsymbol{s}_{0:k}$. To this end, we consider at first the constant velocity model (CVM) [25], the most simple, yet effective, target velocity predictor. In the simplest case, the CVM velocity inputs are determined with their expected values computed on a set of velocity measurements. Then, the Average/Adaptive velocity model (AVM) based on the inertia of the motion choices adopted by human beings is adopted [26]. Since the AVM does not make use of any filtering, it is highly sensitive to measurement noise. Hence, a low-pass filter or a smoother is usually needed [26]. The approach followed in this paper is to use the Smoothing Cubic Spline (SCS) interpolation based on the last m estimates of the EKF, i.e., $\boldsymbol{s}_{k-m+1:k}$, using the well-known De Boor's approach [27] prior to computing the mean velocity values within such a window.

The third velocity estimation model adopted in this paper is a time series forecast model based on a Long short-term memory (LSTM) artificial recurrent neural network [28]. LSTM predictors outperforms the conventional low-pass filters and can further improve the AVM accuracy [26]. In particular, the LSTM predictor exploits the AVM velocities $\overline{v}_{k-1}$ and $\overline{\omega}_{k-1}$ as inputs to synthesise $\overline{v}_k$ and $\overline{\omega}_k$.

B. EM Robust Tracking

A major known hurdle for all multi–hypotheses tracking algorithms is related to the fact that not all the observations drawn from an estimated distribution are associated to the global hypothesis during the tracking phase. This may lead to the impossibility to estimate the actual covariance matrix of the measurement data associated to the tracked hypotheses. Similarly, the assumption on the target motion, i.e., the parameters of the dynamic model (2), can be inconsistent with the actual target motion (especially if a human target is considered), thus reducing the effectiveness of the EKF. Therefore, we first restrict the model to Cartesian coordinates $p_k = [x_k, y_k]^T$ only and present a solution for the simultaneous online estimate

of all the parameters $\mathcal{X} = \{\mathbf{A}, \mathbf{H}, \mathbf{R}, \mathbf{Q}, p_0, \overline{\mathbf{P}}_0\}$, where p_0 and $\overline{\mathbf{P}}_0$ are the initial state and the covariance matrix, respectively, needed for EKF initialisation and based only on radar measurements. Notice that $\overline{\mathbf{P}}_0$ refers only to p_0, while the full state $\boldsymbol{s}_k$ and the associated covariance matrix $\mathbf{P}_0$ are needed for the EKF. It is important to emphasise that, the target orientation θ_k and its variance can always be reconstructed from p_k, p_{k+1}, $\overline{\mathbf{P}}_k$ and $\overline{\mathbf{P}}_{k+1}$ [29]. We re-state the adopted model using a position-based Linear time-invariant Dynamic System (LDS) excited by white random noises as follows:

$$\begin{aligned} p_{k+1} &= \mathbf{A}p_k + \nu_k, \\ \boldsymbol{z}_k &= \mathbf{H}p_k + \delta_k. \end{aligned} \tag{8}$$

In order to identify the parameters of set $\mathcal{X}$, the EM algorithm can be used [22], [23]. First, a preliminary estimate $\widehat{p}_k$, its covariance $\widehat{\overline{\mathbf{P}}}_k$ and the Jacobian $\mathbf{J}_k$ can be obtained by using a Rauch-Tung-Striebel (RTS) smoother as follows [30]:

$$\begin{aligned} \widehat{p}_k &= p_k + \mathbf{J}_k\big(\widehat{p}_{k+1} - \mathbf{A}p_k\big), \\ \widehat{\overline{\mathbf{P}}}_k &= \overline{\mathbf{P}}_k + \mathbf{J}_k\big(\widehat{\overline{\mathbf{P}}}_{k+1} - \overline{\mathbf{P}}_k^-\big)\mathbf{J}_k^T, \\ \mathbf{J}_k &= \overline{\mathbf{P}}_k\mathbf{A}^T(\overline{\mathbf{P}}_k^-)^{-1}, \end{aligned} \tag{9}$$

where $\overline{\mathbf{P}}_k^-$ is the covariance matrix (resulting from a first-order approximation) of the predicted state, while p_k and $\overline{\mathbf{P}}_k$ are are updated state estimate and the respective covariance matrix resulting from the the Kalman filter associated to the RTS. Then, in the **E-step** of the EM algorithm, the following expected values are computed, i.e.,

