A multi-target detection and position tracking algorithm based on mmWave-FMCW radar data

Farhad Shamsfakhr [a], David Macii [a,*], Luigi Palopoli [b], Michele Corrà [c], Alessandro Ferrari [d], Daniele Fontanelli [a]

[a] *University of Trento, Department of Industrial Engineering, via Sommarive, 9, Trento, 38123, Italy*
[b] *University of Trento, Department of Information Engineering and Computer Science, via Sommarive, 9, Trento, 38123, Italy*
[c] *Tretec S.r.l., via Praga, 20, Trento, 38121, Italy*
[d] *VNG Ingegneria, via A. Degasperi, 77, Trento, 38123, Italy*

ARTICLE INFO

ABSTRACT

Detecting and tracking the position of multiple targets indoors is a challenging measurement problem due to the inherent difficulty to cluster correctly the sensor data associated to a given target and to track the position of each cluster with adequate accuracy. This problem is critical especially in rooms filled with fixed or moving objects hampering target detection and whenever the paths of different targets cross one another. In this paper, a robust Multiple Targets Tracking (MTT) algorithm exploiting the clouds of points collected from a mmWave-FMCW radar is presented. The proposed solution consists of four main steps. First, the possible outliers of a raw radar data set are removed using a neural network model. Next, the cleaned-up radar data are clustered using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm. Then, a Kalman Filter (KF) is used to track the position of the centroid of each cluster. Finally, a Structured Branching Multiple Hypothesis Testing (SBMHT) algorithm is applied and updated over reasonably short time intervals to decide which detected tracks are supposed to be confirmed and which ones instead should be discarded. The proposed MTT technique was validated experimentally using the data sets collected from a 60-GHz TI IWR6843 radar platform. The reported results show that the developed algorithm, if properly tuned, is faster and returns more accurate results than other MTT techniques. In particular, the percentage of detection errors is negligible and the planar positioning accuracy is within about 30 cm with 90% probability when up to five targets move freely within the same room.

1. Introduction

The growing interest in indoor tracking and localisation of people and vehicles in indoor environments has recently gained momentum primarily due to the widespread adoption of mobile smart devices, which provide wireless and effortless services to users [1]. Several wireless technologies are utilised for indoor agent tracking, with WiFi and Ultra-Wideband (UWB) technologies taking the lion's share [2,3]. WiFi technologies have an indisputable advantage: they are deployed everywhere to provide connectivity services. However, when they are used for localisation purposes, accuracy is far from satisfactory [4,5]. Conversely, indoor localisation based on UWB signals ensures superior accuracy and enhanced robustness [6], but it requires a specific UWB infrastructure [7,8]. Moreover, the need to wear electronic devices with a unique identifier raises understandable privacy concerns.

To address this issue, device-free, memoryless tracking methods have become increasingly popular [9,10]. Such techniques are non-invasive, and allow users to move freely without having to be tethered to their identifying devices.

The Millimetre-Wave Frequency Modulated Continuous Wave (mmWave-FMCW) radars have gained popularity over the last few years mainly due to the evolution of antenna arrays [11]. Indeed, they can be used not only for indoor positioning, but also for vital signs detection [12]. A single mmWave-FMCW radar can potentially track multiple, unknown users at the same time. As a consequence, mmWave-FMCW radars do not suffer from privacy concerns [13]. Thus, they could revolutionise target localisation and tracking in indoor environments, even because their detection range is much longer than the range of Ultra-High Radio Frequency Identification (UHF-RFID)

---

* Corresponding author.
E-mail addresses: fshamsfakhr@gmail.com (F. Shamsfakhr), david.macii@unitn.it (D. Macii), luigi.palopoli@unitn.it (L. Palopoli), michele.corra@3tec.it (M. Corrà), alessandro.ferrari.vng@gmail.com (A. Ferrari), daniele.fontanelli@unitn.it (D. Fontanelli).

readers [14–16]. The data generally returned by a mmWave-FMCW radar are the coordinates of the clouds of points reflected by some target moving in the space in front of the radar's antenna array [17]. Unfortunately this points cloud can be quite noisy. In [18], a PointNet auto-encoder neural network is adopted to enhance the points cloud of data collected from a mmWave-FMCW radar. In [19] instead a localisation framework based on the fusion of UWB and mmWave-FMCW radar data is proposed both in outdoor and indoor scenarios. When multiple targets move in the same environment, three additional critical problems must be addressed. First of all, the amount of outliers in the data set is much greater than in the case of a single target. Therefore, removing such outliers is critically important for performance. Secondly, the radar points must be clustered and properly associated to different moving targets. Finally, the time-varying position of such clusters should consistently represent the position of the corresponding targets. Thus, to address the aforementioned issues, a Multi-Target Tracking (MTT) algorithm is needed. These kinds of algorithms are designed to simultaneously track the positions of multiple objects over time in dynamic environments [20]. Among the existing solutions, the Structured Branching Multiple Hypothesis Tracking (SBMHT) offers significant computational savings compared to other approaches [21], as it will be explained in Section 2 and shown in Section 6. However, the standard SBMHT approach is not very reliable in small indoor environments. This is why in this paper we propose a solution to improve both multi-target detection reliability and tracking performance accuracy in indoor scenarios.

Therefore, the key elements of novelty of this work are: a better outlier removal scheme before data clustering and a robust multi-hypotheses tracking approach, which requires just a preliminary tuning of the algorithm parameters in the chosen environment.

The rest of the paper is structured as follows. Section 2 presents some related work in the field of radar-based multi-target indoor positioning. Section 3 introduces the multi target detection problem using radars, while Section 4 describes the theoretical framework and the proposed solution. Section 5 discusses the importance of parameters tuning and explains the main settings of the experimental setup. The detection and tracking performance based on several experimental results are reported in Section 6. Finally, in Section 7 the main conclusions are drawn and the future work is outlined.

2. Related work on radar-based positioning

To discriminate the valid data from the noise floor, a heuristic threshold is usually applied to the raw mmWave-FMCW signals. The Constant False Alarm Rate (CFAR) detector is the simplest and most common (and yet powerful) filtering technique to decrease the number of clouds points to a manageable level (some further details on CFAR filtering are reported in Section 5.1) [22]. Nevertheless, the position points returned by the radar include sparse observations due to reflections that do not belong to any target. The amount of such "outliers" or "false alarms" in small rooms may be very large [23]. The combination of: data sparsity, sporadic missing detection in the case of distant targets, noise and multipath propagation effects, lack of knowledge of the true shape of the clusters of points associated to a given target, make the performance of the traditional outlier detection methods dramatically poor when applied to radar data [24,25]. As suggested by Keller et al. [26], efficient methods for analysing these types of clusters often require self-adaptive thresholding mechanisms for labelling non-clustered and clustered data. The main reason is that the optimal thresholds depend on the overall density of the cluster points. So a method that is appropriate for a low number of moving targets could not be suitable when the number of targets grows.

As a consequence, classic parametric clustering methods, like the K-means algorithm [27], the Gaussian Mixture Model (GMM) criterion [28], and the agglomerative hierarchical technique [29] turn to be hardly effective for outliers detection when applied to radar data. On the other hand, supervised cluster analysis and outlier detection techniques for radar tracking have recently attracted increasing attention [30,31].

Many MTT systems based on radars typically rely on Multiple Hypothesis Tracking (MHT) techniques to overcome the aforementioned challenges and to tackle the data association problem [32]. Over the past decades, various MHT approaches have been employed for data association, detection, and tracking (see [33,34] for a comprehensive review). In [21] the problem of making multiple data association hypotheses can be considerably mitigated by dividing the entire set of tracks and observations into separate clusters. Furthermore, many other approaches rely on the idea that a big tracking problem can be split into a set of smaller problems to be addressed independently [35]. Fontana et al. [36] propose a clustering and merging approach based on a Poisson multi-Bernoulli mixture (PMBM) filter, which looks suitable for multiple target tracking with a large number of agents. The research work by He et al. [37] introduces an innovative multi-sensor, multi-target tracking technique, followed by a fusion stage using clustering and statistical tests to group local tracks into clusters, with a global estimation based on the generalised covariance intersection (GCI) algorithm. Extensive simulations confirm its effectiveness for multi-target tracking. In [36] a data-driven clustering algorithm is adopted to divide the data association problem into sub-problems and to derive the clustered PMBM posterior density via Kullback–Leibler divergence minimisation. Although this method mitigates the computational burden when the number of targets grows, the tracking performance in the presence of occlusions is still lower than using standard Track-Oriented Multiple Hypothesis Tracking (TOMHT) approaches.

One big challenge of the MHT-based approaches is the difficulty to ensure consistency, as it is commonly assumed that multiple tracks in a given hypothesis cannot share the same measurements [38]. This means that if one track relied on a detected point to update the position of a target, any other track using the same point would conflict with the first association. This assumption may not hold true when targets are close to each other and when the distance between them is comparable with sensor spatial resolution. This problem is recognised to be particularly tough [39]. Most of the existing MHT approaches are not able to provide accurate results in such situations, as the amount of false tracks can be particularly high and dominate most hypotheses. In addition, the family of MHT approaches suffers from another important drawback: the need to keep track of and to test a vast number of hypotheses. This issue persists even when most hypotheses are discarded immediately. In [40] a multi-hypotheses fractional belief propagation (MHFBP) algorithm for radar-based MTT is proposed. This technique effectively addresses computational challenges and outperform the classic MHT, the feature-aided MHT (FA-MHT), and the MHT-belief propagation (MHT-BP) with improved tracking performance and reduced computational burden in diverse scenarios. However, even this method does not offer any effective way to reduce the number of false detected tracks. As correctly explained by Smiti et al. [41], the supervised cluster analysis and outlier detection approaches are more effective in detecting abnormal events when they do not derive from sporadic points that are faraway from the centre of a cluster. Moreover, such solutions are also preferable in the case of small clusters with temporal and spatial local anomalies.

