IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received December 22, 2016, accepted January 17, 2017, date of publication February 13, 2017, date of current version March 13, 2017.

Digital Object Identifier 10.1109/ACCESS.2017.2667720

Radar Signal Processing for Jointly Estimating Tracks and Micro-Doppler Signatures

THOMAS WAGNER, (Graduate Student Member, IEEE), REINHARD FEGER, (Member, IEEE), AND ANDREAS STELZER, (Member, IEEE)

Institute for Communications Engineering and RF-Systems, Johannes Kepler University Linz, 4040 Linz, Austria

Corresponding author: T. Wagner (thomas.wagner@jku.at)

This work was supported by the Austrian COMET-K2 Programme of the Linz Center of Mechatronics (LCM), and was funded by the Austrian federal government and the federal state of Upper Austria.

ABSTRACT The aim of the radar systems is to collect information about their surroundings. In many scenarios besides static targets there are numerous moving objects with very different characteristics, such as extent, movement behavior or micro-Doppler spread. It would be most desirable to have algorithms that extract all information on static and moving object automatically, without a system operator. In this paper, we present measurements conducted with a commercially available high-resolution multi-channel linear frequency-modulated continuous-wave radar and algorithms that do not only produce radar images but a description of the scenario on a higher level. After conventional spectrum estimation and thresholding, we present a clustering stage that combines individual detections and generates representations of each target individually. This stage is followed by a Kalman filter based multi-target tracking block. The tracker allows us to follow each target and collect its properties over time. With this method of jointly estimating tracks and characteristics of each individual target in a scenario, inputs for classifiers can be generated. Which, in turn, will be able to generate information that could be used for driver assistance or alarm trigger systems.

INDEX TERMS Automotive radar, surveillance radar, range-Doppler, clustering, tracking.

I. INTRODUCTION

Starting a few years ago, radar sensors were introduced in civil surveillance tasks, mainly in the following three different fields.

1) One application area is automotive safety. Demands in automotive scenarios, especially towards protection of vulnerable road users like pedestrians and bicyclist, have significantly increased in the last few years. Therefore, cars are equipped with radar sensors which are used for driver assistance systems. These sensors enable the driver assistance system to avoid crashes if obstacles are detected in front of the cars. In addition to the detection of immediate obstacles, more elaborate analysis should be able to detect and track the surrounding of the vehicle. This additional information allows a prediction whether a further traffic participant or a vulnerable road user might become an obstacle in the near future [1].

2) Another field of application for radar sensors is making road-side infrastructure intelligent. Some countries, e.g., the United Kingdom, migrate from time-controlled, i.e. pedestrian light controlled (pelican), to the demand-controlled, i.e. pedestrian user friendly intelligent (puffin), crossing for intersections [2], [3]. An efficient traffic control with fully automated traffic lights for pedestrian crossings requires surveillance of traffic participants. To reduce traffic hold-ups, curb-side sensors are required to detect when a pedestrian intends to cross the road and when the pedestrian has crossed over.

Street light control systems could use the same principle. With the evolution towards LED-based illumination more complex light control systems rather than simple sunset based toggling are being considered [4]–[6]. During the night, lamps are dimmed to save energy and reduce light pollution. However, in situations where traffic participants are in the dark, the street lights should automatically turn to full brightness.

3) Comparable demands are expected for property surveillance. Due to their low price, many existing surveillance systems rely on infrared based motion sensors. These suffer from a high false alarm rate, since they can be triggered by moving foliage or animals like roaming cats. Due to mass production for the automotive industry, radar sensors are becoming affordable for other commercial applications. Using such sensors also in surveillance systems, would have

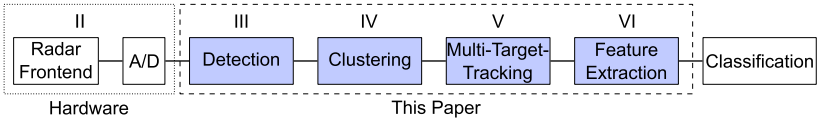

FIGURE 1. Block diagram from sampling the IF data to classification. The numbers above the blocks correspond to the sections in this paper.

the additional advantage of determining the cause of motion detection triggers. False alarm rate can therefore be lowered compared to conventional motion detection systems.

Radar sensor based systems can also address privacy issues which are crucial to any surveillance system. Camera based surveillance does not only allow tracking and classification, but also identification of individuals, e.g., by facial recognition. Following the 'privacy by design' approach, individuals cannot be identified by a radar sensor.

In all of the three cases described above, it is not possible to employ classical operator based radar systems, since there is either not enough reaction time (crash avoidance in automotive scenario) or it is too impractical. Therefore, there is demand for radar systems which can not only visualize raw data but automatically process the data to provide a description of scenario and the individual targets within. This description should not only contain the position of the targets over time but also information suitable for classification.

High resolution radar measurements, also called micro-Doppler and/or micro-range systems, allow for extracting target properties, i.e. signatures, for classification. These signatures are used for classification, e.g., by using support vector machines [7], empirical mode decomposition [8], or statistical pattern recognition [9]. Signatures suitable for classification include extent, range-rate, and various statistics of the micro-Doppler spread like mean of the Doppler signal, standard deviation (STD) of the Doppler signal, total bandwidth of the Doppler signal, normalized STD of the Doppler signal, average Doppler frequency, periodicities in the Doppler signal, as well as range-weighted target energy [10]. Heuel *et al.* [11] used the extent in the range/Doppler map for classification of pedestrians. Bartsch *et al.* [12] used the physical extent and Doppler spread from a single range/Doppler/Direction of Arrival (DOA) measurement. Models of a human's gait are broadly available [13]–[17]. References [7], [8], [10] also analyzed various other activities besides walking. In [18], Doppler spread measurements originating from scans at different frequencies are used. A comparison of the micro-Doppler signatures of trees in wind to pedestrians is presented in [19]. Also, micro-Doppler characteristics of unmanned aerial vehicles (UAVs) have been studied [20], [21] and successfully distinguished from the signature of birds [22]. Recently, bicyclists have been studied in [23].

A basic underlying assumption in all of the aforementioned publications is the presence of a single contributing target. In an actual scenario, however, it is more likely that several targets are present in the field-of-view simultaneously. Thus, a mapping of detections to targets must be made prior to feature extraction. This mapping can be done efficiently within a multi-target tracking (MTT) framework. However, a high-resolution measurement in either range (due to high bandwidth), range-rate (due to long observation time) or DOA (due to large aperture) will produce multiple detections per physical target. Having multiple detections per target violates a basic assumption of radar based tracking systems, namely that a target should be only visible in one cell at a time, i.e. it produces a single detection at each time step. Cluster algorithms come into play in order to generate information of each independent physical target out of the detection list at each time step.

In this paper, we fill the gap between sampling data and feeding the classifier by presenting algorithms to condense suitable inputs for the available classification methods using raw analog-to-digital converter (ADC) data of a radar sensor. Fig. 1 depicts an overview of the necessary processing steps from data-sampling to classification for the case of multiple targets, each consisting of multiple detections, being present simultaneously in a scenario. Moreover, we present results of our processing technique based on measurement data from a commercially available multichannel linear frequency-modulated continuous-wave (LFMCW) radar system with a uniform linear array (ULA). The presentation of the algorithms in this paper is organized as follows:

- measurement and data preparation in Section II,
- spectrum calculation and thresholding to get a list of detections in Section III,
- clustering of detections to represent physical targets in Section IV,
- tracking with cluster information in Section V,
- feature extraction in Section VI, and
- track analysis in Section VII.

In Section VIII, we will address the computational demand of our framework and draw a conclusion in Section IX.

II. MEASUREMENT PRINCIPLE AND SCENARIO

We use a 77-GHz multichannel LFMCW radar platform as depicted in Fig. 2. The platform is similar to [24] and based on Infineon's monolithic microwave integrated circuits (MMICs) [25]–[27] and baseband technology from Inras GmbH [28]. The RF frontend consists of four transmit channels and a linear antenna array for reception with $N_{\mathrm{A}} = 8$ antennas. A single TX channel was used in this

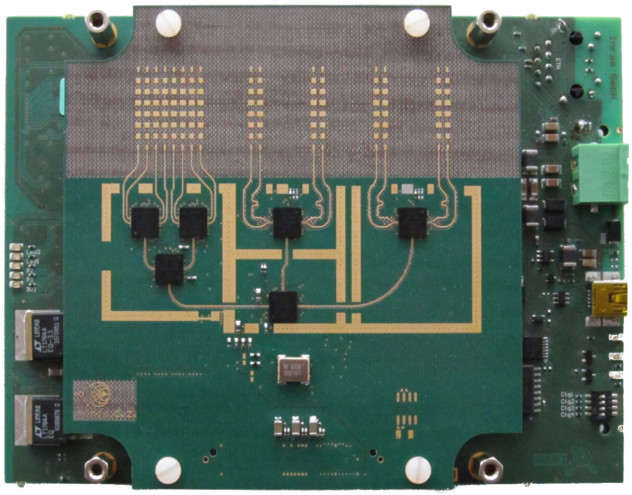

FIGURE 2. Image of the radar system. The PCB on top is the RF frontend which holds the RF components, e.g., chips, distribution networks, and antennas. The board at the bottom is the IF board comprising the power supply, the ADCs, the control units for the RF chips and the USB connection to a PC.

application since multiple channel could only operate in time division multiple access (TDMA) mode. Using TDMA would reduce the unambiguous range-rate region.

We carried out several measurements in different scenarios of which four will be presented in this paper:

A) an outdoor measurement in a pedestrian area,
B) an outdoor measurement of a cat in a backyard,
C) an indoor measurement of a small UAV, and
D) a measurement in an underground parking lot.

For the measurements in the underground parking lot we had access to a hardware revision with support for higher data-rates than the above-ground measurements. Values for the faster hardware are listed in parenthesis. The following parameters were used for the LFMCW measurements: $B = 1\,\text{GHz}$ sweep bandwidth, $3.\dot{3}\,\text{MHz}$ ($10\,\text{MHz}$) sample rate for each receive channel, $64\,\mu\text{s}$ chirp duration of an up-chirp resulting in $N = 210$ (620) samples per up-chirp, and $T_\text{c} = 130\,\mu\text{s}$ ($97\,\mu\text{s}$) chirp repetition interval.

We used $N_\text{C} = 200$ ($N_\text{C} = 268$) consecutive chirps to form a data cube with a size of $N \times N_\text{C} \times N_\text{A}$ at each time step, where N is the number of samples during one chirp and N_A is the number of antennas. This corresponds to an observation interval of approximately 26 ms, for each range/Doppler/DOA map. We have chosen the update interval to be equal to the observation interval, resulting in 38.5 data cubes per second.

With aforementioned parameters, the maximum range is approximately 16 m (48 m) and the unambiguous region in range-rate is approximately $\pm 7.5\,\text{m/s}$ ($10.2\,\text{m/s}$). The bin-width after a fast Fourier transform (FFT) is $\Delta R = 15\,\text{cm}$ in range, $\Delta v = 7.5\,\text{cm/s}$ in range-rate, and 22.5° in DOA at broad-sight. Due to the measurement rate of 38.5 Hz, variations of the scenario, e.g., target motions, can be observed up to the unambiguous frequency of $1/2 \cdot 38.5\,\text{Hz} \approx 19\,\text{Hz}$. These values have been chosen to fit well to the available models of pedestrians [13]–[16], which state an average velocity of $1.5-3\,\text{m/s}$ with peaks of approximately $\pm 3\,\text{m/s}$ around the average and a step frequency of approximately 2 Hz.

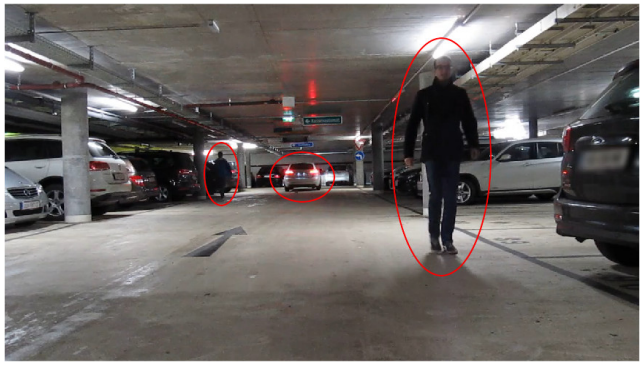

FIGURE 3. Camera view of the measurement scenario used for illustration. In the underground parking lot at the university, two persons are walking next to parked cars while another vehicle is passing by.

With this setup we conducted experiments in various scenarios. The results of these are presented in detail in Section VII. For illustration of the algorithms presented in the following sections, we have chosen the scenario where two pedestrians and a car are moving in an underground parking lot. Fig. 3 shows a picture of this measurement scenario.