$$\begin{aligned} e_k &= \mathrm{E}\left\{p_k\right\} = \widehat{p}_k, \\ e_k^{k-1} &= \mathrm{E}\left\{p_k p_{k-1}^T\right\} = \mathbf{J}_{k-1}\widehat{\overline{\mathbf{P}}}_k + \widehat{p}_k\widehat{p}_{k-1}^T, \\ e_k^k &= \mathrm{E}\left\{p_k p_k^T\right\} = \widehat{\overline{\mathbf{P}}}_k + \widehat{p}_k\widehat{p}_k^T, \\ e_{k-1}^k &= \mathrm{E}\left\{p_{k-1} p_k^T\right\} = \mathbf{J}_k\widehat{\overline{\mathbf{P}}}_{k-1} + \widehat{p}_{k-1}\widehat{p}_k^T. \end{aligned} \tag{10}$$

Finally, in the **M-step** the following matrices are updated by using a maximum likelihood estimator processing the last $k \geq 2$ measurement data, i.e.

$$\begin{aligned} \mathbf{A}^+ &= \sum_{i=2}^{k} e_i^{i-1}\left(\sum_{i=2}^{k} e_{i-1}^{i-1}\right)^{-1}, \\ \mathbf{Q}^+ &= \frac{1}{k-1}\sum_{i=2}^{k} e_i^k - \mathbf{A}^+ e_{i-1}^i - e_i^{i-1}\mathbf{A}^+ + \mathbf{A}^+ e_{i-1}^{i-1}\mathbf{A}^{+T}, \\ \mathbf{H}^+ &= \sum_{i=1}^{k} \overline{\boldsymbol{z}}_i e_i^T\left(\sum_{i=1}^{k} e_i\right)^{-1}, \\ \mathbf{R}^+ &= \frac{1}{k}\sum_{i=1}^{k} \overline{\boldsymbol{z}}_i\overline{\boldsymbol{z}}_i^T - \mathbf{H}^+ e_i\overline{\boldsymbol{z}}_i^T - \overline{\boldsymbol{z}}_i e_i^T\mathbf{H}^+ + \mathbf{H}^+ e_i^i\mathbf{H}^+. \end{aligned} \tag{11}$$

Notice that the elements of $\overline{\boldsymbol{z}}_i$ refer just to the measured target position in (3), namely the coordinates of $\bar{p}_k$.

The EM algorithm is quite sensitive to the initial conditions and the initial choices of the corresponding process and measurement covariance matrices [31], which may lead to convergence problems. To account for this problem, we use the subspace approach for parameters initialisation: given a set of vectors of measurement $\overline{\boldsymbol{z}}_1, \ldots, \overline{\boldsymbol{z}}_k$, the following Hankel matrix is built, i.e.,

$$\mathcal{H} = \begin{bmatrix} \overline{\boldsymbol{z}}_{1,\,x} & \overline{\boldsymbol{z}}_{2,\,x} & \cdots & \overline{\boldsymbol{z}}_{K-1,\,x} \\ \overline{\boldsymbol{z}}_{1,\,y} & \overline{\boldsymbol{z}}_{2,\,y} & \cdots & \overline{\boldsymbol{z}}_{K-1,\,y} \\ \overline{\boldsymbol{z}}_{2,\,x} & \overline{\boldsymbol{z}}_{3,\,x} & \cdots & \overline{\boldsymbol{z}}_{K,\,x} \\ \overline{\boldsymbol{z}}_{2,\,y} & \overline{\boldsymbol{z}}_{3,\,y} & \cdots & \overline{\boldsymbol{z}}_{K,\,y} \end{bmatrix}, \tag{12}$$

where $\overline{\boldsymbol{z}}_{i,\,x}$ and $\overline{\boldsymbol{z}}_{i,\,y}$ refer to the x and y components of the i-th measured point $\overline{\boldsymbol{z}}$. If U, S an V are the matrices resulting from the singular value decomposition (SVD) of $\mathcal{H}$, the initial value of $\mathbf{H}$ is set equal to $\mathbf{H}_0 = U_{1:2,\,1:2}$ (the notation $U_{1:2,\,1:2}$ denotes the first two rows and two columns of the matrix U). As far as matrix $\mathbf{A}$ is concerned, we first need to calculate the extended observability matrix $\mathbf{W} = S_{1:2,\,1:2}V_{1:2,\,1:2}^T$, and then

$$\mathbf{A}_0 = \mathbf{W}_{1:2,\,2:K}\mathbf{W}_{1:2,\,1:K-1}^{\dagger}, \tag{13}$$