In our research, we noticed that the problem of clustering indoor radar data when multiple targets are present is similar to the complex molecular data analysis based on biological specimens [42]. Recently, a robust, fast and accurate outlier rejection method for microscopy data localisation was presented in [43]. That method can detect outliers in clusters of measurement data having nonuniform and non-stationary densities. Taking inspiration from this method, in this paper a novel and effective technique to remove outliers from raw radar data is presented. This approach is combined with an efficient method based on the SBMHT algorithm (i.e., a computationally efficient version of MHT) to detect and to track multiple targets [21]. As described in the following, this algorithm provides a novel solution in the panorama of indoor positioning techniques, and it is essential to ensure the robust tracking of multiple targets in the same environment.

3. The multi-target detection problem

Let n_o be the number of stationary or moving objects in an environment. The state (i.e., the planar position) of each object o at the kth time step (with $o \in \mathcal{O} = \{1, \ldots, n_o\}$) is described by vector $s_{o,k} = [x_{o,k}, y_{o,k}]^T$.

Let $\mathbf{S}_k = [s_{1,k}^T, \ldots, s_{n_o,k}^T]^T$ be the vector including the state of all the objects. Each data frame returned by the mmWave-FMCW radar due to the signal reflections caused by an object with coordinates (x_k, y_k) includes: Doppler frequency shift, object velocity, distance and bearing angle between the radar antennas array and the object. Thus, distance, bearing angle and Doppler frequency shift measured at the kth time step can be modelled as follows:

$$z_k = \begin{bmatrix} \bar{r}_k \\ \bar{\phi}_k \\ \bar{d}_k \end{bmatrix} = \begin{bmatrix} \sqrt{x_k^2 + y_k^2} + \gamma_{r,k} \\ \tan^{-1}(y_k/x_k) + \gamma_{\phi,k} \\ d_k + \gamma_{d,k} \end{bmatrix}, \tag{1}$$

where $\gamma_{r,k}$, $\gamma_{\phi,k}$ and $\gamma_{d,k}$ are the uncertainty contributions associated with range, bearing and frequency shift measurements, respectively. Depending on a variety of factors, such as radar settings (e.g., frequency, bandwidth and rate of change of frequency of chirp signals), Doppler effect resolution and CFAR threshold, the amount of data frames associated to the detected targets may differ from the number of objects that are actually present in the room and it may also change with time. Hence, considering that there are n_o objects in the scene that may or may not be detected by the radar and that multiple measurements may refer to the same object, the radar measurement data in (1) can be labelled with an additional index q, i.e. $z_{q,k}$, with $q \in \mathcal{Q}_k = \{1, \ldots, n_k\}$, with n_k being the number of target measurements. For instance, in a room with $n_o = 2$ objects, at the kth time step we may have $n_k = 10$ measurements: 7 of them refer to first target, while the other 3 refer to the second one. Therefore, the joint measurements vector $\mathbf{Z}_k = [z_{1,k}^T, z_{2,k}^T, \ldots, z_{n_k,k}^T]^T$ can be defined at time step k. Since each object $o \in \mathcal{O}$ may generate an arbitrary number of noisy measurements, the set $\mathcal{Q}_{o,k} \subseteq \mathcal{Q}_k$ includes the indices $q \in \mathcal{Q}_{o,k}$ of the measurements $z_{q,k}$ pertaining to object o only. In the following, a single measurement result $z_{i,k}$ shall be attributed to one object only.

It is worth emphasising that despite the threshold-based CFAR filtering, the raw radar data may be affected by several outliers, i.e., points due to noise or other random phenomena that are spread quite randomly throughout the field of detection. In the case of multiple targets, not only the valid radar data, but also the amount of outliers tend to grow, thus making data clustering difficult. However, the random events causing such outliers are quite uncorrelated in time, whereas the radar points related to a true target are strongly correlated in both time and space. Thus, if the radar data are accumulated over m consecutive frames, the points associated with existing targets tend to exhibit a good persistence. For this reason, the set of most recent radar data (i.e., those collected in time step intervals $\iota = \{k-m+1, \ldots, k\}$) can return more accurate clustering results, as the radar points associated to existing targets can be distinguished from the outliers, as it will shown in Section 5.2.

3.1. Problem formulation

Given the set of measurement data $\mathbf{Z}_k$ and assuming that each moving object lies on the same two-dimensional plane, our goal is to track the position of one or more targets over time $\forall k = 1, 2, \ldots$, i.e., to estimate paths $t_{o,k} = \{\hat{s}_{o,j}\}_{j=1}^k$, $o = 1, \ldots, n_o$, while minimising the estimation error $\|s_{o,j} - \hat{s}_{o,j}\|$, $\forall j$ [44]. To this end, the measurement data $z_{q,k}$ must be first detected and correctly associated to a given object o (i.e., for $q \in \mathcal{Q}_{o,k}$) [34]. Since the MTT problem based on radar data does not require to place any device on the target, associating the radar data to the correct target is very hard [45,46]. In fact, this is one of the most important contributions of this paper.
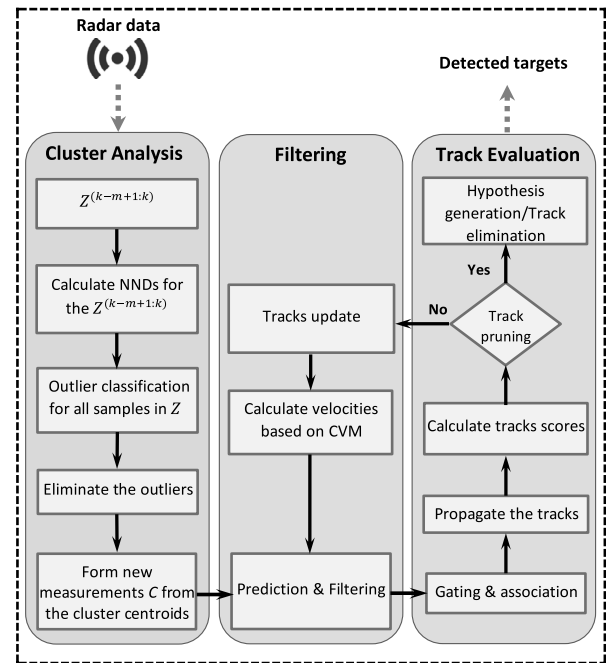

Fig. 1. Overview of the algorithm for multiple targets detection and tracking.

The detection problem aims at finding the elements $\hat{\mathbf{S}}_k$ of the estimated state that best approximate the actual state $\mathbf{S}_k$ in every measurement instant. To formalise the problem, we can consider a square real-valued matrix $W_k : \hat{\mathbf{S}}_k \times \mathbf{S}_k \to \mathbb{R}^{n \times n}$ representing the pairwise Euclidean distance between the centroids of $\hat{\mathbf{S}}_k$ and $\mathbf{S}_k$. To make W_k square, the smaller vector between $\hat{\mathbf{S}}_k$ and $\mathbf{S}_k$ is zero-padded to the length n of the longer set. Also, to find the best association, a linear assignment $f : \hat{\mathbf{S}}_k \to \mathbf{S}_k$ and two index functions $\delta : \mathbf{S}_k \times \{1, \ldots, n_o\}$ and $\hat{\delta} : \hat{\mathbf{S}}_k \times \{1, \ldots, n\}$ must be sought to minimise the cost function

$$J(f, \delta, \hat{\delta}) = \sum_{\hat{s} \in \hat{\mathbf{S}}_k} W_k(\hat{\delta}(\hat{s}), \delta(f(\hat{s}))). \tag{2}$$

Notice that if $n = n_o \geq n_s$, the inverse function $f^{-1} : \mathbf{S}_k \to \hat{\mathbf{S}}_k$ can be simply used in (2).

4. Algorithm description

A flowchart describing the proposed algorithm is shown in Fig. 1. A detailed description of the different functions is provided in the following.

4.1. Cluster analysis

The first three steps of the proposed MTT algorithm are: background (i.e., obstacle points) removal, outliers removal and targets' data clustering. Generally, two main approaches exist to remove the spurious radar points due to clutter and fixed obstacles in a given environment. In the first group of solutions, the region of interest is extracted from the map. As a result, all the radar points that lie outside the region of interest are regarded as obstacle points and are discarded immediately [47]. In the second group of solutions instead, the information about the map of the room (including both obstacles and free space) wherein targets can actually move is used to update the track score. Thus, the localisation process is based on a Markov chain model [48]. In this paper, the former approach is adopted. Of course, some radar observations due to static objects may not be removed and could survive the outlier removal process. However, since these points are usually static and do not belong to any moving target, they are

inherently discarded by the multi-hypotheses approach described in Section 4.3. In fact, sooner or later these points are associated to some false hypothesis that is finally rejected.

To address the outliers removal problem, a supervised classification approach is used in this paper [43]. Let $\mathbf{Z}_t$ be the vector of the m latest radar measurement data collected up to the kth time step. Starting from $\mathbf{Z}_t$, arrays of ordered sequences are built by using the Nearest-Neighbour Distance (NND) algorithm to determine whether a measured point obtained from (1) belongs to a cluster or it is an outlier. As mentioned earlier, due to the specific characteristics of radar signals, the target distance from the radar affects the density of points assigned to the target. Thus, the distance of each point from the sensor is added to the feature space of the input data (i.e., $\bar{r}_{j,k}$ in (1)). Using this input pattern, a standard two-layer Feed-Forward Neural Network (FFNN) classifier with 30 neurons per layer was used to distinguish valid data from outliers. This solution proved to be indeed computationally lighter and more effective than a more complicated Long-Short-Term Memory (LSTM) network, which was instead used in [43] to address a similar outlier detection problem, but in a totally different context and with more observations. The neural network hidden layer uses the Rectified Linear Unit (ReLU) activation function and the output relies on a sigmoid activation function. The choice of the ReLU activation function is motivated by both its effectiveness in promoting sparse representations and its good generality.