III. SPECTRUM AND THRESHOLDING

According to the LFMCW principle, the simplified [29] intermediate frequency (IF) signal model for a single point target in the three-dimensional data cube of the k-th observation interval can be approximated as three-dimensional complex exponential function

$$s_{\text{IF},k}[n, n_\text{C}, n_\text{A}] = A\mathrm{e}^{\mathrm{j}2\pi\left(\phi + \psi_\text{R} n + \psi_\text{D} n_\text{C} + \psi_\theta n_\text{A}\right)}. \tag{1}$$

The desired information on range $R = \psi_\text{R}\frac{c}{2}\frac{N}{B}$, range-rate $v_\text{R} = \psi_\text{D}\frac{c}{2T_\text{c}f_0}$, and DOA $\theta = \operatorname{asin}(2\psi_\theta)$ [30] corresponds to the normalized frequencies ψ_R, ψ_D, and ψ_θ, respectively. Here, c is the speed of light and $f_0 = 77\,\text{GHz}$ is the carrier frequency. In a measurement the acquired IF signal

$$x_{\text{IF},k}[n, n_\text{C}, n_\text{A}] = \sum_{m=0}^{M} A_m \mathrm{e}^{\mathrm{j}2\pi\left(\phi_m + \psi_{\text{R},m} n + \psi_{\text{D},m} n_\text{C} + \psi_{\theta,m} n_\text{A}\right)} + w[n, n_\text{C}, n_\text{A}] \tag{2}$$

contains a superposition of M point target reflections, i.e. the model from (1) with different and unknown parameters A_m, $\psi_{\text{R},m}$, $\psi_{\text{D},m}$, $\psi_{\theta,m}$, and ϕ_m. Additionally, additive white Gaussian measurement noise $w[n, n_\text{C}, n_\text{A}]$ is present. For the theoretical case of a single point target, i.e. $M = 1$, correlating the measured $x_{\text{IF},k}[n, n_\text{C}, n_\text{A}]$ with the model $s_{\text{IF},k}[n, n_\text{C}, n_\text{A}]$ and a subsequent maximum search could be employed to estimate the unknown parameters of the target. Due to the structure of the model (1) the correlation can be efficiently carried out

using the FFT. For a real world scenario with $M > 1$, the windowed three-dimensional FFT

$$\hat{P}_k(\psi_\mathrm{R}, \psi_\mathrm{D}, \psi_\theta) = \sum_{n}\sum_{n_\mathrm{C}}\sum_{n_\mathrm{A}} a_\mathrm{R}[n]a_\mathrm{D}[n_\mathrm{C}]a_\theta[n_\mathrm{A}]x_{\mathrm{IF},k}[n, n_\mathrm{C}, n_\mathrm{A}] \times \mathrm{e}^{-\mathrm{j}2\pi\psi_\mathrm{R}n}\mathrm{e}^{-\mathrm{j}2\pi\psi_\mathrm{D}n_\mathrm{C}}\mathrm{e}^{-\mathrm{j}2\pi\psi_\theta n_\mathrm{A}} \quad (3)$$

is commonly used as an estimator for the spectrum. Parameters of the peaks in the spectrum are used to deduce the parameters of the targets.

For the range and Doppler dimension, we successfully used typical windows a_R, a_D, like the von Hann window, and the Nuttall-Tukey window with similar results. However, due to some strong reflections near DC, we preferred the Nuttall-Tukey window over the von Hann window to reduce the blind range close to zero range. In the angular dimension we used a Taylor window a_θ with 2 taps and -20 dB sidelobe damping. Additionally, we extended the angular dimension from 8 to $Z_\mathrm{A} = 64$ samples by padding with zeros prior to the calculation of the FFT. The zero-padding technique is a computationally efficient interpolation of the spectrum estimate (3) to reduce the effects of discretization.

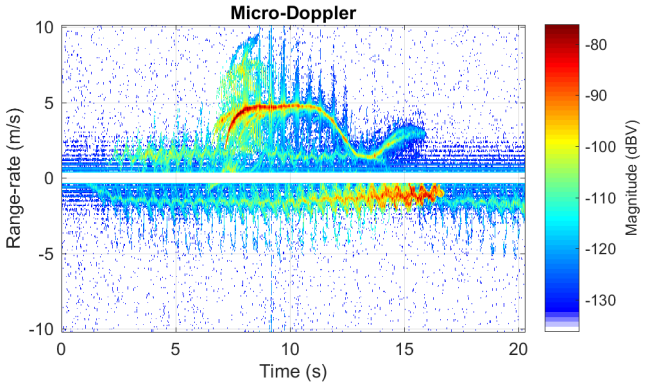

FIGURE 4. Visualization of the Doppler spectrum over time. There are three individual signatures that need to be separated. At negative range-rate the pedestrian moving towards the radar sensor is visible and with positive range-rates another pedestrian and the car are identifiable but overlain.

We computed the spectrum from the measurements of the scenario depicted in Fig. 3 according to (3). The resulting range-rate over time is depicted in Fig. 4. Range and DOA dimensions have been eliminated by a maximum search. The individual micro-Doppler signatures that need to be separated are perceptible. A different view on the data, which is more suitable for target separation, is the range/Doppler/DOA map as given in Fig. 5 for $t \approx 12.6$ s. Next to static reflections located around zero range-rate the two persons at a distance of approximately 6 m and 20 m and at a car at 25 m are very well visible.

Often one of the many constant false alarm rate (CFAR) algorithms [31], [32] is used for the task of applying a threshold in radar applications. However, this family of algorithms is not applied here for three reasons. Firstly, CFAR thresholds

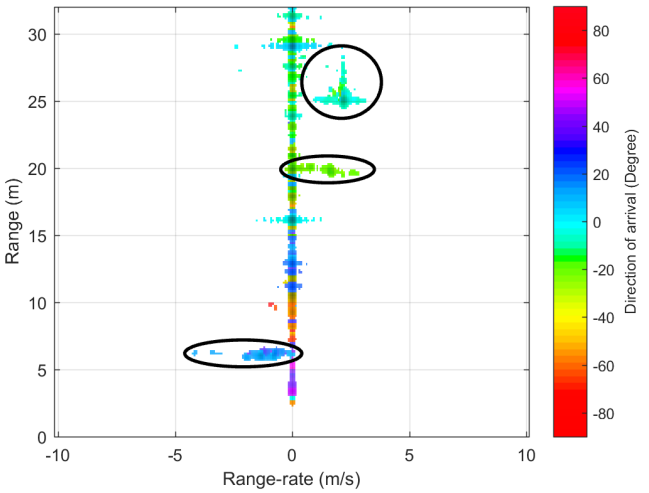

FIGURE 5. Plot of the three dimensional spectrum for the measurement scenario of Fig. 3. The intensity of each pixel represents the magnitude of the spectrum and the color of each pixel represents the DOA. Lots of reflections around zero range-rate originating from non-moving reflections are visible. The three moving targets are clearly visible and marked with ellipsoids.

are intended to follow the noise floor. In our measurements, however, the noise floor is flat. Secondly, CFAR algorithms should avoid detections in cluttered areas. Due to the high resolution in range and range-rate of our setup, the considered targets occupy more than a single cell. Thus, in range and range-rate our expected targets appear like the clutter that CFAR algorithms are typically designed to suppress. Thirdly, the DOA dimension has only 8 samples, which is not enough to apply a CFAR algorithm in a meaningful way. Therefore, a CFAR based threshold does not apply here.

In our test-scenarios, pedestrians were able to walk up closely, i.e. less than 3 m towards the radar frontend. As received power increases significantly with lower distance, the reflections of close pedestrians have much larger extent in the three-dimensional spectrum than their distant counterparts. To avoid for different detection and clustering strategies depending on distance, we used a threshold level with an $1/R^2$-shape with 16 dB noise distance at R_max.

A different thresholding technique for DOA is required, due to the large difference of the sidelobe levels and number of samples compared to range and range-rate. In the 3D datacube we select each range/range-rate index that has any of its corresponding DOA bins above threshold. Then we select every bin along the DOA dimension which is above the sidelobe level (SLL) (here the -20 dB of the Taylor window) and the threshold. This method allows us to fully utilize the processing gain of the FFT in each of the dimensions.

Fig. 6 shows the detections, i.e. elements above the threshold in the spectrum depicted in Fig. 5. These correspond to each of the pixels in Fig. 5 which are not white. A high number of reflections above threshold are present, but we are mainly interested in the ones originating from the three moving targets. Feeding a tracker with all of the detections would be computationally very demanding and most likely produce unusable results.

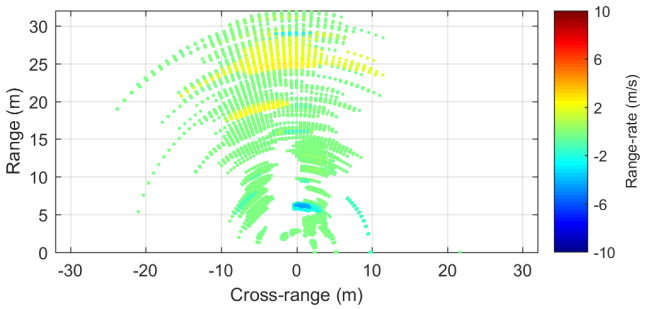

FIGURE 6. Plot of all detections, i.e. elements after applying the threshold on the spectrum of Fig. 5, from the measurement scenario of Fig. 3. The detections are converted to x-y space for better perception. Feeding a tracker with all of the detections would be computationally intractable.

The task to combine the individual detections to represent physical targets is handled by the clustering stage as will be presented in the following section.

IV. CLUSTERING

The final goal in this work is to produce tracks and properties of the targets associated with the tracks. Literature about tracking systems for radar or sonar applications mostly deals with the assumption that a single target will produce a single detection. This assumption is not valid for micro-range and micro-Doppler systems, where a target consists of several detections. To combine these two worlds we need to assess which detections belong to which target. This step is called *clustering*. Clustering allows to combine detections to represent a physical target. The cluster center can then be used as input for a tracker. There are several ways to define a center of a cluster, e.g., the location of the strongest reflection, the arithmetic mean over a cluster's elements, a weighted arithmetic mean, etc. Once the cluster center is calculated, it can be used as input for a Kalman-filter based tracker.

The field of clustering mainly originates from data mining in big databases. There are numerous classes of clustering algorithms broadly available. Unfortunately, due to their origin, many algorithms have severe disadvantages which make them suboptimal for radar applications. Firstly, the number of clusters, i.e. the number of targets in a radar measurement, must be known a priory. This information is not available for radar measurements. Secondly, many clustering algorithms have shape constraints and tend to produce spherical clusters only. Thirdly, optimization of available clustering algorithms often go towards large databases with more than 100 dimensions. Additionally, our data has a property that is not covered in the literature on clustering. As Fig. 5 depicts, the pedestrian at a range of approximately 20 m has gaps along the range-rate dimension. These gaps are particular features of a pedestrian's reflection. In our measurements, we observed gaps up to $v_{\text{gap}} = 1\,\text{m/s}$.

What is needed in our radar tracking application is a clustering algorithm that is robust against outliers and estimation errors, can cope with gaps in range-rate, and is fast and suitable for online calculation. We found two applicable clustering algorithms, the family of density-based spatial clustering of applications with noise (DBSCAN) and the inner loop of the family of grid-based clustering algorithms, which is basically a flood-fill.

A. DBSCAN FAMILY

The DBSCAN family of clustering algorithms aims to find structures of high densities and arbitrary shapes in sparse data-sets. A detailed description is presented in [33] and [34]. Considering the detections from Fig. 6, this fits our needs.

DBSCAN is quite fast since it only needs to address each element once. The original DBSCAN is robust against outliers but heavily influenced by different densities over the search area. A measurement scenario with lot of static clutter and few moving targets, as it is the case in our considered scenarios, show high density differences. The density of detections is high around zero range-rate and low where single targets are located. In this case DBSCAN most likely combines all elements near the static-clutter-ridge at $v = 0\,\text{m/s}$ wrongfully to a single cluster.

Purging all static detections, as it will be described in Section IV-D can alleviate the problem but results in loss of information. We tackled above mentioned problem of DBSCAN by introducing the extended density-based spatial clustering of applications with noise (EDBSCAN) [35] algorithm. EDBSCAN introduces a size attribute for each detection. The original EDBSCAN approach works well in a sparse environment and low dynamic range. In a dense environment with strong reflections off static targets and weak reflections off pedestrians, this approach becomes unstable. In such a case, static targets with high power are merged, while extended targets with lower power—the pedestrians we are interested in—will be split into several clusters.

Here we introduce a new size metric for EDBSCAN. We discriminate between static and moving detections. Static detections have a range-rate less than $v_{\text{min}} = (v_{\text{mp}} + v_{\text{gap}})/2 \approx 0.8\,\text{m/s}$, where $v_{\text{mp}} \approx 0.6\,\text{m/s}$ is the width of the main peak of the FFT along the range-rate dimension using the Nuttall-Tukey window. For static detections the size of the ellipsoid equals the width of a bin of the underlying FFT. The size of the ellipsoid for moving detections is twice the size of an FFT-bin in range and v_{gap} in range-rate. With this setting, just a small gap in range-rate is sufficient to separate moving from static targets while moving targets can have larger gaps up to v_{gap}. Using the new size metric, EDBSCAN produces a good representation of the static environment, while also combining the reflections of moving pedestrians reliably.

B. GRID-BASED CLUSTERING

Like every other DBSCAN variant, the EDBSCAN based clustering approach suffers from the computational demand of calculating distance metrics between all elements. The calculation of distance metrics can be avoided by exploiting the grid-structure of the FFT. After applying the threshold, a flood-fill algorithm can be used to determine clusters [36].