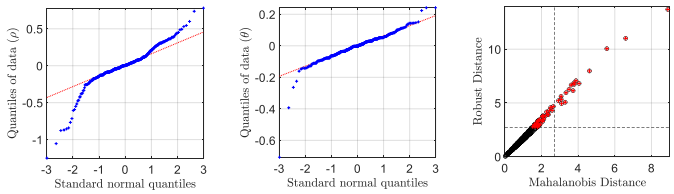
where the $\dagger$ symbol denotes the pseudo-inverse. Hence, the EM algorithm starts with the initial vector of parameters $\mathcal{X} = \{\mathbf{A}_0, \mathbf{R}_0, \mathbf{H}_0, \mathbf{Q}_0, p_0, \mathbf{P}_0\}$, where $\mathbf{R}_0$, $\mathbf{Q}_0$ and $\mathbf{P}_0$ are set equal to the identity matrix, while $p_0 = \mathbf{W}_{1:2,\,1}$ is the first column of the extended observability matrix. Technically speaking, the convergence of the EM algorithm is evaluated at each iteration in the **E-step** using a log-likelihood function condition, i.e.

$$\big|(\log(f(z_k)) - \log(f(z_{k-1}))\big| < \tau \times \big|\log(f(z_{k-1}))\big|, \tag{14}$$

that, when satisfied, stops the EM algorithm. The measurement likelihood function is given by

$$f(z_k) = \mathcal{N}(z_k|\mathbf{H}^+ p_k, \mathbf{H}^+\overline{\mathbf{P}}_k\mathbf{H}^{+^T} + \mathbf{R}^+)$$

where τ is a convergence parameter threshold.



(a) (b) (c)

Fig. 3. Radar measurement analysis. (a, b) Measurement quantiles versus the theoretical standard Gaussian distribution for target range and angle of arrival (AoA), respectively. (c) Robust covariance estimation.

IV. Experimental results

The efficiency of the proposed tracking methodology was analysed on the field by using an FMCW radar platform based on a System-on-Chip (SOC) TI IWR6843. The environment used for testing was a 12×8 m^2 arena monitored by an OptiTrack system equipped with 14 calibrated cameras tracking one or more targets at 125 Hz with accuracy in the order of a few mm. Hence, the data collected by the OptiTrack was regarded as ground truth. The radar platform returns measurements at a maximum rate of 10 frames per second. In

TABLE I
FINAL ESTIMATED LDS PARAMETERS USING EM.

Trajectory	**A**	**H**
eight-shaped	$\begin{bmatrix} 0.99 & 16 \cdot 10^{-4} \\ -12 \cdot 10^{-4} & 0.99 \end{bmatrix}$	$\begin{bmatrix} 0.030 & -0.41 \\ 0.40 & 0.031 \end{bmatrix}$
random	$\begin{bmatrix} 0.99 & -18 \cdot 10^{-3} \\ 5 \cdot 10^{-4} & 0.99 \end{bmatrix}$	$\begin{bmatrix} 0.019 & -0.41 \\ 0.40 & 0.037 \end{bmatrix}$
diagonal	$\begin{bmatrix} 0.99 & 74 \cdot 10^{-4} \\ -13 \cdot 10^{-4} & 0.99 \end{bmatrix}$	$\begin{bmatrix} 0.075 & -0.40 \\ 0.40 & 0.081 \end{bmatrix}$

the experiments, we used the EM algorithm with a maximum number of iterations set to 2000 and with the convergence threshold $\tau = 10^{-4}$ in (14).

We first analysed the performance of the TOMHT-EKF based on the RTS smoother applied to the raw FMCW radar measurement data. Then, the impact of the LDS parameter estimation on tracking accuracy improvement was estimated. For the first scenario (i.e., standard TOMHT without parameters prediction), the corresponding measurement covariance matrix $\mathbf{R}$ was estimated by using 357 FMCW radar measurements data which were collected in 8 different locations with known coordinates. The QQ plots in Figure 3(a)-(b) shows the dataset quantiles versus the theoretical quantile of a standard gaussian distribution. As can be seen from the figure, neither target angle nor its range follow a normal distribution. Nonetheless, since the radar measurement data are affected by outliers, the robust covariance estimation method was employed to estimate the measurement covariance matrix [32]. In this way, we found that

$$\mathbf{R} = \begin{bmatrix} 0.063 & 0.0003 \\ 0.0003 & 0.0565 \end{bmatrix}$$

The above covariance matrix was employed in the EKF for the TOMHT algorithm. The DD plot in Figure 3(a) shows the standard Mahalanobis distance of the data versus the robust mahalanobis distance, thus identifying the portion of the data used for covariance matrix estimation.