The adopted training algorithm is based on the stochastic gradient descent method which is crafted to optimise performance and it is more adaptable than other techniques in the case of binary classification tasks. Of course, the algorithm works if a sufficient number of measured points is available, as discussed later in this paper. Once the outliers are detected and removed from the observations set, the data are then grouped by the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm [49]. This algorithm is preferable to alternative clustering methods because of: (i) its simplicity, efficiency and adaptability; (ii) its robustness to noise, and, last but least, (iii) its ability to identify clusters of varying shapes and sizes with no prior knowledge of the number of clusters [50]. Thus, for a given $\mathbf{Z}_t$, a set of n_c clusters $\mathbf{C}_k = [c_{1,k}^T, \ldots, c_{n_c,k}^T]^T$ is determined at time step k, with $c_{i,k} = [x_{i,k}^c, y_{i,k}^c]^T$ being the centre of each cluster. Each cluster contains a subset of refined radar data.

4.2. Filtering and prediction

In general, the clusters resulting from the previous step do not just coincide with the actual targets. First of all, the number of estimated clusters n_c may differ from the actual amount of objects n_o due to a variety of factors, including radar measurement data quality, radar resolution, the value of n_o itself, the targets' locations $s_{i,k}$ (for $i = 1, \ldots, n_o$) and the distance between them. Moreover, the clustering performance is affected by the multipath propagation of the backscattered signals. As a result of these issues, the estimated centre of the ith cluster tends to deviate from the actual target's position, i.e., the value of $\|s_{i,k} - c_{i,k}\|$ differs from zero and often increases. To mitigate the aforementioned problems, it is required to keep track of the set of consistent hypotheses around $s_{i,k}$ for $i = 1, \ldots, n_o$ and $\forall k$, using the estimated trajectories of each target $t_{i,k}$. As discussed in Section 1, MHT refers to a group of algorithms that can generate numerous hypotheses on the locations of a target. These hypotheses are evaluated by using the data received over time, with some of them being deemed more plausible than others. In the MHT algorithms, the hypotheses are updated using every new set of data. Therefore, the new hypotheses inherit the prior probabilities from their parent hypotheses. The low-probability hypotheses are gradually removed. However, the same measurement data can be associated to different tracks since the algorithm considers multiple hypotheses and, until a track is terminated, a measurement may be part of multiple hypotheses. If some radar-based measurements are lost or no longer associated with the same target and no dynamic models were used to

describe the motion of the moving targets, the evolution of a track would be suddenly lost as well, which is not acceptable. To avoid this problem, a basic dynamic model describing the motion of each target can be used. In this way, even if some measurements are missing, then the evolution of a track can be estimated in open loop. Since trajectory $t_{i,k}$ of the ith target is just a sequence of estimated positions $\widehat{s}_{i,k}$, a standard linear Kalman Filtering (KF) can be used to this purpose, i.e.,

$$\widehat{\tilde{s}}_{k+1} = \widehat{s}_k + \Delta t \begin{bmatrix} v_{x,k} \\ v_{y,k} \end{bmatrix},$$
$$\widehat{\tilde{c}}_{k+1} = H_{k+1}\widehat{\tilde{s}}_{k+1}, \tag{3}$$

where:

- superscript $\cdot^-$ denotes the predicted quantities;
- Δt is the interval between two subsequent radar scans. Without loss of generality, Δt coincides with the sampling period adopted to discretise the system dynamics.
- Since we are assuming that the available measurements at time step $k + 1$ are the predicted cluster positions $\widehat{c}_{k+1}$ derived from the radar data, the matrix H_{k+1} defines a 2×2 linear mapping: for instance, assuming that the cluster matches an actual object in the scene, H_{k+1} is the identity matrix I_2.

Note that index i (referring to the ith target) is omitted in (3) to improve readability. As thoroughly investigated in [51] and empirically verified in [52], the Constant Velocity (CV) model (which relies on the average velocity of a target over the last m steps), although simple, can well describe the motion of a human target [53]. Thus, it will be employed in this work, as well. The covariance matrix of the prediction estimation errors is then given by

$$\mathbf{P}_{k+1}^- = \mathbf{P}_k + \Delta t^2 Q, \tag{4}$$

where $Q = \sigma_q^2 I_2$ is the estimation error covariance matrix of the velocities resulting from the average of the last m steps. Denoting the cluster position measured as explained in Section 4.1 with

$$\bar{c}_{k+1} = c_{k+1} + \eta_{k+1}, \tag{5}$$

where c_{k+1} is the actual cluster position at time step $k + 1$ and η_{k+1} represents the white, zero-mean and normally distributed measurement errors with covariance matrix $R = \sigma_r^2 I_2$, the innovation term based on the estimated cluster position $\widehat{c}_{k+1}$ returned by (3) is

$$\tilde{c}_k = \bar{c}_{k+1} - \widehat{c}_{k+1}. \tag{6}$$

Therefore, the covariance matrix associated with (6) is

$$\Sigma_{k+1} = H_{k+1}\mathbf{P}_{k+1}^- H_{k+1}^T + R, \tag{7}$$

and the Kalman gain is given by

$$K_{k+1} = \mathbf{P}_{k+1}^- H_{k+1}^T \Sigma_{k+1}^{-1}. \tag{8}$$

Finally, the Kalman filter equations for updating the target location and the estimation error covariance matrix are, as usual,

$$\widehat{s}_{k+1} = \widehat{\tilde{s}}_{k+1} + K_{k+1}\tilde{c}_{k+1},$$
$$\mathbf{P}_{k+1} = \mathbf{P}_{k+1}^- - K_{k+1}H_{k+1}\mathbf{P}_{k+1}^-. \tag{9}$$

Unlike a standard Kalman filter, the innovation term (6) and the covariance matrix Σ_{k+1} are employed to compute the Mahalanobis distance

$$\mathbf{M}^2 = \tilde{c}^T \Sigma_{k+1}^{-1} \tilde{c}, \tag{10}$$

which takes into account the difference between the predicted target position and the new measurements. This parameter is then used for efficient gating and target-cluster association, as explained in [52].

4.3. Track evaluation and hypotheses formation

After tracks initialisation and target detection gating and branching, a hypothesis testing for each track must be performed. To this end, the SBMHT algorithm calculates a score for each track in the form of a log-likelihood ratio that relies on the prior probability of target existence, the false alarm rate, and the sequence of detected events [21]. Let $t_{i,k} \in \mathcal{T}$ be the ith track in the set $\mathcal{T}$ for $i = 1, \ldots, n_i$ (with $n_i \leq n_o$) and $h_{i,k} \in [0, 1]$ be the probability that the ith track $t_{i,k}$ is valid at time step k. Therefore, tracks confirmation and elimination are performed continuously at specific time intervals by checking the track scores recursively. The score calculation relies on the Bayes' rule [54], i.e.

$$\Pr\left[h_{i,k}|t_{i,k} \subseteq \mathcal{T}\right] = \frac{\Pr\left[t_{i,k} \subseteq \mathcal{T}|h_{i,k}\right]\Pr\left[h_{i,k}\right]}{\Pr\left[t_{i,k} \subseteq \mathcal{T}\right]}, \tag{11}$$

where $\Pr\left[h_{i,k}|t_{i,k} \subseteq \mathcal{T}\right]$ is the probability that the sequence of observations assigned to the track $t_{i,k}$ represents a true target hypothesis. As explained in [39], such a posterior probability can also be expressed as the Bayesian average

$$\Pr\left[h_{i,k}|t_{i,k} \subseteq \mathcal{T}\right] = \frac{\Gamma_{i,k}\Pr\left[h_{i,k}|t_{i,k-1} \subseteq \mathcal{T}\right]}{1 + (\Gamma_{i,k} - 1)\Pr\left[h_{i,k}|t_{i,k-1} \subseteq \mathcal{T}\right]}, \tag{12}$$

where

$$\Gamma_{i,k} = \begin{cases} \frac{\hat{P}_D e^{-\mathbf{M}^2/2}}{\beta_{FA}(2\pi)\sqrt{|\Sigma_{i,k}|}} & \text{if the measurement is} \\ & \text{assigned to } t_{i,k} \text{ at time step } k \\ 1 - \hat{P}_D & \text{if no measurement is} \\ & \text{assigned to } t_{i,k} \text{ at time step } k \end{cases} \tag{13}$$

In (13), $\hat{P}_D$ is the expected probability of detection, β_{FA} is the false alarm rate, $\mathbf{M}^2$ is the normalised Mahalanobis distance defined in (10) and $|\Sigma_{i,k}|$ is the determinant of the matrix obtained from (7). Note that the first case in (13) can be approximated by [21]

$$\Gamma_{i,k} = \frac{1}{2\pi|\Sigma_{i,k}|^{1/2}} \exp\left[-\frac{\tilde{c}_{i_k}^T \Sigma_{r_k}^{-1} \tilde{c}_{i_k}}{2}\right]. \tag{14}$$

Assuming that the product $\Gamma_{i,k}\Pr\left[h_{i,k}|t_{i,k-1} \subseteq \mathcal{T}\right]$ is negligible, using the log-likelihood form and following the same steps described in [21, 39], (12) can be rewritten recursively as follows, i.e.

$$\mathcal{L}_{i,k} \approx \mathcal{L}_{i,k-1} + \log(\Gamma_{i,k}) \tag{15}$$

where

$$\mathcal{L}_{i,k-1} = \log\left(\frac{\Pr\left[h_{i,k}|t_{i,k-1} \subseteq \mathcal{T}\right]}{1 - \Pr\left[h_{i,k}|t_{i,k-1} \subseteq \mathcal{T}\right]}\right), \tag{16}$$

and the initial score is given by

$$\mathcal{L}_{i_0} = \log\left(\frac{\Pr\left[t_{i_0} \subseteq \mathcal{T}\right]}{1 - \Pr\left[t_{i_0} \subseteq \mathcal{T}\right]}\right) = \log\left(\frac{\beta_N}{\beta_{FA}}\right). \tag{17}$$

Hence, the track probabilities are initialised using spatial densities that are the estimated using the probabilities of a new target detection β_N and a false alarm rate β_{FA}. The initialisation of such probabilities in (17) will be discussed in the next section. The calculated score is finally used for track confirmation or elimination.