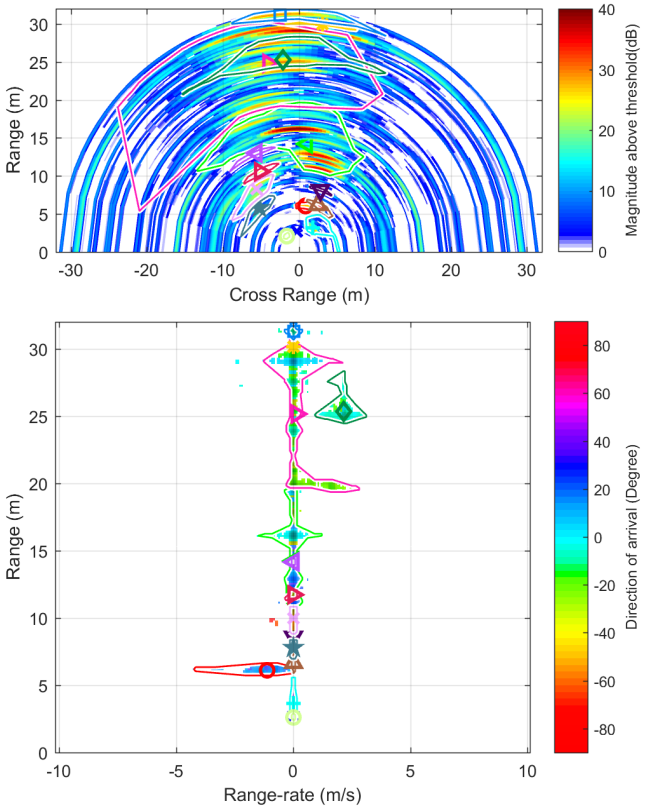

FIGURE 7. Clustering results of the detections showed in Fig. 6 with EDBSCAN applied. The background shows the bins of 3D-FFT which are above the threshold level. Each individual cluster is indicated by a separate marker at the location of its center and an outline with corresponding color. Most of the individual detections belonging to a physical target are correctly represented as individual clusters. At a distance of approximately 20m the problem of a moving pedestrian being merged with non-moving targets is visible.

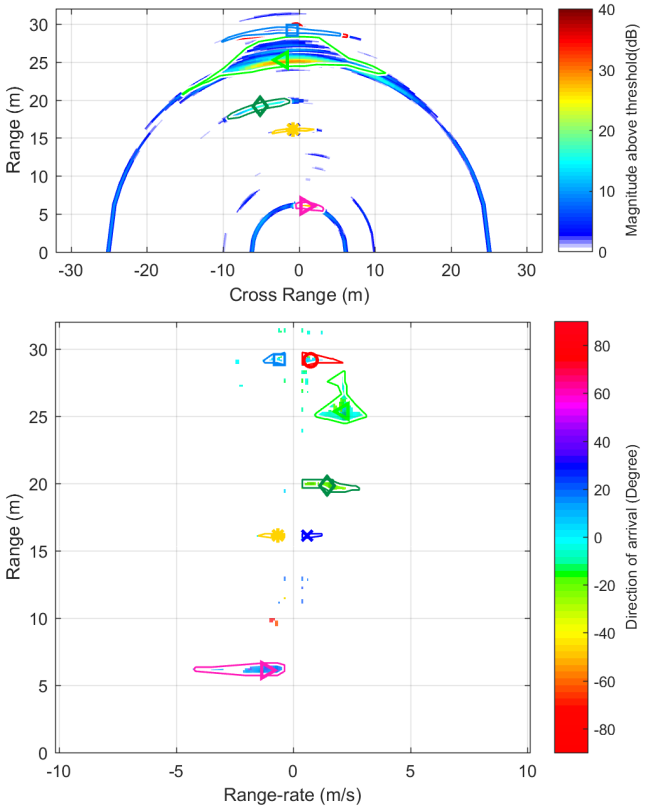

FIGURE 8. Clustering after purging of static targets. Compared to Fig. 7, all the moving targets are correctly clustered.

With a step size greater than one, the flood-fill algorithm can account for gaps in the clusters. We set the step size to 13 in the range-rate dimension which accounts for a gap of $v_{\text{gap}} = 1\,\text{m/s}$. Setting a minimum cluster size makes this algorithm robust against outliers.

Compared to EDBSCAN the gap size parameter is global for grid-based clustering. A gap size individual for each element would correspond to an asymmetric distance metric and unintended behavior, i.e. the order of execution would determine whether detections are combined to a cluster or not.

C. CLUSTERING RESULT COMPARISON

Grid-based clustering is able to produce similar results as density based clustering. In theory, grid-based clustering has a higher chance (compared to EDBSCAN) to produce bogus clusters by merging static with moving elements. This is due to the global gap size parameter. See, Section IV-D for further explanation of this phenomenon. However, in our measurements we did not observe a significant influence of this difference on the tracking stage.

On the other hand, grid-based clustering is much faster if the number of detections exceeded a few hundred, which was nearly always the case in our measurements. If computational demand can be tolerated, super-resolution algorithms might be used for, e.g., DOA estimation. In such cases EDBSCAN is a good choice since it does not rely on gridded data.

Fig. 7 shows EDBSCAN clustering applied to the detections of Fig. 6. The center of a cluster is calculated as the weighted mean of its elements, where the weight is an element's magnitude. To reduce the number of false-alarms we only use clusters with 50 or more elements for further processing.

D. CLUSTERING WITH AND WITHOUT STATIC ELEMENTS

For scenarios where pedestrians can get in close proximity (in terms of resolution) with static targets, reflections of the moving and the static targets might get merged into a single cluster. This is the case in the underground parking lot, were there are a lot of significant reflections from pillars and parked cars. Fig. 7 depicts a common clustering problem that can occur in such a scenario. The pedestrian at a distance of approximately 20 m was nearby an extended static target, i.e. the combined reflections of pillars and cars. The detections of the pedestrian and the static targets are combined and form a large cluster with an L-shape in range/Doppler dimension. The cluster center of such an L-shaped cluster does not represent anything physically present and hence will be a problematic input for the tracking algorithm.

To avoid such L-shaped clusters in dense scenarios, prior clustering it can be advantageous to remove all detections below a minimal range-rate of $v_{\text{min}} = v_{\text{mp}}/2 \approx 0.3\,\text{m/s}$, which is half the width of the main peak of the used Nuttall-Tukey window in Doppler dimension. This measure purges all non-moving point-targets. Fig. 8 shows the same setup as Fig. 7 but after pruning static detections. EDBSCAN produces a smaller number of clusters and deformed clusters do not occur any more. However, modulated static targets, e.g., due to reflections on moving objects, may leave some remains and produce some spurious clusters.

Without the static reflections, subsequent analysis has no information about low-velocity components. This has implications on performance of the tracker and the quality of the resulting tracks. Too much clutter, i.e. static targets, will result in either bogus tracks or no tracks. However, like static targets, moving targets with a nearly constant distance to the radar frontend will also be purged. Thus, tracks might be stopped prematurely. This could be alleviated by complex gating strategies like the ones already available for air-craft tracking [37]. Another negative side-effect of purging detections with low Doppler frequency is that subsequent micro-Doppler analysis will also lack the low-velocity component.

E. TIME DEPENDENCY

Doing such clustering at consecutive timesteps introduces another problem. Imagine the case where the i-th target in a field-of-view is represented by the j-th cluster at timestep k. In timestep $k+1$, the ordering of the clusters might be different. Thus, a target can not be followed reliably over time using clustering solely. We will treat this problem in the next section.

V. TRACKING

The Kalman filter [38] is a fast recursive linear filter which can be used to estimate the state of a dynamic system from a time series of noisy measurements. Kalman filters can estimate the internal state for model variables that cannot be measured directly, e.g., cross-range velocity. Additionally, they have a recursive structure, low computational demand, and are robust against measurement inaccuracies and correlations in a multi-target scenario. Therefore, the Kalman filter based tracking is one of the most widely used tracking algorithm in the radar community [37].

A. TRACKING A SINGLE TARGET USING PSEUDO MEASUREMENTS

The Kalman filter framework requires the definition of a state vector $\mathbf{x}_k$, a state noise process, the variance of measurements and the transition from measurement space to state space.

Various possibilities are available to select the entries of the state vector $\mathbf{x}_k$. Viable state vector entries are, e.g., position, velocity, acceleration, heading, and turn-rate of a target. Given a certain measurement variance, increasing the number of elements in the state vector will also increase the corresponding state variances, and of course computation time. Consequently, one should choose the number of state variables as low as necessary to achieve a low variance of the estimated parameters. Heading and turn-rate are mainly used for slow-turning vehicles like cars, ships, or airplanes. Pedestrians or animals can have highly fluctuating heading. Hence, heading and turn rate are not employed in the tracker used in this paper. Like in many other applications, we assume a constant-velocity model in the x-y plane. With these considerations the state vector

$$\mathbf{x}_k = \begin{bmatrix} x & v_{\text{x}} & y & v_{\text{y}} \end{bmatrix}^{\text{T}} \tag{4}$$

contains the position and the velocity for each observable track j in two dimensions.

The state noise process contains the contribution of the unknown accelerations of a target. We model these as a zero mean Gaussian noise process with a covariance matrix

$$\tilde{\mathbf{Q}} = \begin{bmatrix} \sigma_{a,\text{x}}^2 & 0 \\ 0 & \sigma_{a,\text{y}}^2 \end{bmatrix} \tag{5}$$

holding the acceleration's variances $\sigma_{a,\text{x}}^2$ and $\sigma_{a,\text{y}}^2$. Using an inertial measurement unit (IMU) we empirically assessed that $\sigma_{a,\text{x}} = \sigma_{a,\text{y}} = 8\,\text{m/s}^2$ is a reasonable value for pedestrians. However, we observed that this value also works well for tracking the cat and the UAV.

The measurement space contains the measurements corresponding to a certain track. For the radar sensor from our setup, these are range, range-rate and DOA for each target. However, including the range-rate in the Kalman filter framework would introduce two problems. Firstly, the Kalman filter's update stage would become highly non-linear and even linearization, e.g., with the extended Kalman filter (EKF) framework [39], [40], is not necessarily stable anymore [41], [42]. Thus, higher order filters, like the unscented Kalman filter (UKF), would be necessary. However, these non-linear filters have significantly higher computational demand. Secondly, we know that targets like pedestrians are widely spread in the Doppler dimension. Of course, the previous clustering stage, in theory, should combine all velocity components of a target and calculate the true range-rate. However, due to the heavily fluctuating radar cross section (RCS) [43], [44] of pedestrians and large difference between, e.g., hands and torso velocity, the range-rate estimate might be too inaccurate.

Our approach for a tracker does not include the range-rate as possible input for the tracker. However, note that the information is not lost, since it was already used in the clustering stage. Despite the simplification to avoid the highly non-linear transform of the range-rate to v_x and v_y, by not including the range-rate into the state vector the remaining transform of a measurement

$$\tilde{\mathbf{z}}_k = \begin{bmatrix} r_k & \theta_k \end{bmatrix}^{\text{T}} \tag{6}$$

to state variables (4) is still non-linear. The idea to overcome this problem, is to use functions of measurements, so-called

pseudo-measurements, as input for the Kalman filter. Here we use x-y coordinates

$$\mathbf{z}_k = f(\tilde{\mathbf{z}}_k) = [r_k \cos(\theta_k) \quad r_k \sin(\theta_k)]^{\mathrm{T}} \tag{7}$$

as functions of the measured range r_k and DOA θ_k.

The measurement variance is also needed. In literature, the approach of a linearized transform of the measurement variance is often employed, e.g., in [45]. However, two facts make this standard approach unfavorable. Firstly, the underlying assumption for this approach is that the measurement variances in range and DOA are known. However, methods to calculate the measurement variances are only available for targets that follow a very simple model, e.g., the Cramér-Rao lower bound for point-targets. For extended targets, as we have to deal with here due to our high resolution radar system, there is no reliable way to determine the measurement variances. Secondly, our primary goal is to track persons in the x-y plane. The appearance of persons to a radar sensor cannot be modeled in a simple way. So, we empirically assessed an approximate variance

$$\mathbf{R} = \begin{bmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{bmatrix} \quad \text{with} \quad \sigma_x = \sigma_y = 0.5\,\text{m}. \tag{8}$$

With all parameters defined for the linear Kalman filter framework, we can use the well-known prediction and update equations from [38] to solve for the state vector $\mathbf{x}_k$ for all timesteps k for a single track. In the next section we will show how to deal with multiple tracks in a scenario.

B. MULTIPLE TARGETS: THE ASSIGNMENT AND TRACK-MANAGEMENT PROBLEM

As we most likely observe multiple targets simultaneously, there is the *assignment problem* that needs to be solved. We need a set of rules for a) when a track starts, b) when a track ends, and c) which measurement should be used to update which track.

At the first timestep $k = 0$, each of the targets, indexed with i, resulting from the clustering stage does start a new track j. This means, a new instance of the state vector $\mathbf{x}_{j,k}$ is created. At the following timesteps the clustering stage generates new lists of targets $\mathbf{z}_{k,i}$. These new targets can either be used to update established tracks or to start a new track. The problem to determine which target i is used to update a track j is called an assignment problem. Solving the assignment problem is crucial for tracking performance, i.e. to get tracks that are robust and sane. In this context, *robust* means that tracks are maintained as long as the target is visible and *sane* means that the computed tracks match the scenario.

The most popular assignment algorithms are, the nearest neighbor (NN) approach, the global nearest neighbor (GNN) approach, the probabilistic data association filter (PDAF) approach, joint probabilistic data association filter (JPDAF) approach, and the nearest neighbor joint probabilistic data association filter (NNJPDAF) approach. In the following, we will briefly recapitulate these possible solutions for the assignment problem, see [37] for a more detailed overview.