In the second scenario (i.e. EM Robust tracking), the experimental results verify that estimating the underlying LDS parameters always yield better performance. Figure 4 and Figure 5 show the tracking results of three sample experimental trajectories in terms of the qualitative trajectory estimation and the percentile of average absolute error for the target position along the trajectory. As can be seen, in all cases the EM algorithm provides more accurate estimation while still remains robust to the wrong hypothesis estimated by the TOMHT-EKF. The final estimated LDS and measurement covariance matrix are reported in Table I and Table II, respectively.

V. CONCLUSIONS

We have presented two different techniques to implement a human localisation and tracking solution based on the data collected by a low-cost phased-array FMCW radar platform. A critical problem was data association, i.e., how to associate the radar measurement data to a specific target while rejecting artefacts and bad data. We proposed two techniques. The first one is based on the combination of Track–Oriented Multi–Hypothesis Tracking (TOMHT) with an Extended Kalman Filter. The second one is based instead on Expectation Maximisation (EM). Both techniques were validated experimentally.

In future works we plan first to perform an in-depth performance comparison between the proposed solutions in different scenarios, and then to use the radar platform for mobile robot applications, too.

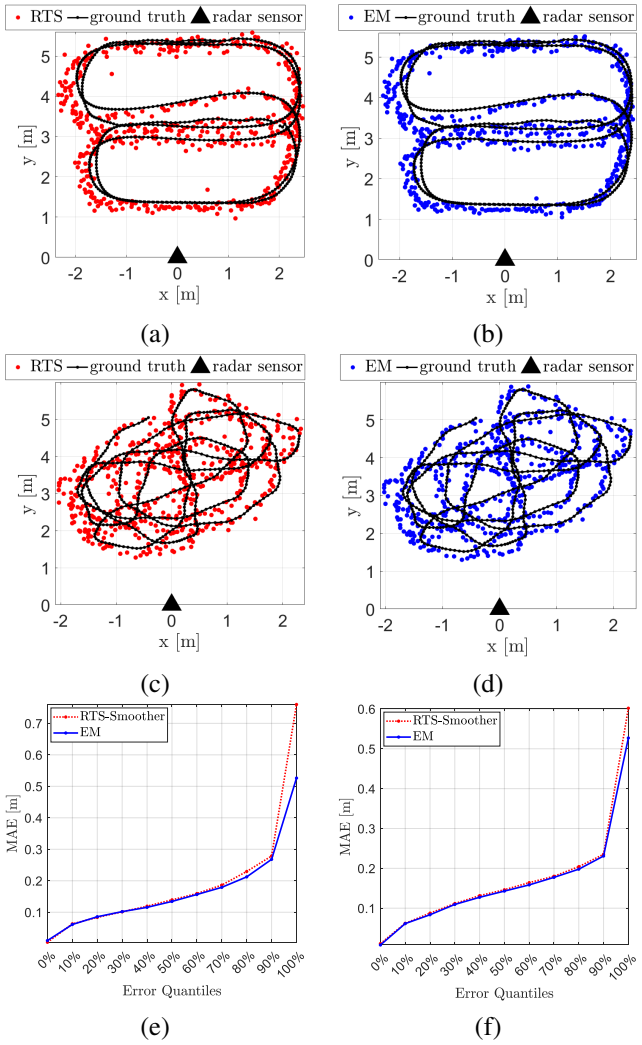


Fig. 4. Tracking performance for an eight-shaped trajectory with (a) RTS-Smoother (b) EM, and for a random shaped trajectory with (c) RTS-Smoother (d) EM. Mean absolute error (MAE) of the target position for the eight-shaped trajectory (e) and for the random trajectory (f).

REFERENCES

[1] C. Laoudias, A. Moreira, S. Kim, S. Lee, L. Wirola, and C. Fischione, "A survey of enabling technologies for network localization, tracking, and navigation," *IEEE Communications Surveys Tutorials*, vol. 20, no. 4, pp. 3607–3644, Fourth quarter 2018.