4.4. Hypotheses evaluation and pruning

The SBMHT algorithm starts with an empty hypotheses set, which is gradually updated with the detected tracks. The tracks score is sequentially updated and evaluated using (15) and for each time window of length λ (referred to as the *pruning length* in the following), it is compared with a decision threshold τ_L. Therefore, at each time step k and for all the tracks with length $\geq \lambda$, a track can either be pruned
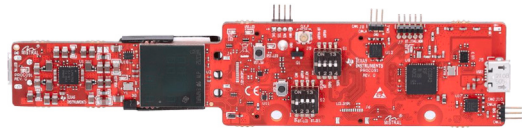

Fig. 2. IWR6843AOPEVM evaluation module [55].

(i.e., $\mathcal{L}_{i,k} \leq \tau_L$) or confirmed (i.e., $\mathcal{L}_{i,k} > \tau_L$). It should be noted that the confirmation of a track does not indicate that the corresponding track is a valid hypothesis, but only that it is consistent with the previously collected data. Moreover, the scores are reinitialised to check if the associated track is consistent with the collected radar data. On the contrary, every time that a track is pruned, it may be considered as a still valid hypothesis depending on its length. Unfortunately, the values of parameters λ and τ_L can just be chosen heuristically. Further details on how these values are set in the specific case study considered in this paper are reported in Section 5.2.

5. Experimental setup and algorithm settings

In this section, first the experimental setup based on the chosen mmWave-FMCW radar (i.e., the Texas Instruments (TI) 60 GHz IWR6843 device) is described. Then, the criteria to set the various parameters of the proposed MTT algorithm are explained.

5.1. Platform, testing environment and data sets

The IWR6843 device is installed on the TI IWR6843AOPEVM USB-powered evaluation module (see Fig. 2). This module is equipped with an advanced Antenna-on-Package (AoP) device. Specifically designed for the 60–64 GHz frequency band, the sensor features a configurable FMCW chirp design. The AoP architecture plays a crucial role in minimising signal losses and simplifying integration, showcasing a technological leap in radar sensing: it includes 4 receiving (RX) and 3 transmitting (TX) antennas, providing a 120° azimuth and a 120° elevation field of view (FoV). Based on the available documentation [56], the maximum radar detection range in line-of-sight conditions is about 49 m and no specific environmental requirements are needed. However, at distances greater than about 10 m, the amount of missing and bad data grows drastically, which makes both multiple-target detection and outlier removal much more challenging. The front-end of the IWR6843 device scans the area in front of the antennas with FMCW chirp signals, that are reflected by possible objects and are finally received as echoes. The frequency shift between the transmitted and back-scattered signals (Doppler shift) is proportional to the distance and the relative velocity of the target. A built-in embedded C674x Digital Signal Processor (DSP) is used for preliminary signal processing. By applying a 2D-FFT, the distance from one or more objects and the related Doppler frequency shift can be estimated in the frequency domain. Since these estimates usually are very noisy, at the output of the 2D-FFT a CFAR threshold filter is applied. To this end, the range-Doppler spectral data are partitioned into cells. Each cell is regarded as a potential target location. The cells surrounding the cell under test are split into guard cells and training cells. The former ones are adjacent to the cell under test and are used to estimate the local noise level, while the latter ones are farther away and are used to estimate the statistical features of the noise. The threshold for detection is computed adaptively by using the average signal level in the guard cells. During the experimental campaign, 15 training cells and 4 guard cells, respectively, were used while collecting the radar data. The CFAR threshold factor was instead set between 1.5 and 2. Such values were tuned heuristically to provide a good trade-off between clutter rejection and target detection sensitivity. Finally, a 3D-FFT (i.e., incorporating the azimuth angle dimension in addition to range and Doppler frequency shift) is applied to the
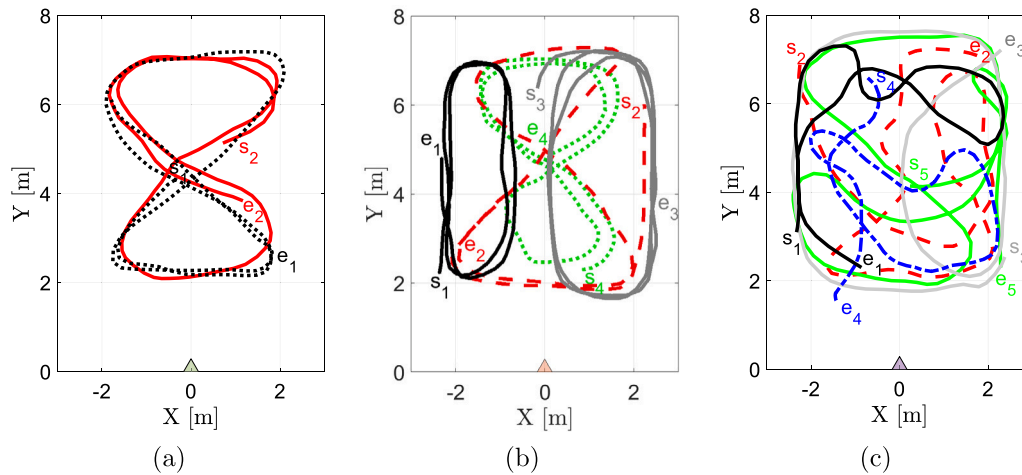
Fig. 3. Three different tracking examples with (a) two (b) four and (c) five people moving simultaneously in the room. The starting and ending points of each path are labelled with letters s and e, respectively. The radar is located in the origin of the Cartesian plane.

filtered data to return the cylindrical coordinates of the cloud of points representing the detected targets. The raw position data are transferred to a PC for further processing at a rate of about 10 Hz.

In the current setup, the TI IWR6843AOPEVM evaluation module was placed on a 2.20 m high post located in the middle of one of the side walls of an empty room monitored by an OptiTrack localisation system equipped with 14 calibrated cameras. The origin of the OptiTrack reference frame was calibrated to make it coincide with the position of the radar post base. The OptiTrack system is able to measure the position of ad-hoc reflective markers placed on moving targets with a standard uncertainty of about 1 mm. The area monitored by both the radar and the Optitrack system was about 50 m^2 wide. According to the data analysis reported in [57], if no spurious target detection occurs, the planar position measurement uncertainty of a single target is about ±35 cm with a 99% confidence level. This value is consistent with the positioning accuracy results obtained in the present multi-target tracking case, as it will be shown in Section 6 (see Fig. 13).

We performed repeated experiments (of duration ranging from 110 s and 270 s) in 6 different scenarios. In each scenario, from 1 up to 5 people move along different paths in the monitored area. Fig. 3(a)–(c) displays three of such experiments involving 2, 4 and 5 people moving simultaneously in the chosen environment. The data sets collected both along these paths and in other 3 scenarios with 1 up to 5 people moving randomly throughout the room (not shown here for the sake of brevity) were used for both algorithm configuration and testing (further details about these aspects are reported in Section 5.2).

5.2. Algorithm parameters configuration

After zero-padding either the sets $\hat{\mathbf{S}}_k$ or $\mathbf{S}_k$ to the longest sequence (as described in Sections 3 and 4), the FFNN mentioned in Section 4.1 was used for outlier detection. While training the FFNN, the key parameters of the proposed algorithm were properly tuned heuristically to maximise outlier detection and to minimise the false target detection rate. Out of the 6 available data sets, those collected with 1, 3 and 5 people moving in the room (18023 observations) were used for training, whereas those collected with 2 and 4 people, as well as the remaining data collected with 5 targets (16388 observations in total) were used for testing. The data sets samples were tagged by the human operator as true or as outliers using the Optitrack system as a reference. Indeed, due to the preliminary CFAR filtering, the number of outliers is limited.

The differences of the Nearest Neighbours Distances (NNDs) between nearby radar points, as well as the amount of nearest neighbours with a small NND value, determine whether a given point belongs to the cluster representing a true target or it is an outlier. Fig. 4 shows
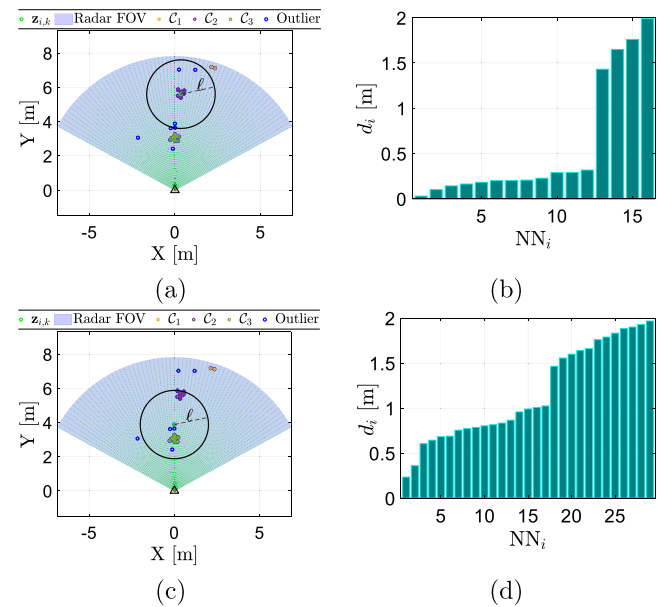

Fig. 4. Qualitative examples or radar scans (on the left side) and corresponding neighbourhood patterns (on the right side) within circles of radius ℓ. The results in (a)–(b) refer to a true cluster observation, whereas those in (c)–(d) refer to the case when an outlier is considered.