1) THE NN APPROACH

This is the simplest and computationally least demanding algorithm. As initial step, a matrix with χ^2-track-scores [37] for all possible assignments is generated. Then a search is carried out for the best measurement-track-pair. The assigned measurement and track are both removed from the matrix. The last two steps are repeated until all assignments are made.

2) THE GNN APPROACH

Like the NN approach, the GNN approach works on the score matrix. The GNN approach, however, all possible measurement-track-pairs by exploring the Hungarian- or Munkres-algorithm. This decreases assignment errors but increases computational demand.

3) THE PDAF APPROACH

For each measurement-track-pair the probability that the measurement represents the correct target is calculated. Based on these probabilities an artificial measurement $z_{k,i}$ is calculated as linear combination of all possible assignments.

4) THE JPDAF APPROACH

The JPDAF is based on the same idea as PDAF but includes the assumption that nearby tracks can interfere. Thus, the association probabilities are computed jointly across all targets and measurements. Since this measure increases the computational demand significantly, approximations have been developed. In this paper we use the Cheap JPDAF [46].

5) THE NNJPDAF APPROACH

The NNJPDAF works comparable to the NN approach but used the assignment probabilities of JPDAF instead of the distances. Any JPDAF can be used to calculate the probabilities. We used the Cheap-JPDAF again.

Note that for all considered methods, the assignments are calculated independently for each timestep k. The multi-hypothesis tracking (MHT) approach [47], [48] forms a tree of hypothesis of assignments. For unclear situations a definite assignment is delayed until new information from subsequent timesteps is available. However, at each timestep k the tree grows by another level and is becoming broader with a high number of clutter measurements. Maintaining the tree introduces a very high computational demand compared to the above mentioned methods, even if unlikely leaves are cut. Hence, we did not consider the MHT further.

To reduce the complexity of the assignment algorithm and avoid false assignments, a restriction towards highly likely target to track association pairs is performed as a first step. This step is also called *gating*. In literature [37], [38], [49], gating is also referred to as the definition of a g-sigma ellipsoid. Typically g is chosen in the range of one to three. We used $g = 3$ throughout this paper. Only detections that are inside the gate, are considered as possible candidates in the assignment stage of the considered track.

The algorithms mentioned in this section have the underlying assumption that only a single measurement inside the

acceptance gate of a track is the correct one. The others are considered as clutter. This is yet another reason for the need of a reliable clustering algorithm.

We are interested in moving targets. Using all the static targets as input to the MTT framework will result in computational overhead and a high number of ghost tracks. The simplest approach for alleviation would be cutting away components with low Doppler by either setting them to zero or applying a high-pass filter. However, a high-pass filter would also eliminate low range-rate components for moving targets which results in unintentionally cropping of moving targets. As alternative, we use the clustering result to eliminate static clutter. All clusters, i.e. targets, where the cluster center is at a range-rate below $v_{\min} = 0.2\,\mathrm{m/s}$ are neglected, in addition to those which were already removed with the optional step described Section IV-D.

C. TRACK DELETION

A valid target might become unobservable by, e.g., occlusion or simply heading out of the field-of-view. Additionally, falsely started tracks, e.g., due to clutter, mirror reflections, wrongly generated clusters, etc., must be eliminated.

Using a NN approach, falsely started tracks will not be updated. In this case the track purging state is simple. At each timestep, all tracks that got no update within the last 250 ms are considered as not running any more and are excluded from further updates. We assessed this specific value empirically. A lower value would lead to track loss when a target is shadowed, e.g., by a pedestrian. On the other hand, a higher value would keep ghost tracks or legitimately stopped tracks active unreasonably long.

Unfortunately, the PDAF approach leaves the assignment and track maintenance problem completely unconsidered. Unlike the NN approach, using the PDAF approach, there are now several possible continuations for a falsely started track.

Firstly, there is a chance that it might be not updated any more. Then the same procedure as for the NN approaches can be used. However, such a situation rarely occurs in case of a PDAF approach. It is much more likely, a track gets an update with a low weight. In this case, the entries of the state covariance matrix $\mathbf{P}$ rise. A track is then purged if the sum of variances of observable states, i.e. the trace of the transformed state covariance matrix, exceeds the limit

$$\sqrt{\mathbf{P}_{j,1,1} + \mathbf{P}_{j,3,3}} > 0.6\,\mathrm{m} \tag{9}$$

Without this measure, stale tracks are kept too long and there is a high risk of wrong track continuity. E.g., a pedestrian stops and a car passes by. Then the track of a pedestrian could be falsely continued with the measurements of the car.

Secondly, falsely started tracks will converge towards existing tracks within a few time steps. This phenomenon is called coalescence. To tackle it, we introduce a track elimination stage which checks for converging tracks by comparing the most recent state vectors. If

$$\left|\left|\left(\mathbf{x}_{k,m} - \mathbf{x}_{k,n}\right) \oslash \left[\Delta R\ \Delta v\ \Delta R\ \Delta v\right]^{\mathrm{T}}\right|\right| < 15, \tag{10}$$

where $\oslash$ is the element-wise division for normalization to the bin-widths ΔR and Δv, is fulfilled for two tracks m and n, they are considered as merging. In such a case, the one with the lowest number of updates is eliminated.

We assessed the comparison values in (9) and (10) empirically during extensive measurement campaigns.

D. COMPARISON OF ASSIGNMENT ALGORITHMS

We tested the different assignment approaches on different scenarios. See the attached videos for a visual comparison.

The simple NN- and GNN-assignments work well for separated point targets, e.g., measurements taken from small objects with low fluctuations like small UAVs. However, pedestrians have high fluctuations of their RCS, are extended targets, and additionally have a high possibility of multiple reflection centers. For these, the NN approach does lead to a more frequent track initiation and loss compared to the PDAF approaches, especially when targets are maneuvering, i.e. taking a turn. Increasing the elements of $\tilde{\mathbf{Q}}$ to allow for higher fluctuation of the target produces more stable tracks. Increasing the acceptance gate g does not work well. Although increasing g does increase the chance that a correct assignment is made, it also increases the chance of wrong assignments and ghost tracks, i.e. tracks following clutter detections, or wrong assignments. The PDAF approaches are less sensitive to this phenomena, however they are more susceptible to form ghost tracks from erroneous or clutter detections. Additionally, allowing more than one assignment in each update stage of the Kalman filter will further combine individual reflections of pedestrians if range-rate is not incorporated in the tracker.

VI. EXTRACTIONS OF FEATURES OF INDIVIDUAL TARGETS AFTER TRACKING

The process to generate input for a classifier is called *feature extraction*. Our work provide the prerequisite for target classification and feature extraction. As the vast majority of publications on classification rely on the fact that only a single target was observed, it is crucial to determine the features of each individual target, even for the typical case that multiple targets were in the field-of-view.

Simple features per target can be derived from the state $\mathbf{x}_{k,j}$ of each track j over time. These features include the traveled distance, maximum velocity or the histogram of the Kalman filter gain [1]. However, they are not sufficient for reliable classification and are thus not considered in literature about classification.

At each update of a track at a timestep k we do not only store the state $\mathbf{x}_{k,j}$ for the j-th track but also all detections that contributed to an update. Thus a track contains information on all contributing clusters along with their elements evolving over time. Having this additional information allows us to thoroughly analyze the features commonly used in classification of all targets individually; even for the common case that multiple targets were in the field-of-view. These features include the extent of a target as used in [1], [12],

and [50] as the size of the clusters, and the widely used micro-Doppler signature. The micro-Doppler signature $\breve{v}_j[k]$, i.e. high resolution range-rate, over time is widely used as basis for feature extractions, e.g., in [8], [10], [11], [18]–[20], [22]. With our framework we can generate

$$\breve{v}_j[k] = \left(\hat{\psi}_{\mathrm{D},1}, \hat{\psi}_{\mathrm{D},2}, \hat{\psi}_{\mathrm{D},3}, \ldots, \hat{\psi}_{\mathrm{D,L}}\right) \tag{11}$$

as the n-tupel of the normalized range-rates obtained from the estimated spectrum (3) of all clusters that contributed to track j at timestep k. Similarly, we generate the n-tupel of magnitudes

$$\breve{a}_j[k] = \left(\hat{a}_1, \hat{a}_2, \hat{a}_3, \ldots, \hat{a}_L\right) \tag{12}$$

associated with the elements of (11). Using $\breve{v}_j[k]$ and $\breve{a}_j[k]$ we can generate virtually any statistics of the micro-Doppler spread over time that is found in literature and already used for the case of a single target in the field-of-view of the radar sensor. Furthermore, combined feature like used in [11], [12] can be extracted using the proposed framework.

To show the possibility of feature extraction after tracking, in this publication we selected the short-time FFT (STFFT) of the sample STD $\mathrm{std}\{\breve{v}_j[k]\}$ and of the weighted arithmetic sample average wmean $\left\{\breve{v}_j[k], \breve{a}_j[k]\right\}$ over the L elements of the tuples from (11) and (12). Both features exploit the fact that parts of a human body change their speed periodically while walking [14]. The swing of the limbs during a step shows micro-Doppler spreads from approximately 0 m/s up to twice the walking pace. The STD is a reliable measure for the width of the Doppler-spread. Pedestrians have a significant frequency component of about 2 Hz, which is not present for other radar targets, like bicyclists or UAVs. The torso has a much higher RCS than the limbs. Therefore, the probability of detection is higher. However, the torso's velocity variation is much lower than the variation from the limbs. So, in addition to the STD the wmean is used to estimate the average velocity of a target.

A track is not necessarily updated at all timesteps k. If there was no update at a certain timestep, we filled the missing data with zeros. The integration time of the STFFT was set to 2 s to capture the expected human step frequency of approximately 2 Hz well.

VII. RESULTS AND ANALYSIS OF THE SEPARATED TRACKS

We have conducted different kinds of experiments, of which we will present four in the following sections:

A) an outdoor measurement in a pedestrian area,
B) an outdoor measurement of a cat in a backyard,
C) an indoor measurement of a small UAV, and
D) a measurement in an underground parking lot.

The measured data from all these scenarios has been processed with the same settings for spectrum calculations, clustering, and tracking. This shows that our set of algorithms is robust against scenario changes. Videos of the experiments and the results can be found in the attached media files.

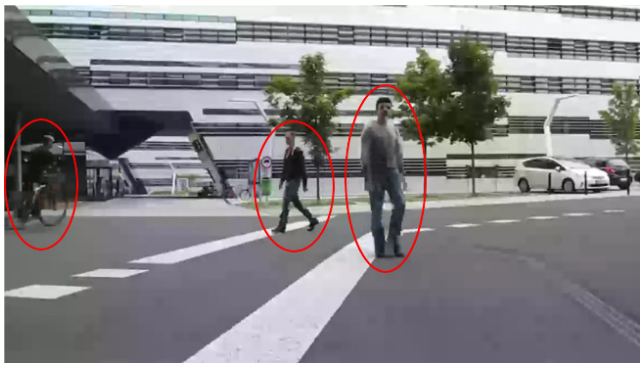

FIGURE 9. Camera view of the measurement scenario used for illustration. Three persons (two walking, one riding a bicycle) are moving around in a pedestrian area at the university.

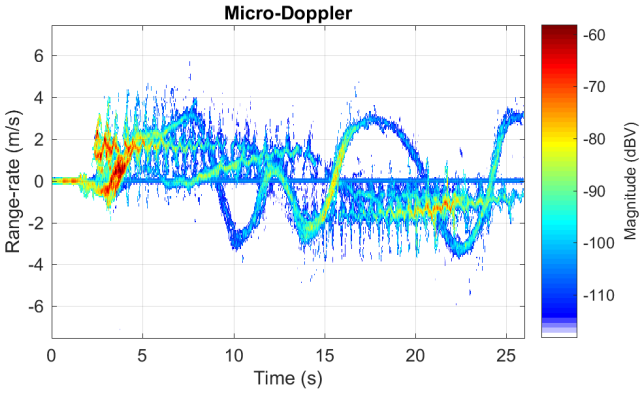

FIGURE 10. Doppler spectrum of the pedestrian area over time. The micro-Doppler signatures of the individual persons are entangled and overlapping.

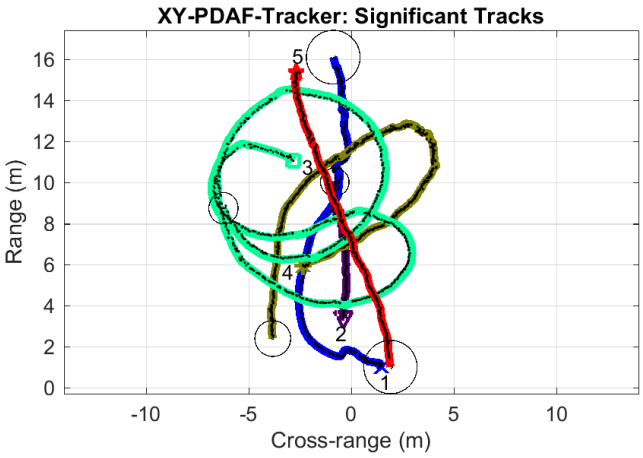

FIGURE 11. Detected tracks of the persons moving around in the pedestrian area. A marker with a number denotes the start of a track, a circle denotes the end of a track. The size of the circle corresponds to the acceptance gate. The black dots denote the measurements used to update the Kalman filters for the tracks.

A. OUTDOOR MEASUREMENT IN A PEDESTRIAN AREA

As first measurement scenario, we present an open space with low clutter in a pedestrian area at the university campus to clearly see the micro-Doppler effects of vulnerable road users. Two pedestrians and one bicyclist were moving around simultaneously in the radar sensor's field-of-view.