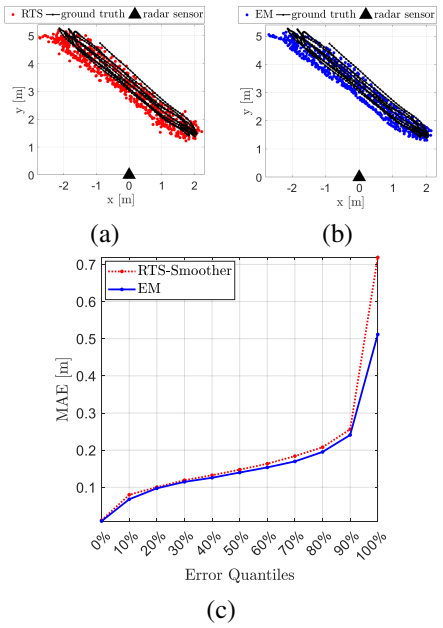


Fig. 5. Tracking performance for a diagonal shaped trajectory. (a) RTS-Smoother (b) EM (c) mean absolute error (MAE) of the target position.

TABLE II
FINAL ESTIMATED PROCESS NOISE AND MEASUREMENT COVARIANCE MATRICES OBTAINED USING EM.

Trajectory	$\mathbf{R}$	$\mathbf{Q}$
eight-shaped	$10^{-4}\begin{bmatrix} 47 & -8 \\ -8 & 35 \end{bmatrix}$	$\begin{bmatrix} 0.055 & -72 \cdot 10^{-4} \\ -72 \cdot 10^{-4} & 0.13 \end{bmatrix}$
random	$10^{-4}\begin{bmatrix} 25 & -5 \\ -5 & 30 \end{bmatrix}$	$\begin{bmatrix} 0.089 & -12 \cdot 10^{-3} \\ -11 \cdot 10^{-3} & 0.085 \end{bmatrix}$
diagonal	$10^{-4}\begin{bmatrix} 44 & -34 \\ -34 & 53 \end{bmatrix}$	$\begin{bmatrix} 0.088 & 0.064 \\ 0.064 & 0.061 \end{bmatrix}$

[2] A. J. Davison, "Real-time simultaneous localisation and mapping with a single camera." in *Iccv*, vol. 3, 2003, pp. 1403–1410.

[3] B. Benfold and I. Reid, "Stable multi-target tracking in real-time surveillance video," in *CVPR 2011*. IEEE, 2011, pp. 3457–3464.

[4] C. Zhang, M. Kuhn, B. Merkl, M. Mahfouz, and A. E. Fathy, "Development of an UWB indoor 3D positioning radar with millimeter accuracy," in *2006 IEEE MTT-S International Microwave Symposium Digest*, San Francisco, CA, USA, Jun. 2006, pp. 106–109.

[5] R. Mendrzik, H. Wymeersch, and G. Bauch, "Joint localization and mapping through millimeter wave MIMO in 5G systems," in *Proc. 2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, UAE, Dec. 2018, pp. 1–6.

[6] Z. Lin, T. Lv, and P. T. Mathiopoulos, "3-D indoor positioning for millimeter-wave massive MIMO systems," *IEEE Transactions on Communications*, vol. 66, no. 6, pp. 2472–2486, Jun. 2018.

[7] A. Guerra, F. Guidi, D. Dardari, A. Clemente, and R. D'Errico, "A millimeter-wave indoor backscattering channel model for environment mapping," *IEEE Transactions on Antennas and Propagation*, vol. 65, no. 9, Sep. 2017.

[8] F. Guidi, A. Guerra, D. Dardari, A. Clemente, and R. D'Errico, "Environment mapping with millimeter-wave massive arrays: System design and performance," in *2016 IEEE Globecom Workshops (GC Wkshps)*, Washington, DC, USA, Dec. 2016, pp. 1–6.

[9] O. Kanhere and T. S. Rappaport, "Position locationing for millimeter wave systems," in *Proc. 2018 IEEE Global Communications Conference (GLOBECOM)*, Abu Dhabi, UAE, Dec. 2018, pp. 206–212.

[10] A. Olivier, G. Bielsa, I. Tejado, M. Zorzi, J. Widmer, and P. Casari, "Lightweight indoor localization for 60-GHz millimeter wave systems," in *Proc. 13th Annual IEEE Int. Conf. on Sensing, Communication, and Networking (SECON)*, London, UK, Jun. 2016, pp. 1–9.

[11] P. Park, S. Kim, S. Woo, and C. Kim, "A centimeter resolution, 10 m range CMOS impulse radio radar for human motion monitoring," *IEEE Journal of Solid-State Circuits*, vol. 49, no. 5, pp. 1125–1134, May 2014.