two examples of radar position data and the corresponding sorted NND sequences. As can be seen from the Fig. 4(a)–(b), a large number of points with small NND values and a small difference between them (e.g., see the NN_i from 1 to 12 where the maximum difference is 7 cm) is likely to denote a true target. On the contrary, the few NNDs with small values shown in Fig. 4(c)–(d), highlight a critical situation, i.e., a possible outlier, since the considered radar point has only one close neighbour and a large number of distant neighbours. As mentioned earlier, since the spatial density of clusters and outliers depends also on the distance from the sensor, using a neural network can be very effective to detect the outliers in ambiguous and complex indoor scenarios. Thus, one pivotal parameter to be tuned in the proposed framework is the NND radius ℓ. While larger values of ℓ provide more information about the surroundings of each track, the increased neighbourhood size complicates the input pattern and makes FFNN training more challenging due to the dimensionality increase of the chosen features.
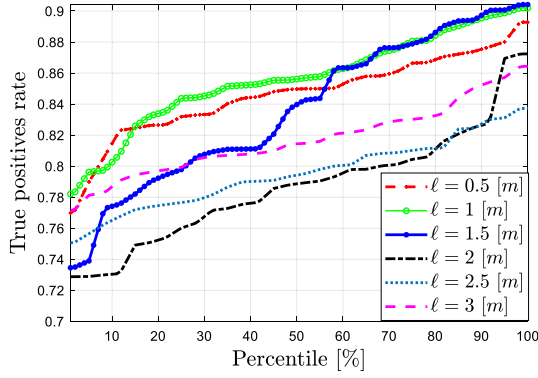
Fig. 5. Empirical true positive outlier detection rate curves for different values of the NND radius ℓ as a function of the percentile of FFNN training data.
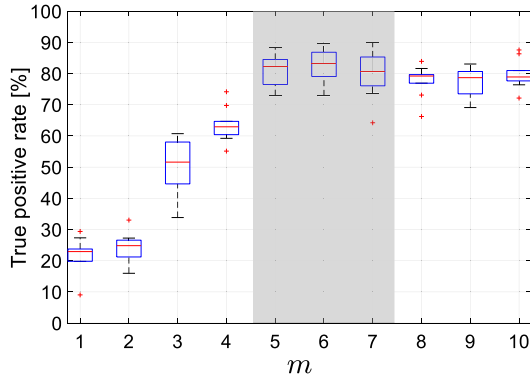

Fig. 6. Box-and-whiskers plot of the true positive outlier detection rate for increasing values of the window length m. The region of interest for the optimal choice of m is highlighted in grey.

Fig. 7. Confusion matrices for outliers detection during FFNN training (a) and testing (b).

To determine an optimal value of ℓ while mitigating the variability due to the limited number of outliers in the considered data set, the FFNN was trained using the Levenberg–Marquardt algorithm multiple times (i.e., about 30 iterations) for increasing values of ℓ. Being theoretically impossible to verify the optimality of ℓ, a heuristic criterion based on the computation of the true positive rate in outlier classification is adopted. Fig. 5 shows the rate of outliers that are correctly detected in the testing phase as a function of the percentile of FFNN trainings for increasing values of ℓ. It can be noticed that both a radius greater than 1.5 m and smaller than 1 m significantly reduce the probability to detect outliers correctly, henceforth ℓ was finally set to 1.5 m. Recalling that, as discussed in Section 3, a single frame is not long enough to accumulate a sufficient amount of data for outlier detection, the optimal value of the window size m was also determined heuristically by training the FFNN model with 10 different subsets of experimental data while increasing the window length. The box-and-whiskers plot of the true positive rate as a function of the window length is shown in Fig. 6. Note that for $m \in [5, 7]$ the best performances are obtained. Based on this results, $m = 6$ consecutive 100-ms-long frames are used for clustering and outlier detection in the rest of this paper. The performance of the FFNN for outlier detection are finally summarised by the confusion matrices shown in Fig. 7(a)–(b) for the training and testing phase, respectively. Approximately, 90% accuracy in detecting the outliers was reached, a number that was deemed effective for the employed FFNN.

Once the outliers are detected, the data are clustered using the DBSCAN algorithm (see Section 4.1). Fig. 8(a)–(c) shows the DBSCAN clustering results for the three training data sets with 2, 4 and 5 targets (16388 observations in total) after outlier detection. To cluster the data
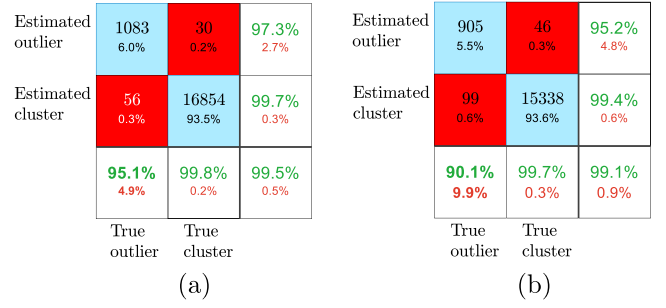
through DBSCAN, the neighbourhood radius $\epsilon^\star$ (which plays a similar role as ℓ for outlier detection) and the density threshold $q^\star$ (i.e., the minimum number of radar points within the neighbourhood radius $\epsilon^\star$ to identify a point as the cluster centre [58]) must be set. Fig. 8(a)–(c) reports the accuracy of DBSCAN clustering on the training data sets for different choices of the corresponding parameters. The results clearly show that the best performances are obtained for $\epsilon^\star = 0.5$ m and $q^\star = 3$. Further results, omitted for the sake of brevity, reveal that if the outliers are not removed before applying the DBSCAN algorithm, not only the best clustering accuracy drops by more than 20%, but also larger values of both $\epsilon^\star$ and $q^\star$ should be used, which negatively affect both sensitivity and selectivity of the clustering algorithm.

In the Kalman Filter described in Section 4.2, the two key parameters to be tuned are σ_q and σ_r, i.e., the process and measurement standard uncertainty values. The former parameter expresses the uncertainty associated with $v_{x,k}$ and $v_{y,k}$ in (3) and it is used to build the covariance matrix Q in (4). The standard deviation σ_r instead refers to the measurement uncertainty of the coordinates of the centroids of the clusters associated with each detected target and it is used to build the covariance matrix R in (7). To compute the values of σ_q and σ_r, the entire algorithm is repeatedly applied to three data sets, while sweeping the values of σ_q and σ_r in the intervals $[0, 8]$ m/s and $[0, 4]$ m, respectively. For each pair of values of σ_q and σ_r, the average absolute error in estimating the true number of targets $\overline{\mathcal{E}_t}$ and the average detection error $\overline{\mathcal{E}_p}$ were estimated over $p = 3600$ time steps (considering all three data sets) for 400 pairs of σ_r and σ_q values:

$$\overline{\mathcal{E}_t} = \frac{1}{p}\sum_{j=1}^{p}|n_j - \hat{n}_j| \qquad \overline{\mathcal{E}_p} = \frac{1}{p}\sum_{j=1}^{p}\left|\left|\mathbf{S}_j - \hat{\mathbf{S}}_j\right|\right|, \qquad (18)$$

where n_j and $\hat{n}_j$ are the numbers of ground truth and detected targets respectively for the jth sample, while $\left|\left|\mathbf{S}_j - \hat{\mathbf{S}}_j\right|\right|$ is the norm of the difference between the ground-truth joint state vector $\mathbf{S}_j$ and $\hat{\mathbf{S}}_j$ as it results from (2). As shown by the error surfaces of $\overline{\mathcal{E}_t}$ and $\overline{\mathcal{E}_p}$ in Fig. 9(a)–(b), both functions reach a minimum value for $\sigma_r = 1.74$ m and $\sigma_q = 8$ m/s. Moreover, the minimum of $\overline{\mathcal{E}_t}$ is global (see Fig. 9(a)). Note that the value of σ_r is much greater than the standard deviation of the positioning error of a single radar measurement (i.e., about 12 cm). However, this setting is needed to keep into account the detrimental effect of multiple targets, as confirmed by the simulation results shown in Fig. 9(a).

The last set of parameters to tune includes both those of the basic SBMHT method described in Section 4.3 and those that should be used for hypotheses evaluation and pruning (see Section 4.4). In this regard, the probability $\hat{P}_D$ in (13) was set to 0.8 by observing the number of clusters correctly assigned to the targets, while the values of parameters $\beta_N = 0.9$ and $\beta_{FA} = 0.1$ (which are needed in (13) and (17)) were directly derived from the results of DBSCAN clustering. Finally, the threshold values to prune the tracks with a low score while keeping those with high scores (namely, the values of parameters λ and τ_L
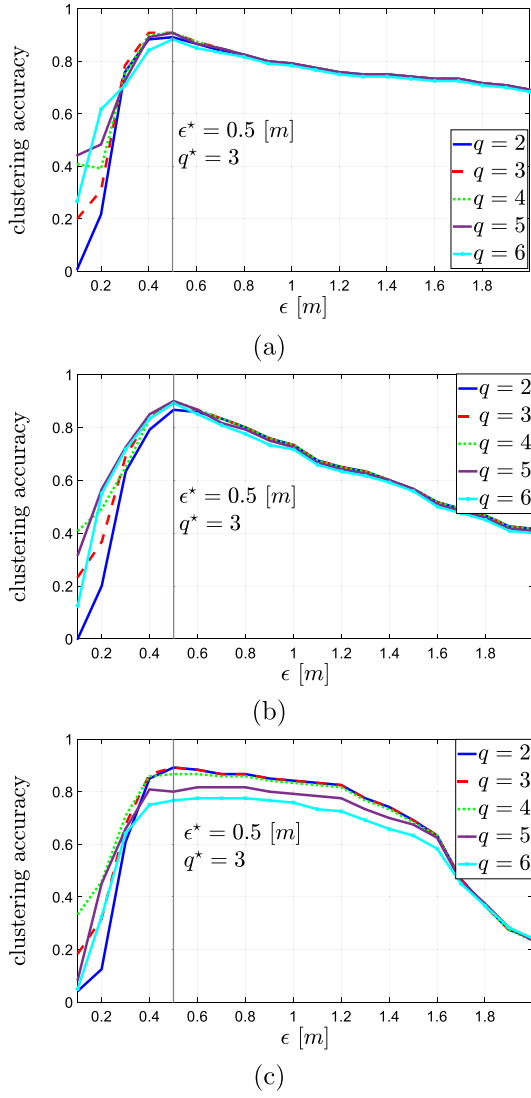
(a)

(b)

(c)

Fig. 8. DBSCAN clustering accuracy after outlier detection. Three data sets per experiment are used in either case, i.e. while tracking 2 (a), 4 (b) or 5 (c) targets. The best values of parameters $\epsilon^\star$ (radius) and $q^\star$ (density threshold) are also highlighted.