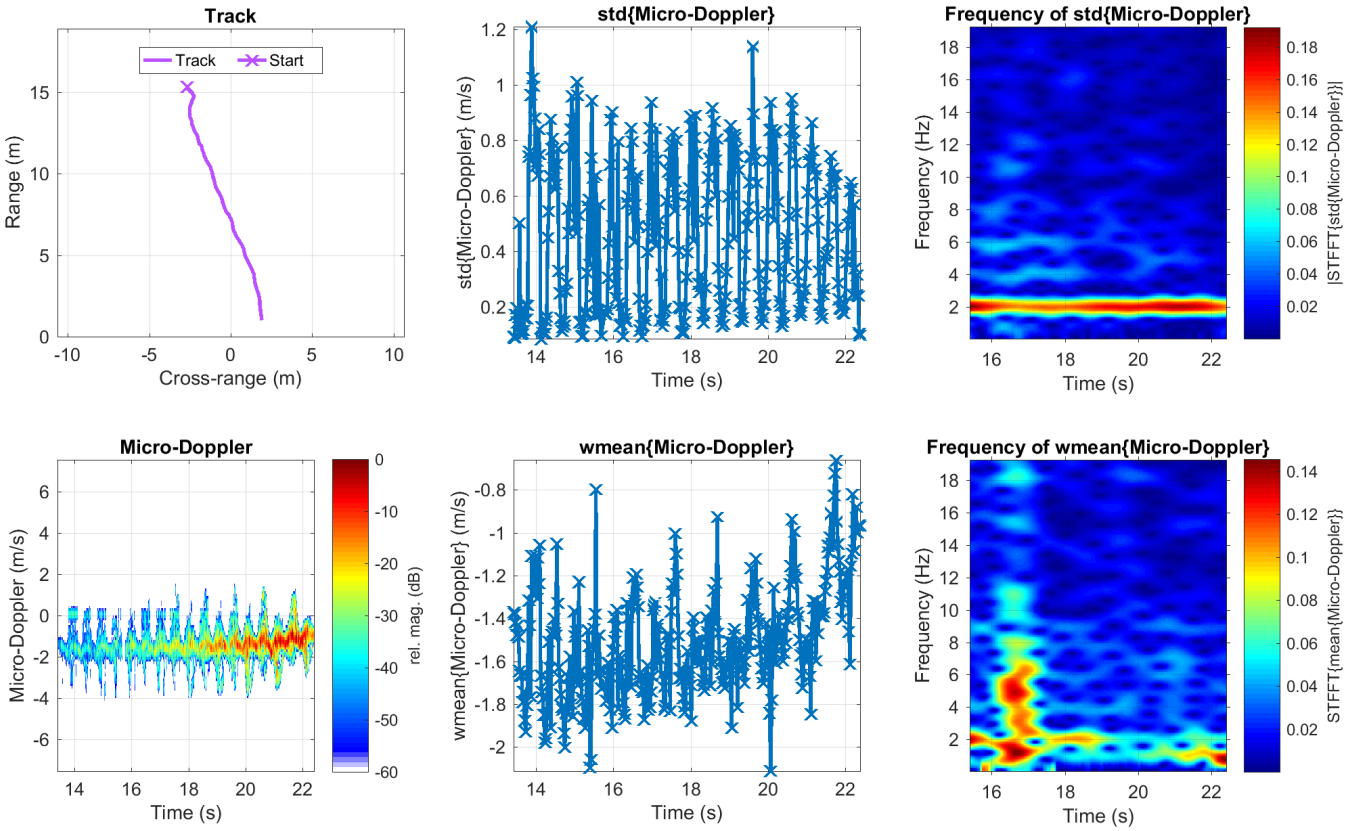

FIGURE 12. Pedestrian walking radially. The pulsating spread is clearly visible in the micro-Doppler plot, which originates from reflections off the swinging arms. In the STFFT plots the periodicity of 2 Hz, which is an indicator for the step frequency, is clearly visible, despite the clustering errors between second 14 and 18. Also in the mean of the micro-Doppler the step frequency can be observed.

Fig. 9 shows a photograph of the scenario. As the persons have been moving around concurrently, their micro-Doppler signatures are entangled and overlapping. This is depicted in Fig. 10.

The acquired data allows us to show that the algorithms are capable of separating multiple targets in a low clutter environment. A video attached to this paper shows that all assignment approaches work well. In Fig. 11 depicts the resulting tracks from the PDAF approach.

Even in cases where a person is not visible to the radar sensor for example when it is occluded by another person, the tracker is able to follow the occluded person. A problem only arises when targets get close and cannot be separated any more. In this case, one of two things will happen. Either only one track gets the update and is continued, or both tracks are continued. The latter case differs whether an NN approach or a PDAF approach is used. The reader is referred to the attached videos. In the NN approach, the update is arbitrarily appended to one of the tracks. This leads to holes in the information used for feature extractions. Using a PDAF approach, both tracks might get the update with the combined information of both targets. This leads to tracking errors like a wrong position estimate or to the case that tracks of two separate targets are stitched together wrongly. Such errors could be avoided by using a sensor with higher resolution.

The next paragraphs focus on the clearly visible micro-Doppler signatures of the individual persons. Fig. 12 shows the track of a pedestrian moving radially towards the radar. The swinging of the arms, which corresponds to the step frequency, is very well visible. As predicted by other authors [13]–[16], the periodicity in the STD of the micro-Doppler spread resembles the step frequency very well. Additionally, the periodicity of the mean, i.e. the motion of the torso, does show a certain periodicity. Between seconds 14 and 18 clustering errors occurred and static reflections became part of the cluster representing the pedestrian. They are visible as spurious velocity components around $v = 0\,\mathrm{m/s}$ in the reconstructed micro-Doppler spectrum. However, they have no significant influence on the standard deviation and mean.

A radial motion like in the previous example is ideal for micro-Doppler extraction. Thus, we take a closer look at a more difficult case, where the second pedestrian performed a direction change. For this example we ran the evaluation twice. Fig. 13 shows the result with and Fig. 14 without static detection removal according to Section IV-D. In both cases the tracker was able to follow the reflections of the pedestrian. The step frequency is very well visible in both cases in the micro-Doppler whilst moving radially and obliquely. Around second 14 the pedestrian was moving nearly tangentially with a very low range-rate. Including

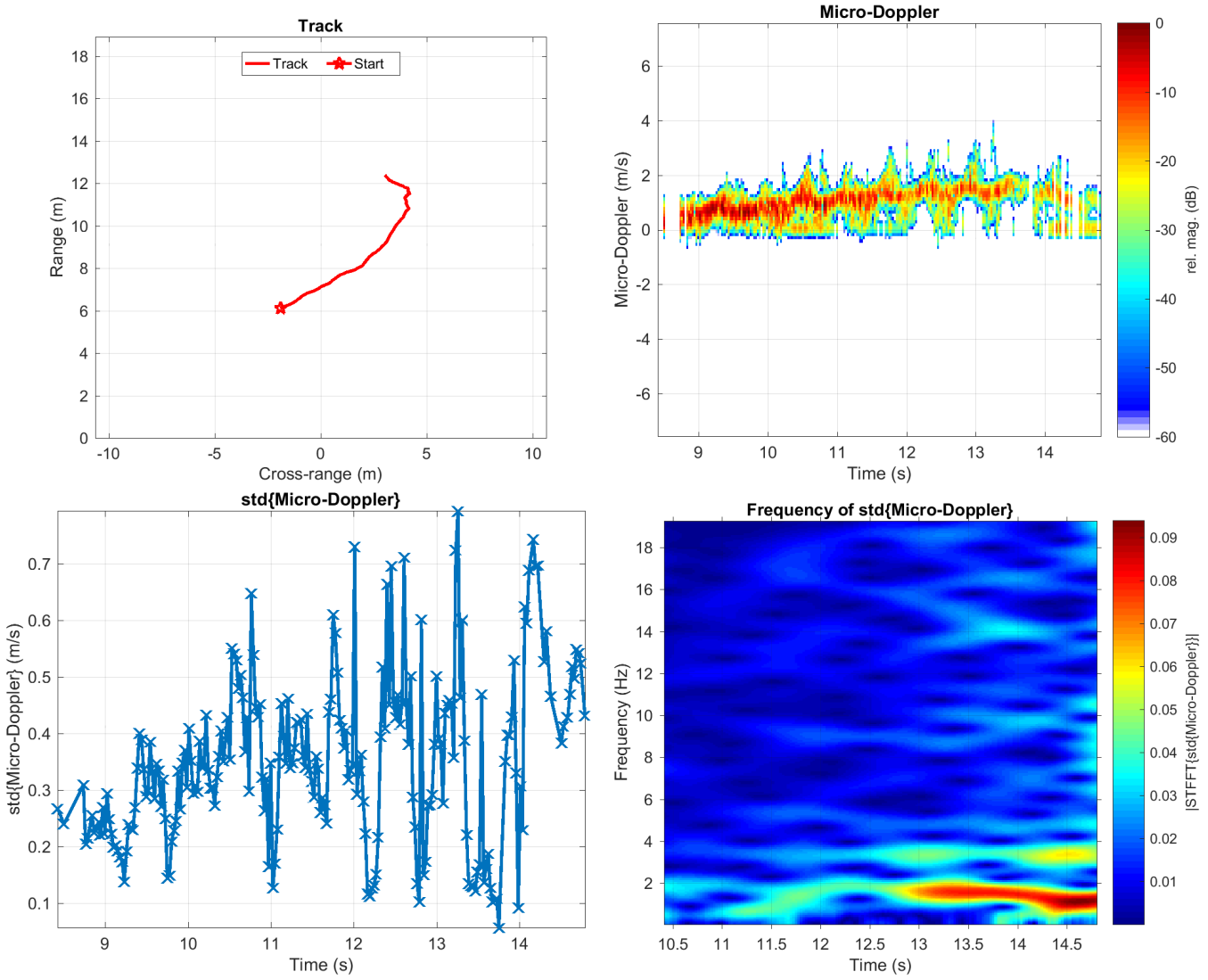

FIGURE 13. Reconstructed Micro-Doppler information of a pedestrian walking in an empty parking space. The swinging of the arms is clearly visible in the STFFT plots despite the non-radial movement. At approximately 14 s the pedestrian was moving nearly tangentially. As static detections have been included, this leads to a track-loss.

static detections leads to an undesired merge of the reflections of the pedestrian with the reflections of static targets and the track was subsequently lost. Purging all static detections prior clustering does prevent this effect as depicted in Fig. 14.

Fig. 15 shows the signature of the bicyclist, whose signature is quite different from the pedestrians. As expected, there is no 2-Hz periodicity, since a bicyclist obviously does not swing its arms. However, the reflections off the spokes are observable and flank the average velocity. This difference could be a suitable input for a classifier.

Note that the data for Fig. 10 to Fig. 15 was captured while all three persons were moving simultaneously in the field-of-view of the radar system. This is only possible due to the successful tracking.

B. OUTDOOR MEASUREMENT OF A CAT IN A BACKYARD

As second experiment, an outdoor scenario as depicted in Fig. 16 was selected. We conducted measurements in a backyard with a lawn. This is a tough scenario since grass is a highly reflective scatterer [51] and a cat has a small RCS. Even with significant clutter from the grass and the surroundings, the algorithms were able to track the cat. The result is shown in Fig. 17. It has to be noted that we were also able to track pedestrians and extract their micro-Doppler signatures in this scenario. Due to the similarity of the results compared to Section VII-A we omit the corresponding plots. However, we can conclude that a) a cat moving on the lawn can be tracked with our system and b) its signature is quite different to a human's signature. In the case of the moving cat, the step frequency can be more clearly seen in the mean of the micro-Doppler rather than in the STD as in the case