[12] H. J. Ng, W. Ahmad, M. Kucharski, J. Lu, and D. Kissinger, "Highly-miniaturized 2-channel mm-wave radar sensor with on-chip folded dipole antennas," in *Proc. IEEE Radio Frequency Integrated Circuits Symposium (RFIC)*, Honolulu, HI, USA, Jun. 2017, pp. 368–371.

[13] R. Ebelt, A. Hamidian, D. Shmakov, T. Zhang, V. Subramanian, G. Boeck, and M. Vossiek, "Cooperative indoor localization using 24-GHz CMOS radar transceivers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 62, no. 9, pp. 2193–2203, Sep. 2014.

[14] A. Antonucci, M. Corrà, A. Ferrari, D. Fontanelli, E. Fusari, D. Macii, and L. Palopoli, "Performance Analysis of a 60-GHz Radar for Indoor Positioning and Tracking," in *International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. Pisa, Italy: IEEE, October 2019, pp. 1–7.

[15] G. Arechavaleta, J.-P. Laumond, H. Hicheur, and A. Berthoz, "On the nonholonomic nature of human locomotion," *Autonomous Robots*, vol. 25, no. 1, pp. 25–35, 2008.

[16] F. Farina, D. Fontanelli, A. Garulli, A. Giannitrapani, and D. Prattichizzo, "Walking Ahead: The Headed Social Force Model," *PLOS ONE*, vol. 12, no. 1, pp. 1–23, 1 2017.

[17] I. V. Komarov and S. M. Smolskiy, *Fundamentals of short-range FM radar*. Artech House, 2003.

[18] V. Winkler, "Range doppler detection for automotive fmcw radars," in *2007 European Radar Conference*. IEEE, 2007, pp. 166–169.

[19] W. Zhang, H. Li, G. Sun, and Z. He, "Enhanced detection of doppler-spread targets for fmcw radar," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 4, pp. 2066–2078, 2019.

[20] Y. Kosuge, Y. Tachibana, and S. Tsujimichi, "A track-oriented multiple hypothesis multitarget tracking algorithm," *Electronics and Communications in Japan (Part I: Communications)*, vol. 82, no. 12, pp. 84–92, 1999.

[21] S. Coraluppi and C. Carthel, "Generalizations to the track-oriented mht recursion," in *2015 18th International Conference on Information Fusion (Fusion)*. IEEE, 2015, pp. 346–350.

[22] Z. Ghahramani and G. E. Hinton, "Parameter estimation for linear dynamical systems," 1996.

[23] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*. Springer, 2006, vol. 4, no. 4.

[24] X. R. Li and Y. Bar-Shalom, "Tracking in clutter with nearest neighbor filters: analysis and performance," *IEEE transactions on aerospace and electronic systems*, vol. 32, no. 3, pp. 995–1010, 1996.

[25] M. Luber, J. A. Stork, G. D. Tipaldi, and K. O. Arras, "People tracking with human motion predictions from social forces," in *2010 IEEE International Conference on Robotics and Automation*, 2010, pp. 464–469.

[26] C. Schöller, V. Aravantinos, F. Lay, and A. Knoll, "What the constant velocity model can teach us about pedestrian motion prediction," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 1696–1703, 2020.

[27] C. De Boor and C. De Boor, *A practical guide to splines*. springer-verlag New York, 1978, vol. 27.

[28] K. Yeon, K. Min, J. Shin, M. Sunwoo, and M. Han, "Ego-vehicle speed prediction using a long short-term memory based recurrent neural network," *International Journal of Automotive Technology*, vol. 20, no. 4, pp. 713–722, 2019.

[29] D. Fontanelli, F. Shamsfakhr, D. Macii, and L. Palopoli, "An Uncertainty-driven and Observability-based State Estimator for Nonholonomic Robots," *IEEE Trans. on Instrumentation and Measurement*, vol. 70, pp. 1–12, January 2021.

[30] H. E. Rauch, F. Tung, and C. T. Striebel, "Maximum likelihood estimates of linear dynamic systems," *AIAA journal*, vol. 3, no. 8, pp. 1445–1450, 1965.

[31] D. Barber, *Bayesian reasoning and machine learning*. Cambridge University Press, 2012.

[32] P. J. Rousseeuw and K. V. Driessen, "A fast algorithm for the minimum covariance determinant estimator," *Technometrics*, vol. 41, no. 3, pp. 212–223, 1999.