(a)

(b)

Fig. 9. Average absolute target detection error (a) and average target position error as a function of increasing values of σ_r and σ_q.

(a)

(b)

Fig. 10. Two alternative example of track scores when three targets are tracked by the SBMHT algorithm. The patterns in (a) show how a properly chosen threshold supports a true hypothesis formation, whereas the results in (b) confirm that excessive threshold values may increase the number of wrong hypotheses.

defined in Section 4.4 for hypotheses evaluation and pruning) were found heuristically in two steps, since to the best of Authors' knowledge no analytical criteria exist to set such parameters. First, the number of frames λ was gradually incremented to find the best trade-off between the detection rate of true tracks (which should be as high as possible) and the false hypotheses detection rate (which instead should be as low as possible). While in [21] the value of λ ranges between 7 and 9, we found that $\lambda = 11$ is the best value for the considered training data set. Afterwards, the value of τ_L was tuned iteratively, till when the best detection performance is achieved for the chosen λ value. In this case $\tau_L = -6$.

Fig. 10 shows the track score of three targets tracked by the SBMHT algorithm as a function of time, for two different values of threshold τ_L. As can be seen from Fig. 10, three valid hypotheses are identified. The scores of the confirmed tracks are reinitialised at the beginning of every time window consisting of $\lambda = 11$ steps. Note that, in the current example, the log-likelihood function of hypotheses $h_{2,k}$ and $h_{3,k}$ drops due to some bad measurement data which affect the update step of the Kalman filter in (9). As a consequence, the filter just relies on the prediction step (see the arrow labelled with "open loop" in Fig. 10). Observe that in Fig. 10(a), the hypothesis $h_{3,k}$ is correctly kept for
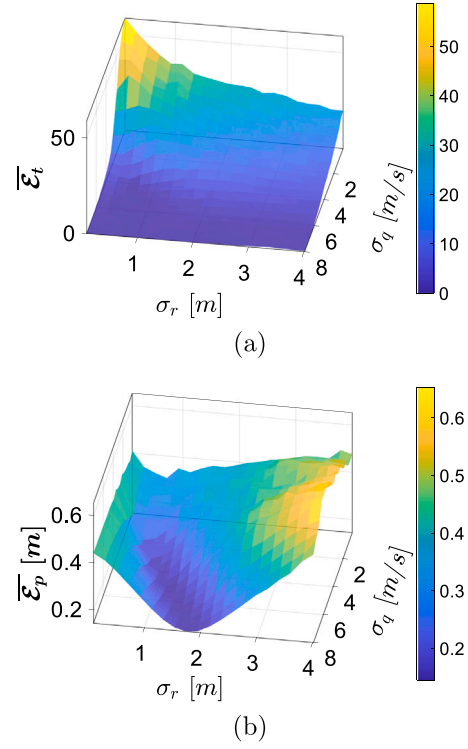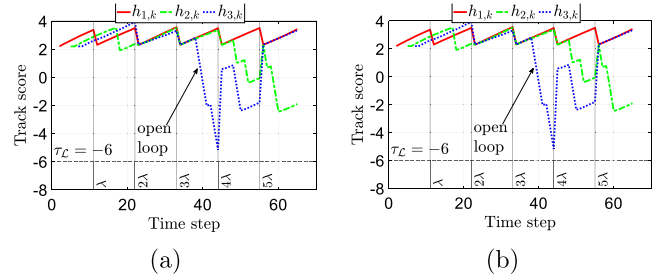
$\tau_L = -6$, while in Fig. 10(b) (that was obtained with a higher threshold, i.e., $\tau_L = -3$) hypothesis $h_{3,k}$ is first pruned and then reinstantiated as hypothesis $h_{4,k}$. This example shows the importance of tuning the key parameters of the SBMHT algorithm.

The list of parameters of the overall MTT algorithm and the chosen values in the considered case study are finally summarised in Table 1. Among them, the most critical for the algorithm performance are: the NDD radius ℓ and the hypotheses pruning parameters λ and τ_L.

6. Experimental results

In the following, we report an experimental characterisation of the proposed approach as well as a comparison with the results obtained with other MHT techniques, i.e., the Global Nearest Neighbour (GNN) multi-object tracker [59], the Joint Integrated Probabilistic Data Association (JIPDA) algorithm [60], and the TOMHT technique [61]. The analysis is based on three of the data sets described in Section 5.1, i.e., those with two, four, and five people, respectively. Crucial to this comparison is the use of the Multi-Object Trackers Toolbox in Matlab,

Table 1
List of parameters of the proposed MTT algorithm.

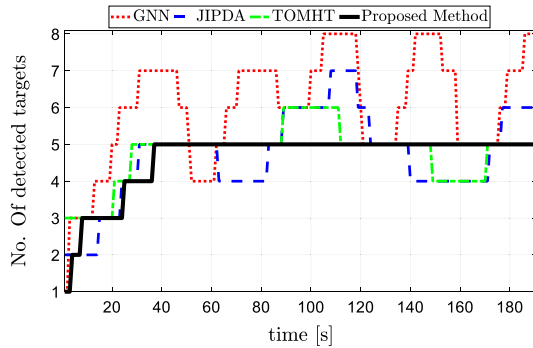| Parameter	Explanation	Value
m	window length	6
ℓ	NND radius	1.5 m
$\epsilon^\star$	DBSCAN neighbour radius	0.5 m
$q^\star$	DBSCAN density threshold	3
σ_q	std. dev. of process noise	8 m/s
σ_r	std. dev. of meas. noise	1.74 m
$\hat{P}_d$	probability of detection	0.8
β_{FA}	probability of false alarm	0.1
β_N	probability of new target	0.9
λ	pruning length	11 frames
τ_L	pruning threshold	−6

Fig. 11. Number of targets detected by different MHT techniques as a function of time when 5 people move randomly in the room.

Fig. 12. Box-and-whiskers plots of the absolute detection errors over multiple experiments using different MHT techniques.

Fig. 13. Empirical cumulative distribution functions of the magnitude of the position error vector computed over three different kinds of users' trajectories.

which includes several built-in algorithms. As shown in [62], there are around 50 parameters to be set for the above-mentioned trackers in the Toolbox. So a full listing of these parameters cannot be reported for a matter of space. Nevertheless, the most important parameters of the benchmark algorithms used for comparison as well as the corresponding values are summarised in Table 2. Such parameters were empirically tuned to achieve the best possible tracking performance with the considered data sets. In all cases, the same KF based on the input constant velocity model was used to describe targets' motion to ensure a fair comparison [51]. Fig. 11 shows a meaningful example, i.e., the number of targets detected by different MHT techniques over time in the most challenging scenario, namely when 5 people move randomly [see Fig. 3(c)]. The fluctuating behaviour of different techniques depends on their sensitivity in accepting or rejecting current and new hypotheses. In particular, a rise in a pattern denotes the acceptance of a given hypothesis (i.e., a possible new target), whereas a decrease in the pattern means the rejection of the hypothesis (i.e., the target is no longer within the sensor range) Using Fig. 3(c) as a reference, the results in Fig. 11 vividly confirm the resilience of the proposed solution in the case of challenging tracking scenarios. Indeed, the fluctuating patterns obtained with the other algorithms highlight their vulnerability in complex situations, especially when the paths of different targets cross each other. On the contrary, the proposed approach exhibits stability in target detection, paving the way to consistent and accurate tracking.

Fig. 12 provides a deeper insight about the detection performance of the considered MHT techniques, as it shows the box-and-whiskers plots of the absolute detection errors over the three data sets used for testing, i.e., with 2, 4 and 5 people moving in the room at the same time. Observe that both the JIPDA and the GNN technique exhibit visible performance limits in detecting the true number of targets, while the detection errors achieved with both the TOMHT technique and the proposed approach are almost negligible, although the TOMHT results are still affected by several outliers. This is confirmed by the mean value of the distribution of the absolute position errors (shown just
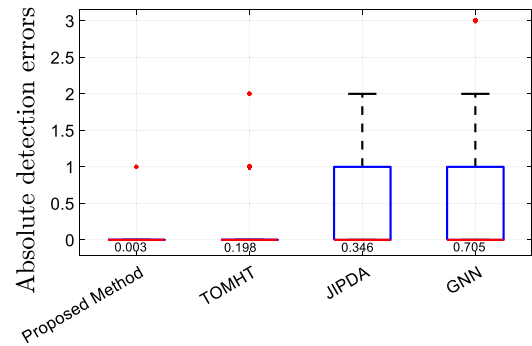
below the boxes) that, in the TOMHT case, is about 60 times greater than when the proposed approach is used.
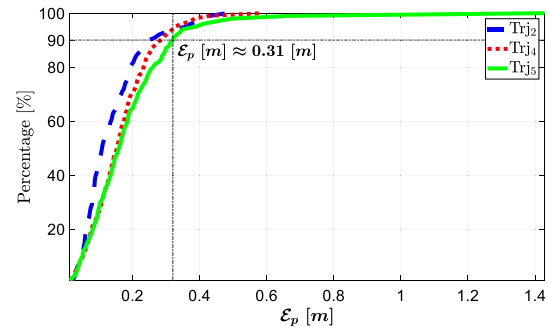
Fig. 13 shows the empirical cumulative distribution functions (CDF) of the magnitude of the position error vector computed over three different kinds of users' trajectories. Observe that the 90th-percentile of the position error is 31 cm. This is an excellent result when up to five people move simultaneously in the same room. Further results (not reported for the sake of brevity) confirm that the positioning accuracy of our approach is much better than the accuracy of the GNN method and it is comparable on average to the TOMHT and the JIPDA algorithms, which nonetheless are affected by more outliers maybe due to the larger multi-target detection uncertainty. Also, quite importantly, the fact that CDF curves in Fig. 13 are almost overlapped suggest that the performance of the proposed algorithm, once its parameters are properly tuned, is rather robust to the increasing number of targets.