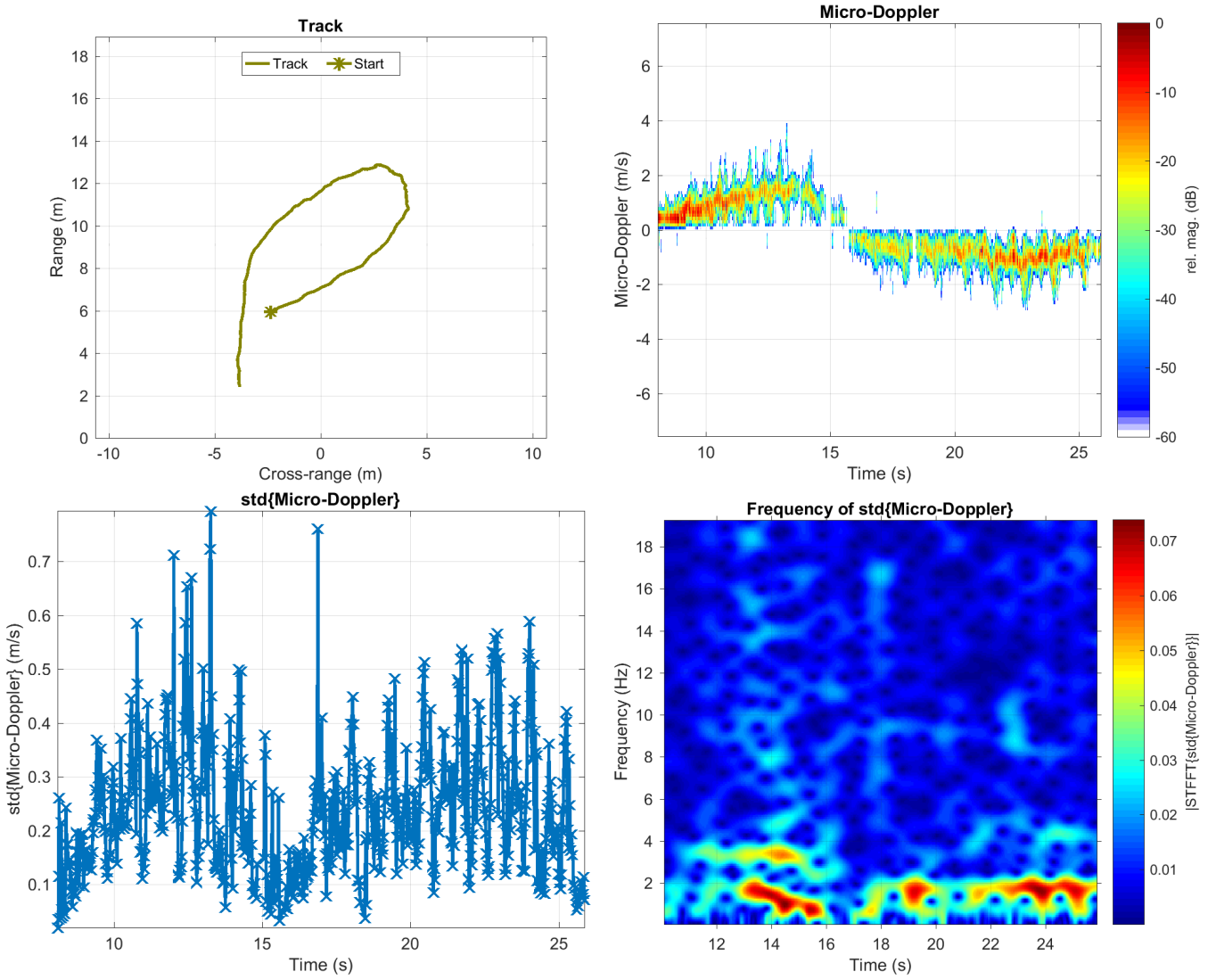

FIGURE 14. Same setting as in Fig. 13 but static detection have been purged prior clustering. The tracker can follow the turn around second 14, however spikes of the micro-Doppler profile going to zero range-rate are purged too.

for the pedestrian. The reason is the limited sensitivity of our radar sensor which was not capable to detect the short feet of the cat in the comparably tall grass. However, we are confident that the separation between humans and small animals, which is crucial for surveillance systems, is possible when the presented algorithms are accompanied with a suitable classifier.

C. INDOOR MEASUREMENT WITH A SMALL UAV

As third scenario, an indoor location was selected. We used the small UAV Blade 200 QX with a side-length of 142 mm and a height of 90 mm as a flying target. Fig. 18 shows a photograph of the scenario.

This scenario was selected since it has a high amount of static clutter reflections and a single, simple target with a very different micro-Doppler signature compared to the ones previously shown. Due to the sharp and unprotected rotor blades no pedestrian was allowed in the scenario for safety reasons. The high number of static reflections makes it crucial to purge static detections prior clustering according to Section IV-D. Fig. 19 shows a portion of a track of the UAV while it was flying in front of the radar sensor. Similar to [20], the basic characteristics of the UAV are perceptible. The UAV's only internal motion is due to the fast rotation of the rotor blades. There are no other moving parts. Most of the time they reflect the impinging waves into an arbitrary direction. However, at a certain angle of approach, the reflections coming back to the radar system are visible. Due to the high rotation speed of up to ≈20 000 rpm and the rotor diameter of 90 mm, the reflections of the rotor blades cover the whole unambiguous range-rate region.

In the bottom plot of Fig. 19 we ceased purging static reflections to visualize the difference. The cluster algorithm wrongfully generates L- or T-shaped clusters by

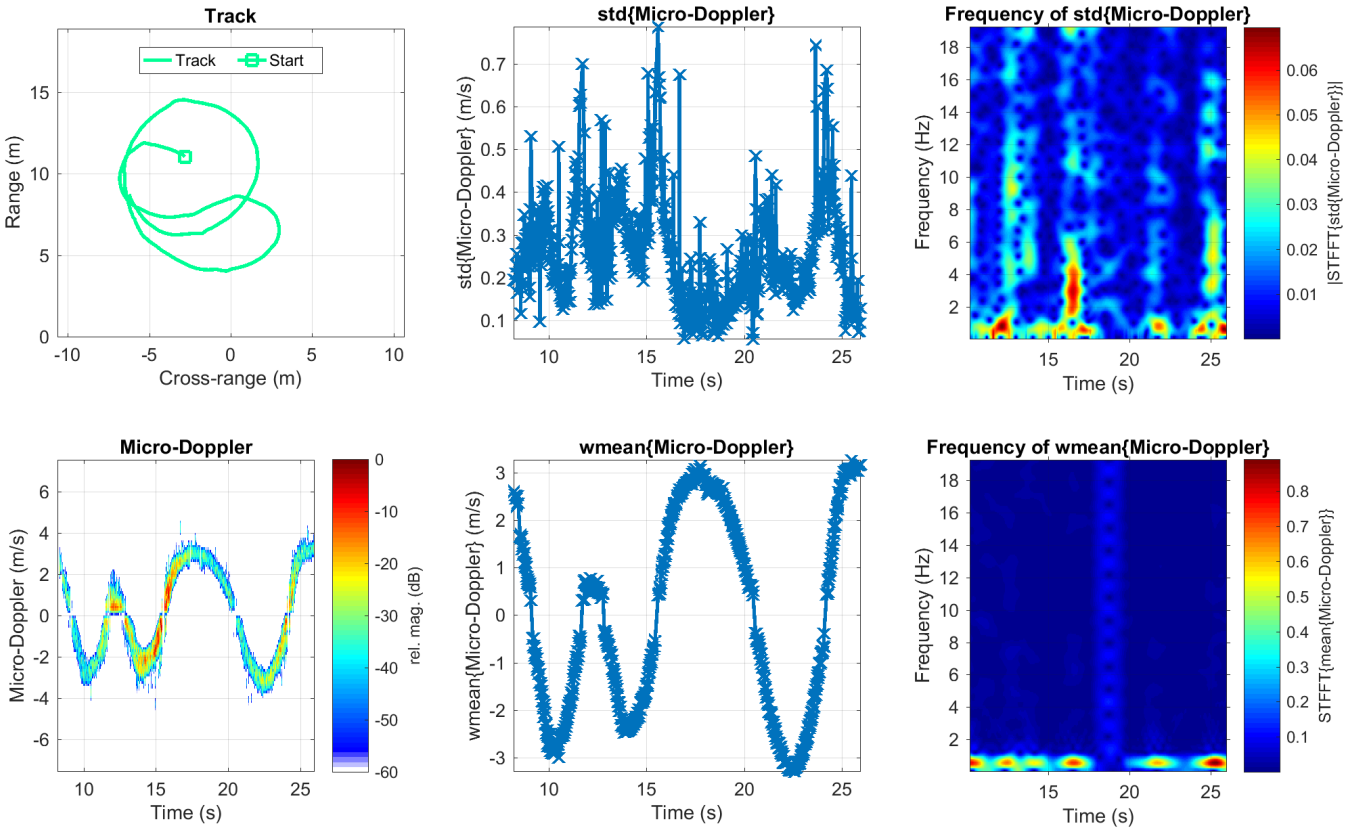

FIGURE 15. Bicyclist taking a turn. Unlike the pedestrian, the micro-Doppler of the bicyclist has no outstanding periodicity visible. However, a skirt of reflections around the average velocity might appear. These reflections originate from the spokes of the wheels and are especially visible around second 15.

FIGURE 16. Measurement scenario *cat in a backyard*. The cat is a small target within a high clutter, due to the grass, scenario.

merging the detections of the moving UAV with reflections of static targets. Quite a lot of missing updates are visible as gaps. They occur when the range-rate of the center of the L- or T-shaped cluster is too low to be used for a tracker update. Additionally, bogus clusters (where static detections have been included) produce the ridge around zero velocity.

The reflections off the rotor-blades and the motion of the UAV itself modulate the reflections of static targets appear like reflections of moving targets. This is caused by the forward scattering effect and will give rise to a number of ghost tracks. However, our framework is still capable of tracking the UAV. See the attached videos for more details.

D. UNDERGROUND PARKING LOT

The fourth scenario was the underground parking lot at the university from which the examples from the previous chapters were taken. The numerous static objects, like parking cars or supporting pillars, on both sides of the driveway make it mandatory to purge static detections prior clustering in order to get stable tracks. On the left side of the driveway a pedestrian was moving away from the radar sensor. On the right side of the driveway a pedestrian was moving towards the radar sensor. In the middle of the driveway a car was heading away from the radar sensor. Due to its reflective surface, the car gives raise to numerous multipath detections and ghost targets. We were able to track both pedestrians and the car and reconstruct their individual micro-Doppler signatures, despite this difficult scenario. Fig. 20 depicts the result using the NNJPDAF approach. See the attached video for details.

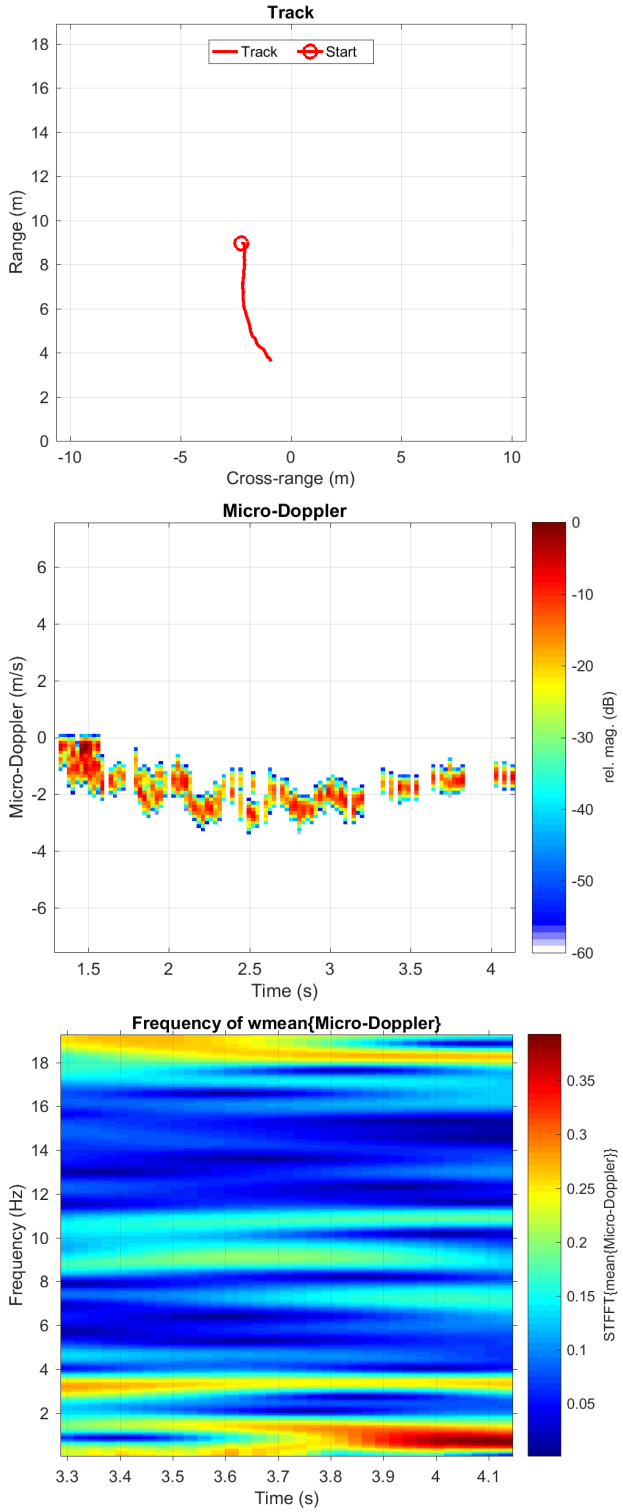

FIGURE 17. Cat moving on a lawn towards the radar system. Compared to the pedestrian, the step frequency in the time dependency of the mean of the velocity can be seen. The standard deviation did not produce usable results since the clutter of the grass was too high to the capture the feet of the cat.

In Fig. 21, the track and the reconstructed micro-Doppler signature of one of the pedestrians is depicted. Despite the cluttered environment with a high number of multipath reflections, the pedestrian could be tracked successfully. The micro-Doppler plot clearly shows the signature of a single pedestrian, despite the fact that a car and another pedestrian also have been moving in the scenario.

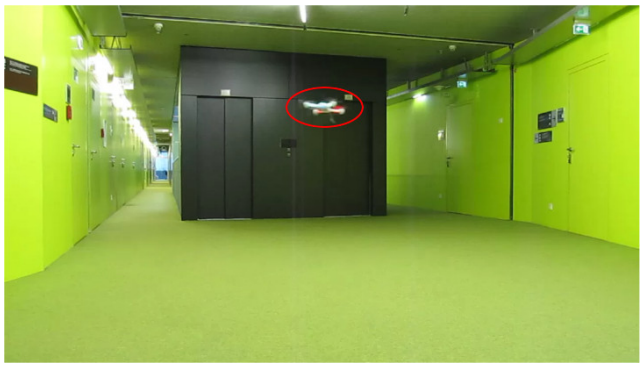

FIGURE 18. Measurement scenario *UAV flying indoors*. The UAV is a fast target and the scenario produces lots of multi-path detections.

Fig. 22 shows the track and the reconstructed micro-Doppler signature of the car. A difference between the NN- and PDAF-based trackers can be seen between second 7 to 10. The car has multiple strong and separated reflection centers. When the car entered the scenario, this resulted in two individual clusters spawning two tracks for the car. The NN approach allows only for a single cluster to be used as update of a track. Thus, the assignments made to the wrongfully started track are missing between second 7 and 10. The JPDAF did not leave such a hole.

VIII. COMPUTATIONAL DEMAND

We implemented all algorithms presented in this paper in Matlab 2016a, running on an Intel i7-2600k desktop CPU. The fastest evaluation we could achieve was a real-time[1] evaluation using single precision FFTs with $Z_A = 32$ and grid-clustering for the measurement scenarios with the low data-rate, i.e. the scenario with the bicyclist, the cat and the quadcopter.

The computational demand is mainly governed by the 3D-FFTs, the additions and multiplications during the calculation of the distance metrics for EDBSCAN, and the matrix inversions during the calculation of the chi-square track score in the PDAF tracking stage. Fluctuation of necessary CPU time originates from the dependency of the computational demand on the number of detections. While calculation of the FFTs can be done in constant time, the computational demand of the clustering stage and the PDAF tracker depends on the number of detections. Computational demand, of course, varies with the threshold level and also varies over the discrete time k and the scenario.

[1] We use the definition of real-time in digital signal processing. This means that the processing time is shorter than the time it takes to capture the data.

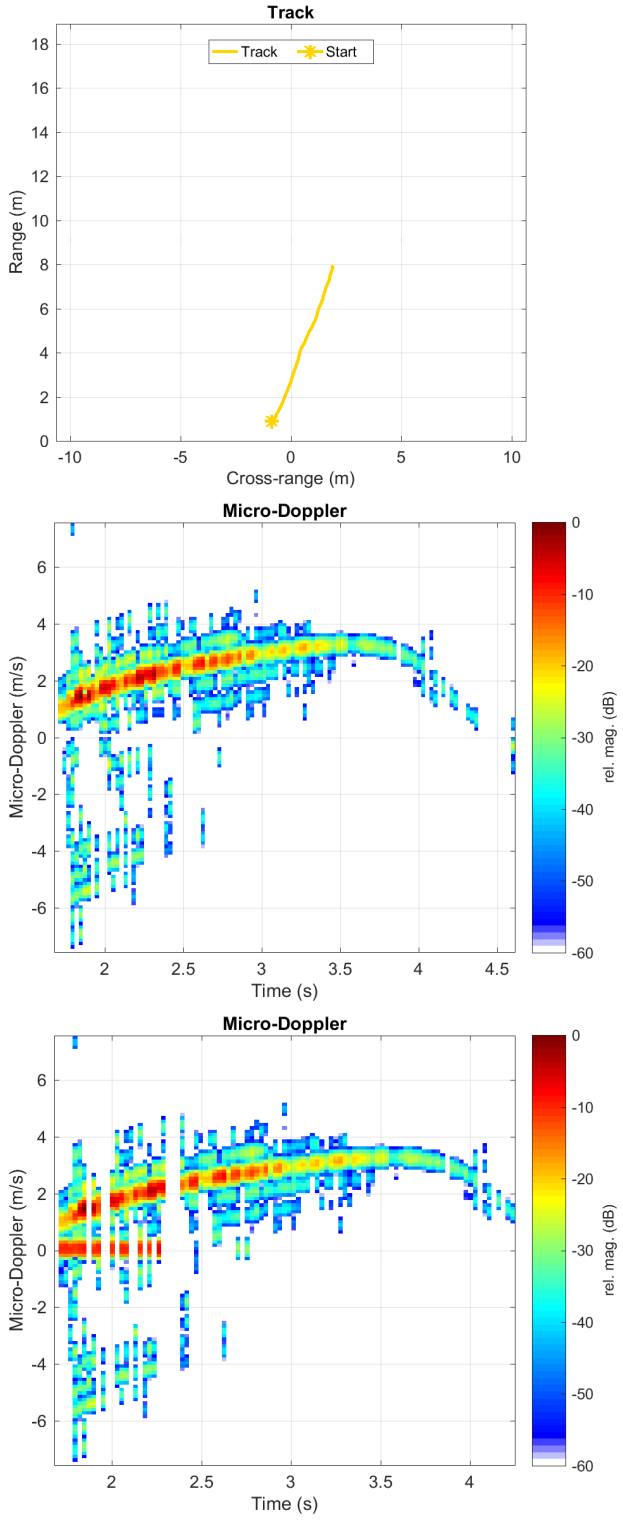

FIGURE 19. Snapshot of a tracking result of a small UAV Blade 200 QX flying indoors. Most of the time the rotor blades are not visible. Thus the UAV is mainly a point target. If the fast revolving rotor blades are at an angle where they reflect back to the radar system, their velocity will cover the whole unambiguous region.

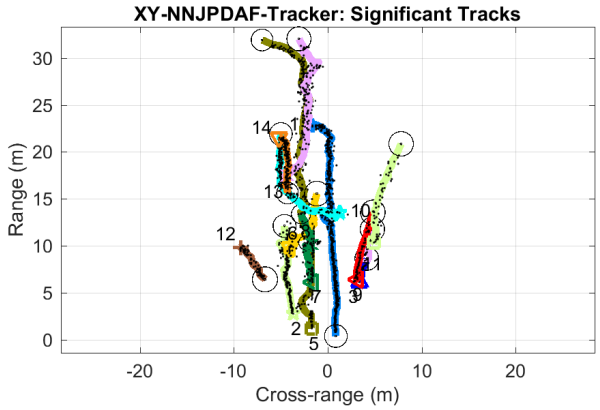

FIGURE 20. Detected tracks of the persons and the car in the underground parking lot. Due to the vast number of multipath reflections in this scenario, more ghost tracks have been formed compared to the measurements in the pedestrian area. However, the car and both pedestrians could be successfully detected and tracked with our framework.

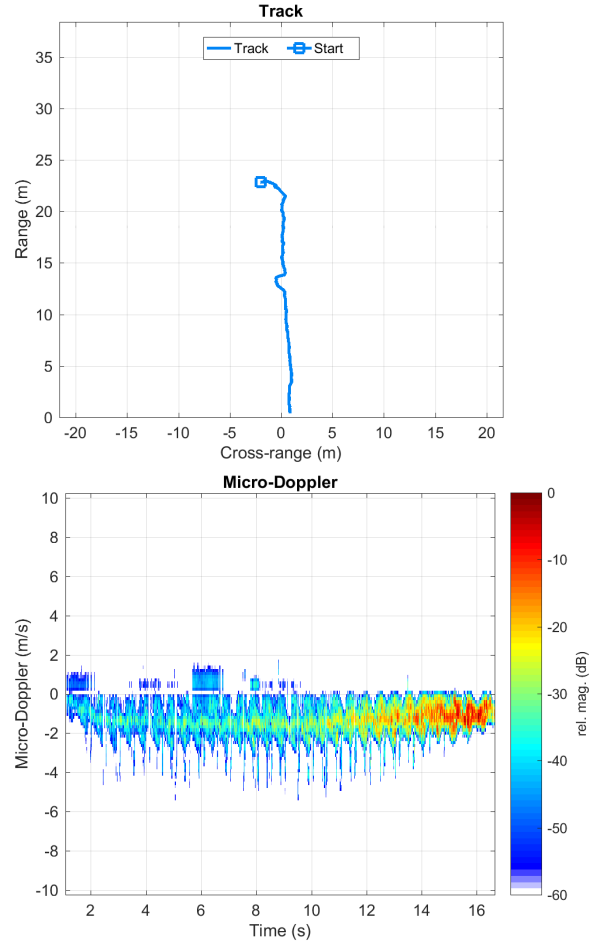

FIGURE 21. The track and reconstructed micro-Doppler signature of a pedestrian in the underground area. Despite the high amount of clutter in this scenario, the pedestrian's track could be followed. The reconstructed micro-Doppler shows a clear signature of a single pedestrian, although the car and another pedestrian where moving simultaneously.

A significant speed up can be expected by offloading the calculations of the FFTs, the thresholding, the distance metric calculations as well as the matrix inversions to graphics processing units (GPUs) or even field-programmable gate arrays (FPGAs). Also doing fast neighborhood searches during the calculation of the EDBSCAN by exploiting an

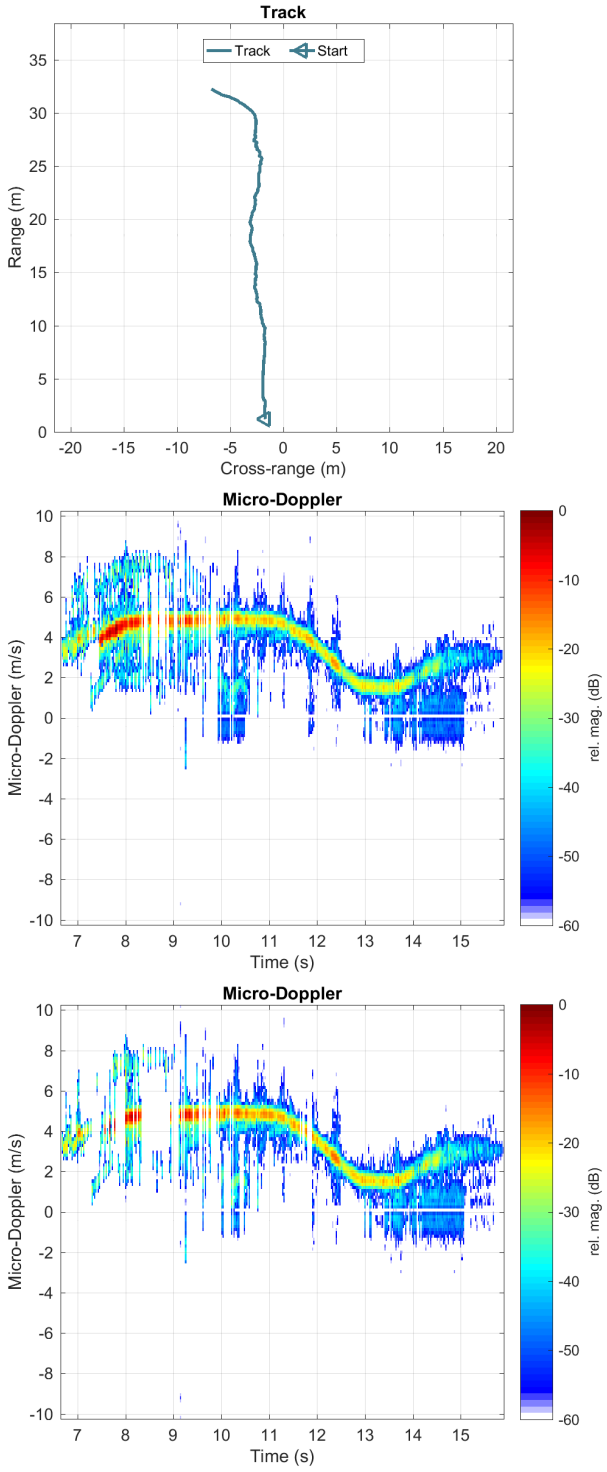

FIGURE 22. The track and the reconstructed micro-Doppler signature of the car in the underground parking lot. The reflections of the rim are clearly visible from second 7 to 9. The JPDAF assignment method was used to generate the top micro-Doppler plot, whereas for the bottom one, the NNJPDAF was used. Between second 7 and 10, the NN approach did some assignments to a wrongfully started track, thus leaving a small gap in the micro-Doppler signature.

R* tree data-structure [33], [34] can reduce the necessary CPU time. Finally, computation time reduction can be achieved by choosing a programming language that efficiently allows pipelining.

However, implementing such optimizations is out of the scope of the work presented in this paper.

IX. CONCLUSION AND OUTLOOK

In this paper, we have shown a method for jointly estimating micro-Doppler signatures and tracks of multiple targets present simultaneously in a scenario. Features suitable for classification of the individual targets could be extracted. A CFAR algorithm is not needed. Two types of clustering stages for target recognition have been presented. Grid clustering allows fast computation while EDBSCAN can be configured in more detail.

An MTT framework with different assignment methods has been presented which allows for tracking individual targets over time. The principle has been shown to work at least for pedestrians, bicyclists, UAVs, cars, and cats in indoor as well as outdoor scenarios. All tested assignment approaches produce satisfying results. However, NN approaches have a slightly higher chance to leave holes in the micro-Doppler signature whereas PDAF approaches have a slightly higher chance to wrongfully add a cluster to a track. Post-processing, like classification, might be sensitive to this different behavior.

Depending on the measurement scenario, real-time evaluation is possible on common PC hardware. Computational demand is mainly governed by the calculation of the 3D-FFTs, which can be offloaded to hardware in future works.

With the presented methods, micro-Doppler plots, suitable for classification [7], [10] can be produced even when multiple targets are in the scenario. As possible candidates for feature extraction, we exemplarily presented the mean and STD (along with their STFFTs) of the micro-Doppler signatures of the moving target after separation. However, other signatures can be easily generated with the algorithms present in this paper.

Future work include classification based upon the acquired information and implementing our suggested code optimizations.

ACKNOWLEDGMENT

This work has been partially supported by the Austrian COMET-K2 programme of the Linz Center of Mechatronics (LCM), and was funded by the Austrian federal government and the federal state of Upper Austria.