As a final comparison, the box-and-whiskers plots of the computation times of the different algorithms are reported in Fig. 14. A PC equipped with an Intel Core i7 CPU running at 2.90 GHz, 32 GB of RAM and Windows 11 as operating system was used to run all tests. The better tracking performance of the TOMHT with respect to the GNN technique is achieved at the price of a much longer processing time. When the JIPDA approach is used instead, a single detection can update multiple tracks at the same time. However, the computational burden reduction is limited by the fact no multiple hypotheses can be kept over multiple scans. As shown in Fig. 14, the proposed algorithm exhibits the smallest mean execution time (i.e., about 100 ms), as well as the lowest variability (i.e., about ±40 ms). The lower computational burden of the SBMHT approach underlying our method is due to the fact that it tends to forget some hypothesised tracks and it does not consider all the hypotheses in every data association (which is preferable whenever the risk of false alarms is rather high). Moreover, the adopted methodology to check and to prune the branches associated with the various hypotheses sequentially is particularly effective.

Table 2

List of the most important parameters of other MTT algorithms implemented in the Multi-Object Trackers MATLAB Toolbox.

| Parameter	Description	TOMHT	JIPDA	GNN
Confirmation Threshold	Minimum score required to confirm a track	16	0.95	16
Deletion Threshold	Maximum score drop for track deletion	−7	.1	−7
Detection Threshold	Probability of detection used in track score	0.9	≥ 1	.9
False-alarm Rate	False alarm probability used for track scoring	0.1	–	0.1
Min. Branch	Minimum probability required to keep a track	0.003	–	–
Pruning Method	Technique to prune branches of the same track	Hyp. testing	None	None
Filter Function	Function used by the tracker to propagate the tracks	CVEKF[a]	CVEKF[a]	CVEKF[a]
Branch Length	Maximum number of track branches per track	9	None	None
Gate	Threshold for target assignment	1.3	3.2	3.2

[a] CVEKF = Constant Velocity Extended Kalman Filter.
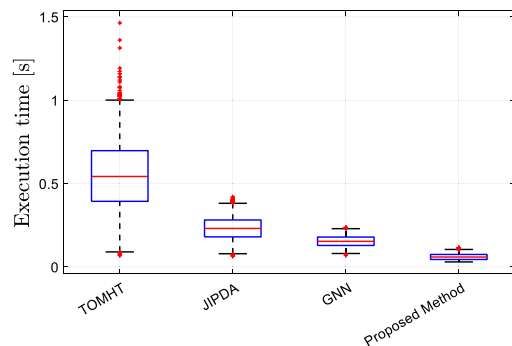
Fig. 14. Box-and-whiskers plots of the execution times of different MHT algorithms.

Considering that the reported execution times of all algorithms were obtained in MATLAB using a single core of an Intel Core i7 microprocessor, considerable performance improvements on a cheaper embedded platform could be achieved if the algorithm were implemented using a lower-level programming language (e.g., in C/C++). In this case, the real-time processing of streams of radar data refreshed every 100 ms is definitely feasible.

7. Conclusions and future work

This paper presents an algorithm to detect and to track multiple targets in indoor environments using the data collected from a mmWave-FMCW radar. The key and most important feature of the proposed solution is its high accuracy in detecting and tracking multiple targets. Detection accuracy and robustness in complex scenarios (i.e., when multiple targets move simultaneously in the same room) are indeed superior to those of other state-of-the-art algorithms. This important result is achieved through a more effective clustering approach (able to discard a large amount of outliers) and by a final Structured Branching Multiple Hypothesis Testing (SBMHT) algorithm, which keeps the risks of both tracking inexistent targets and suddenly losing the existing ones extremely low, i.e., well below 1%. On the contrary, when other MHT techniques are used, target detection errors of about ±1 units are rather likely. As a consequence, the targets' planar positioning uncertainty may be greatly affected by such detection errors as well. Instead, the planar positioning accuracy of the proposed solution using a 60-GHz TI IWR6843 device as a radar platform and a calibrated Optitrack vision system as reference is within about 30 cm with 90% probability when up to five targets move across the same room. Moreover, the proposed approach is faster and more deterministic than other existing MHT algorithms. This is indeed a relevant achievement in view of a possible embedded implementation, which is indeed the next goal, possibly supported by some companies interested in the practical applications of a fully integrated device. Future research directions will also explore the possibility to include the data collected from other sensors (e.g., RFID passive tags or UWB signals) to perform not only target detection, but also identification. In this way, the risk of tracking a wrong target may be further decreased.

CRediT authorship contribution statement

Farhad Shamsfakhr: Conceptualization, Data curation, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **David Macii:** Writing – original draft, Visualization, Validation, Resources, Methodology. **Luigi Palopoli:** Writing – review & editing, Supervision, Methodology, Funding acquisition. **Michele Corrà:** Investigation, Resources, Writing – review & editing. **Alessandro Ferrari:** Investigation, Resources, Writing – review & editing. **Daniele Fontanelli:** Writing – original draft, Supervision, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] N.M. Tiglao, M. Alipio, R. Dela Cruz, F. Bokhari, S. Rauf, S.A. Khan, Smartphone-based indoor localization techniques: State-of-the-art and classification, Measurement 179 (2021) 109349, http://dx.doi.org/10.1016/j.measurement.2021.109349.

[2] B.N. Ozdemir, A. Ceylan, Constructing a precise radio map and application of indoor positioning with dual-frequency Wi-Fi fingerprinting method, Measurement 163 (2020) 107997, http://dx.doi.org/10.1016/j.measurement.2020.107997.

[3] V. Brunacci, A. De Angelis, G. Costante, P. Carbone, Development and analysis of a UWB relative localization system, IEEE Trans. Instrum. Meas. 72 (2023) 1–13, http://dx.doi.org/10.1109/TIM.2023.3305661.

[4] S. Boonsriwai, A. Apavatjrut, Indoor WIFI localization on mobile devices, in: 2013 10th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, IEEE, 2013, pp. 1–5.

[5] S. Sadowski, P. Spachos, K.N. Plataniotis, Memoryless techniques and wireless technologies for indoor localization with the internet of things, IEEE Internet Things J. 7 (11) (2020) 10996–11005.

[6] V. Magnago, P. Corbalán, G. Picco, L. Palopoli, D. Fontanelli, Robot localization via odometry-assisted ultra-wideband ranging with stochastic guarantees, in: Proc. IEEE/RSJ International Conference on Intelligent Robots and System, IROS, IEEE, Macao, China, 2019, pp. 1607–1613, http://dx.doi.org/10.1109/IROS40897.2019.8968019.

[7] L. Santoro, M. Nardello, D. Brunelli, D. Fontanelli, UWB-based indoor positioning system with infinite scalability, IEEE Trans. Instrum. Meas. 72 (1005711) (2023) 1–11, http://dx.doi.org/10.1109/TIM.2023.3282299.

[8] G. Cerro, L. Ferrigno, M. Laracca, G. Miele, F. Milano, V. Pingerna, UWB-based indoor localization: How to optimally design the operating setup? IEEE Trans. Instrum. Meas. 71 (2022) 1–12, http://dx.doi.org/10.1109/TIM.2022.3194924.

[9] J. Wang, Q. Gao, M. Pan, Y. Fang, Device-free wireless sensing: Challenges, opportunities, and applications, IEEE Netw. 32 (2) (2018) 132–137.

[10] R. Zhang, X. Jing, S. Wu, C. Jiang, J. Mu, F.R. Yu, Device-free wireless sensing for human detection: The deep learning perspective, IEEE Internet Things J. 8 (4) (2020) 2517–2539.

[11] L. Piotrowsky, S. Kueppers, T. Jaeschke, N. Pohl, Distance measurement using mmWave radar: Micron accuracy at medium range, IEEE Trans. Microw. Theory Tech. 70 (11) (2022) 5259–5270, http://dx.doi.org/10.1109/TMTT.2022. 3195235.

[12] W. Xue, R. Wang, L. Liu, D. Wu, Accurate multi-target vital signs detection method for FMCW radar, Measurement 223 (2023) 113715, http://dx.doi.org/ 10.1016/j.measurement.2023.113715.

[13] W. Li, R. Chen, Y. Wu, H. Zhou, Indoor positioning system using a single-chip millimeter wave radar, IEEE Sens. J. 23 (5) (2023) 5232–5242, http://dx.doi.org/10.1109/JSEN.2023.3235700.

[14] L. Yang, J. Cao, W. Zhu, S. Tang, Accurate and efficient object tracking based on passive RFID, IEEE Trans. Mob. Comput. 14 (11) (2015) 2188–2200.

[15] F. Shamsfakhr, A. Motroni, L. Palopoli, A. Buffi, P. Nepa, D. Fontanelli, Robot localisation using UHF-RFID tags: A Kalman smoother approach, Sensors 21 (3) (2021) 717.

[16] F. Shamsfakhr, D. Macii, D. Fontanelli, A. Motroni, P. Nepa, L. Palopoli, A. Buffi, RFID-based robot localisation: An unconstrained optimisation problem by exploiting RSSI, in: 2022 IEEE International Instrumentation and Measurement Technology Conference, I2MTC, IEEE, 2022, pp. 1–6.

[17] Z. Zhang, X. Wang, D. Huang, X. Fang, M. Zhou, Y. Zhang, MRPT: Millimeter-wave radar-based pedestrian trajectory tracking for autonomous urban driving, IEEE Trans. Instrum. Meas. 71 (2021) 1–17.

[18] W. Piao, X. Zhao, Z. Li, MMGraphSLAM: Autonomous indoor positioning based on millimeter wave GraphSLAM, Measurement 220 (2023) 113300, http://dx.doi.org/10.1016/j.measurement.2023.113300.

[19] J. Jung, S. Lim, E. Lee, M.-O. Heo, S.-C. Kim, Localization based on distributed heterogeneous radar sensor network, Measurement 204 (2022) 112042, http://dx.doi.org/10.1016/j.measurement.2022.112042.

[20] J. Chen, Z. Xie, P. Dames, The semantic PHD filter for multi-class target tracking: From theory to practice, Robot. Auton. Syst. 149 (2022) 103947.

[21] J.R. Werthmann, Step-by-step description of a computationally efficient version of multiple hypothesis tracking, in: Signal and Data Processing of Small Targets 1992, vol. 1698, SPIE, 1992, pp. 288–300.

[22] H. Rohling, Radar CFAR thresholding in clutter and multiple target situations, IEEE transactions on aerospace and electronic systems 1983 (4) (1983) 608–621.

[23] X. Huang, H. Cheena, A. Thomas, J.K. Tsoi, Indoor detection and tracking of people using mmwave sensor, J. Sensors 2021 (2021) 1–14.

[24] P.J. Rousseeuw, M. Hubert, Robust Statistics for Outlier Detection, Wiley Interdiscipl. Rev. Data Mining Knowl. Discover. 1 (1) (2011) 73–79.

[25] Y. Zhang, N. Meratnia, P. Havinga, Outlier Detection Techniques for Wireless Sensor Networks: A survey, IEEE Commun. Surveys Tutor. 12 (2) (2010) 159–170.

[26] D. Kellner, J. Klappstein, K. Dietmayer, Grid-based DBSCAN for clustering extended objects in radar data, in: 2012 IEEE Intelligent Vehicles Symposium, IEEE, 2012, pp. 365–370.

[27] Y. Zhang, H. Liu, W. Fu, H. Deng, Multi-radar tracking based on weighted K-means clustering fusion, in: 2008 IEEE International Conference on Granular Computing, IEEE, 2008, pp. 813–816.

[28] I. Bilik, J. Tabrikian, A. Cohen, GMM-based target classification for ground surveillance Doppler radar, IEEE Trans. Aerosp. Electron. Syst. 42 (1) (2006) 267–278.

[29] Z. Kang, S.J. Landry, An Eye Movement Analysis Algorithm for a Multielement Target Tracking Task: Maximum Transition-based Agglomerative Hierarchical Clustering, IEEE Trans Hum. Mach. Syst. 45 (1) (2014) 13–24.

[30] J. Pegoraro, M. Rossi, Real-time people tracking and identification from sparse mm-wave radar point-clouds, IEEE Access 9 (2021) 78504–78520.

[31] O. Schumann, M. Hahn, J. Dickmann, C. Wöhler, Supervised clustering for radar applications: On the way to radar instance segmentation, in: 2018 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility, ICMIM, IEEE, 2018, pp. 1–4.

[32] S.S. Blackman, Multiple Hypothesis Tracking for Multiple Target Tracking, IEEE Aerosp. Electron. Syst. Mag. 19 (1) (2004) 5–18.

[33] D. Musicki, B. La Scala, Multi-target tracking in clutter without measurement assignment, IEEE Trans. Aerosp. Electron. Syst. 44 (3) (2008) 877–896.

[34] W. Luo, J. Xing, A. Milan, X. Zhang, W. Liu, T.-K. Kim, Multiple object tracking: A literature review, Artif. Intell. 293 (2021) 103448.

[35] J. Roy, N. Duclos-Hindie, D. Dessureault, Efficient cluster management algorithm for multiple-hypothesis tracking, in: Signal and Data Processing of Small Targets 1997, vol. 3163, SPIE, 1997, pp. 301–313.

[36] M. Fontana, Á.F. García-Fenández, S. Maskell, Bernoulli merging for the Poisson multi-Bernoulli mixture filter, in: 2020 IEEE 23rd International Conference on Information Fusion, FUSION, IEEE, 2020, pp. 1–8.

[37] S. He, H.-S. Shin, A. Tsourdos, Multi-sensor multi-target tracking using domain knowledge and clustering, IEEE Sens. J. 18 (19) (2018) 8074–8084.

[38] T. Sathyan, T.-J. Chin, S. Arulampalam, D. Suter, A multiple hypothesis tracker for multitarget tracking with multiple simultaneous measurements, IEEE J. Sel. Top. Sign. Proces. 7 (3) (2013) 448–460.

[39] S.S. Blackman, Multiple-Target Tracking with Radar Applications, Dedham, 1986.

[40] L. Cao, D. Zheng, Z. Zhao, T. Wang, D. Wang, C. Fu, J. Gu, Convex variational inference for multi-hypothesis fractional belief propagation based data association in multiple target tracking systems, IEEE Sens. J. 21 (17) (2021) 19121–19133.

[41] A. Smiti, A critical overview of outlier detection methods, Comp. Sci. Rev. 38 (2020) 100306.

[42] R.K. Blashfield, M.S. Aldenderfer, The Literature on Cluster Analysis, Multivar. Behav. Res. 13 (3) (1978) 271–295.

[43] D.J. Williamson, G.L. Burn, S. Simoncelli, J. Griffié, R. Peters, D.M. Davis, D.M. Owen, Machine Learning for Cluster Analysis of Localization Microscopy Data, Nat. Commun. 11 (1) (2020) 1493.

[44] A. Yilmaz, O. Javed, M. Shah, Object tracking: A survey, Acm Comput. Surveys (CSUR) 38 (4) (2006) 13–es.

[45] X. Wang, T. Li, S. Sun, J.M. Corchado, A survey of recent advances in particle filters and remaining challenges for multitarget tracking, Sensors 17 (12) (2017) 2707.

[46] A. Milan, K. Schindler, S. Roth, Challenges of ground truth evaluation of multi-target tracking, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2013, pp. 735–742.

[47] G. Gennarelli, G. Vivone, P. Braca, F. Soldovieri, M.G. Amin, Multiple extended target tracking for through-wall radars, IEEE Trans. Geosci. Remote Sens. 53 (12) (2015) 6482–6494, http://dx.doi.org/10.1109/TGRS.2015.2441957.

[48] Q. Yu, G. Medioni, I. Cohen, Multiple target tracking using spatio-temporal markov chain monte carlo data association, in: 2007 IEEE Conference on Computer Vision and Pattern Recognition, 2007, pp. 1–8, http://dx.doi.org/10.1109/CVPR.2007.382991.

[49] R.J. Campello, D. Moulavi, A. Zimek, J. Sander, Hierarchical density estimates for data clustering, visualization, and outlier detection, ACM Trans. Knowl. Discover. Data (TKDD) 10 (1) (2015) 1–51.

[50] E. Schubert, J. Sander, M. Ester, H.P. Kriegel, X. Xu, DBSCAN revisited, revisited: Why and how you should (still) use DBSCAN, ACM Trans. Database Syst. 42 (3) (2017) http://dx.doi.org/10.1145/3068335.

[51] C. Schöller, V. Aravantinos, F. Lay, A. Knoll, What the constant velocity model can teach us about pedestrian motion prediction, IEEE Robot. Autom. Lett. 5 (2) (2020) 1696–1703.

[52] F. Shamsfakhr, M. Corrà, A. Ferrari, D. Macii, L. Palopoli, D. Fontanelli, Algorithms for enhanced indoor positioning and tracking based on a 60-GHz radar platform, in: 2022 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd4. 0 & IoT), IEEE, 2022, pp. 184–189.

[53] A. Antonucci, G. Rosati Papini, P. Bevilacqua, L. Palopoli, D. Fontanelli, Efficient prediction of human motion for real-time robotics applications with physics-inspired neural networks, IEEE Access 10 (2021) 144–157, http://dx.doi.org/10.1109/ACCESS.2021.3138614.

[54] C. Kim, F. Li, A. Ciptadi, J.M. Rehg, Multiple hypothesis tracking revisited, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 4696–4704.

[55] IWR6843AOP evaluation module, 2022, https://www.ti.com/tool/IWR6843AOPEVM, (Accessed: 30 September 2010).

[56] 60 GHz mmwave sensor EVMs description, 2022, https://www.ti.com/lit/ug/swru546e/swru546e.pdf?ts=1686903092404&ref_url=https%253A%252F%252Fwww.google.com%252F. (Accessed: 30 September 2010).

[57] A. Antonucci, M. Corrà, A. Ferrari, D. Fontanelli, E. Fusari, D. Macii, L. Palopoli, Performance analysis of a 60-GHz radar for indoor positioning and tracking, in: 2019 International Conference on Indoor Positioning and Indoor Navigation, IPIN, 2019, http://dx.doi.org/10.1109/IPIN.2019.8911764.

[58] K. Khan, S.U. Rehman, K. Aziz, S. Fong, S. Sarasvady, DBSCAN: Past, present and future, in: The Fifth International Conference on the Applications of Digital Information and Web Technologies, ICADIWT 2014, IEEE, 2014, pp. 232–238.

[59] A. Sinha, Z. Ding, T. Kirubarajan, M. Farooq, Track quality based multitarget tracking approach for global nearest-neighbor association, IEEE Trans. Aerosp. Electron. Syst. 48 (2) (2012) 1179–1191.

[60] S.H. Rezatofighi, A. Milan, Z. Zhang, Q. Shi, A. Dick, I. Reid, Joint probabilistic data association revisited, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 3047–3055.

[61] S.S. Blackman, R. Popoli, Design and Analysis of Modern Tracking Systems, Artech House, 1999.

[62] Trackertomht, 2023, https://it.mathworks.com/help/fusion/ref/trackertomht-system-object.html#responsive_offcanvas. (Accessed: 30 September 2015).