The authors would like to thank Inras GmbH for providing the hardware and support for the measurement system and Infineon Technologies for providing the MMICs.

REFERENCES

[1] S. Heuel and H. Rohling, "Two-stage pedestrian classification in automotive radar systems," in *Proc. IEEE Int. Radar Symp. (IRS)*, Sep. 2011, pp. 477–484.
[2] G. Allister and S. Ladyman, "Puffin crossings good practice guide," U.K. Dept. Transp., London, U.K., Tech. Rep., 2006.
[3] *Traffic Advisory Leaflet: The Installation of Puffin Pedestrian Crossings*, U.K. Dept. Transp., London, U.K., Jan. 2002.
[4] Y. Wu, C. Shi, X. Zhang, and W. Yang, "Design of new intelligent street light control system," in *Proc. 8th IEEE Int. Conf. Control Autom. (ICCA)*, Jun. 2010, pp. 1423–1427.

[5] P. Abrol, "Design of traffic flow based street light control system," *Int. J. Comput. Appl.*, vol. 72, no. 18, pp. 32–37, 2013.
[6] K. S. Sudhakar, A. A. Anil, K. C. Ashok, and S. S. Bhaskar, "Abhale Amol and Chetan Ashok, Kudake and Bhaskar, Shirsath Shravan," *Int. J. Emerg. Technol. Adv. Eng.*, vol. 3, no. 5, pp. 188–189, 2013.
[7] Y. Kim and H. Ling, "Human activity classification based on micro-Doppler signatures using a support vector machine," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 5, pp. 1328–1337, May 2009.
[8] D. Fairchild and R. Narayanan, "Classification of human motions using empirical mode decomposition of human micro-Doppler signatures," *IET Radar, Sonar Navigat.*, vol. 8, no. 5, pp. 425–434, 2014.
[9] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. West Lafayette, IN, USA: Academic, 2013.
[10] R. M. Narayanan and M. Zenaldin, "Radar micro-Doppler signatures of various human activities," *IET Radar, Sonar Navigat.*, vol. 9, no. 9, pp. 1205–1215, 2015.
[11] S. Heuel and H. Rohling, "Pedestrian recognition based on 24 GHz radar sensors," in *Proc. 11th Int. Radar Symp. (IRS)*, 2010, pp. 1–6.
[12] A. Bartsch, F. Fitzek, and R. H. Rasshofer, "Pedestrian recognition using automotive radar sensors," *Adv. Radio Sci.*, vol. 10, no. 4, pp. 45–55, 2012.
[13] E. Bosina, "Bewegungscharakteristika des Fußgängerverkehrs unter besonderer Berücksichtigung der Beschleunigung," M.S. thesis, Dept. (RALI), Univ. Natural Resour. Life Sci., Vienna, Vienna, Austria, Oct. 2012.
[14] F. Fölster, H. Rohling, and H. Ritter, "Observation of a walking pedestrian with a 24GHz automotive radar sensor," in *Proc. German Microw. Conf. (GeMiC)*, 2006, pp. 1–5.
[15] H. Ritter and H. Rohling, "Pedestrian detection based on automotive radar," in *Proc. IET Int. Conf. Radar Syst.*, 2007, pp. 1–4.
[16] E. Schubert, M. Kunert, A. Frischen, and W. Menzel, "A multi-reflection-point target model for classification of pedestrians by automotive radar," in *Proc. IEEE 11th Eur. Radar Conf. (EuRAD)*, Oct. 2014, pp. 181–184.
[17] S. Björklund, H. Petersson, A. Nezirovic, M. B. Guldogan, and F. Gustafsson, "Millimeter-wave radar micro-Doppler signatures of human motion," in *Proc. IEEE Int. Radar Symp. (IRS)*, Sep. 2011, pp. 167–174.
[18] P. Molchanov, A. Vinel, J. Astola, and K. Egiazarian, "Radar frequency band invariant pedestrian classification," in *Proc. IEEE 14th Int. Radar Symp. (IRS)*, vol. 2. Jun. 2013, pp. 740–745.
[19] I. Prokopenko, K. Prokopenko, and I. Martynchuk, "Moving objects recognition by micro-Doppler spectrum," in *Proc. IEEE 16th Int. Radar Symp. (IRS)*, Jun. 2015, pp. 186–190.
[20] J. J. M. de Wit, R. I. A. Harmanny, and G. Prémel-Cabic, "Micro-Doppler analysis of small UAVs," in *Proc. 9th Eur. Radar Conf. (EuRAD)*, Oct./Nov. 2012, pp. 210–213.
[21] O. H. Y. Lam, R. Kulke, M. Hagelen, and G. Möllenbeck, "Classification of moving targets using mirco-Doppler radar," in *Proc. 17th Int. Radar Symp. (IRS)*, May 2016, pp. 1–6.
[22] P. Molchanov, K. Egiazarian, J. Astola, R. I. A. Harmanny, and J. J. M. de Wit, "Classification of small UAVs and birds by micro-Doppler signatures," in *Proc. Eur. Radar Conf. (EuRAD)*, Oct. 2013, pp. 172–175.
[23] E. Schubert, F. Meinl, M. Kunert, and W. Menzel, "High resolution automotive radar measurements of vulnerable road users—Pedestrians & cyclists," in *Proc. IEEE MTT-S Int. Conf. Microw. Intell. Mobility (ICMIM)*, Apr. 2015, pp. 1–4.
[24] R. Feger, C. Pfeffer, W. Scheiblhofer, C. M. Schmid, M. J. Lang, and A. Stelzer, "A 77-GHz cooperative radar system based on multi-channel FMCW stations for local positioning applications," *IEEE Trans. Microw. Theory Techn.*, vol. 61, no. 1, pp. 676–684, Jan. 2013.
[25] H. J. Ng, R. Stuhlberger, L. Maurer, T. Sailer, and A. Stelzer, "Low phase noise 77-GHz fractional-N PLL with DLL-based reference frequency multiplier for FMCW radars," in *Proc. IEEE Eur. Microw. Integr. Circuits Conf. (EuMIC)*, Oct. 2011, pp. 196–199.
[26] C. Wagner *et al.*, "A 77GHz automotive radar receiver in a wafer level package," in *Proc. IEEE Radio Freq. Integr. Circuits Symp. (RFIC)*, Jun. 2012, pp. 511–514.
[27] H. Knapp *et al.*, "Three-channel 77 GHz automotive radar transmitter in plastic package," in *Proc. IEEE Radio Freq. Integr. Circuits Symp. (RFIC)*, Jun. 2012, pp. 119–122.
[28] A. Haderer, *INRAS Products—Radarbook*, INRAS GmbH, Linz, Austria, 2014.
[29] D. E. Barrick, "FM/CW radar signals and digital processing," Nat. Ocean. Atmos. Administration, DTIC Document, Boulder, Colorado, Tech. Rep. AD-774 829, 1973.
[30] H. L. Van Trees, *Optimum Array Processing: Part IV of Detection, Esti-AQ:9 1126 Mation, and Modulation Theory*. New York, NY, USA: SONS INC, 2002.
[31] P. Antonik *et al.*, "Intelligent use of CFAR algorithms," Dept. Electr. Comput. Eng., Syracuse Univ., Syracuse, NY, USA, Tech. Rep. RL-TR-93-75, May 1993.
[32] H. Rohling, "Radar CFAR thresholding in clutter and multiple target situations," in *Proc. IEEE Aerosp. Electron. Syst. Mag.*, vol. 4. 1983, pp. 608–621.
[33] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proc. 2nd Int. Conf. Knowl. Discovery Data Mining*, vol. 96. 1996, pp. 226–231.
[34] J. Sander, "Generalized density-based clustering for spatial data mining," Ph.D. dissertation, Fakultät Math. Informatik, Ludwig Maximilian Univ. Munich, Munich, Germany, Sep. 1998.
[35] T. Wagner, R. Feger, and A. Stelzer, "Modification of DBSCAN and application to range/Doppler/DoA measurements for pedestrian recognition with an automotive radar system," in *Proc. Eur. Radar Conf. (EuRAD)*, Sep. 2015, pp. 269–272.
[36] T. Wagner, R. Feger, and A. Stelzer, "A fast grid-based clustering algorithm for range/Doppler/DoA measurements," in *Proc. Eur. Radar Conf. (EuRAD)*, Sep. 2016, pp. 105–108.
[37] S. S. Blackman, *Multiple-Target Tracking With Radar Applications*, vol. 1. Dedham, MA, USA: Artech House, 1986, p. 463.
[38] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME, D, J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, 1960.
[39] G. L. Smith, S. F. Schmidt, and L. A. McGee, *Application of Statistical Filter Theory to the Optimal Estimation of Position and Velocity on Board a Circumlunar Vehicle*. Washington, DC, USA: NASA, 1962.
[40] A. H. Jazwinski, *Stochastic Processes and Filtering Theory*. New York, NY, USA: Academic, 1970.
[41] F. Daum, "Nonlinear filters: Beyond the Kalman filter," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 20, no. 8, pp. 57–69, Aug. 2005.
[42] D. F. Bizup and D. E. Brown, "The over-extended Kalman filter—Don't use it!" in *Proc. 6th Int. Conf. Inf. Fusion*, vol. 1. 2003, pp. 40–46.
[43] N. Yamada, Y. Tanaka, and K. Nishikawa, "Radar cross section for pedestrian in 76GHz band," in *Proc. Eur. Microw. Conf.*, vol. 2. Oct. 2005, pp. 1–4.
[44] J. Fortuny-Guasch and J.-M. Chareau, "Radar cross section measurements 1161 of pedestrian dummies and humans in the 24/77 GHz frequency bands," 1162 Joint Res. Centre, Via Enrico Fermi 2749, TP 723, 21027 Ispra (VA), Italy, Pub. Office Eur. Union, Tech. Rep. EUR 25762 EN, 2013.
[45] D. Lerro and Y. Bar-Shalom, "Tracking with debiased consistent converted measurements versus EKF," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 29, no. 3, pp. 1015–1022, Jul. 1993.
[46] R. J. Fitzgerald, "Development of practical PDA logic for multitarget tracking by microprocessor," in *Proc. Amer. Control Conf.*, Jun. 1986, pp. 889–898.
[47] D. B. Reid, *A Multiple Hypothesis Filter for Tracking Multiple Targets in a Cluttered Environment*. Lockheed Missiles & Space Company, Tech. Rep. D-560254, 1977.
[48] S. S. Blackman, "Multiple hypothesis tracking for multiple target tracking," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 19, no. 1, pp. 5–18, Jan. 2004.
[49] Y. Bar-Shalom and E. Tse, "Tracking in a cluttered environment with probabilistic data association," *Automatica*, vol. 11, no. 5, pp. 451–460, 1975.
[50] E. Schubert, F. Meinl, M. Kunert, and W. Menzel, "Clustering of high resolution automotive radar detections and subsequent feature extraction for classification of road users," in *Proc. IEEE 16th Int. Radar Symp. (IRS)*, Jun. 2015, pp. 174–179.
[51] G. Brooker, D. Johnson, J. Underwood, J. Martinez, and L. Xuan, "Using the polarization of millimeter-wave radar as a navigation aid," *J. Field Robot.*, vol. 32, no. 1, pp. 3–19, 2015.

THOMAS WAGNER (S'08–GM'12) was born in Linz, Austria, in 1985. He received the Dipl.Ing. (M.Sc.) degree in mechatronics from Johannes Kepler University Linz, Linz, in 2011. Since 2011, he joined the Institute for Communications and Information Engineering, Johannes Kepler University Linz, as a Research Assistant, where he became a member of the Christian Doppler Laboratory for Integrated Radar Sensors.

REINHARD FEGER (GSM'08–M'10) was born in Kufstein, Austria, in 1980. He received the Dipl.Ing. (M.Sc.) degree in mechatronics and the Ph.D. degree in mechatronics from Johannes Kepler University Linz, Linz, Austria, in 2005 and 2010, respectively. In 2005, he joined the Institute for Communications and Information Engineering, Johannes Kepler University Linz, as a Research Assistant. In 2007, he became a member of the Christian Doppler Laboratory for Integrated Radar Sensors, Johannes Kepler University Linz, where he is currently an Assistant Professor with the Institute for Communications Engineering and RF-Systems. His research topics are radar signal processing, and radar system design for industrial and automotive radar sensors.

ANDREAS STELZER (M'00) received the Dipl.Ing. degree in electrical engineering from the Technical University of Vienna, Vienna, Austria, in 1994, and the Ph.D. degree (Hons.) in mechatronics from Johannes Kepler University Linz, Austria, in 2000. In 2003, he became an Associate Professor with the Institute for Communications Engineering and RF Systems, Johannes Kepler University Linz. Since 2007, he has been the Head of the Christian Doppler Research Laboratory for Integrated Radar Sensors. Since 2008, he has been a Key Researcher with the Austrian Center of Competence in Mechatronics, where he is responsible for numerous industrial projects. Since 2011, he has been a Full Professor with Johannes Kepler University Linz, where he is heading the Department for RF-Systems. He has authored or co-authored over 320 journal and conference papers. His research is focused on microwave sensor systems for industrial and automotive applications, radar concepts, SiGe-based circuit design, microwave packaging in eWLB, RF and microwave subsystems, surface acoustic wave sensor systems and applications, and digital signal processing for sensor signal evaluation. He is a member of the Austrian ÖVE. He is also a member of the IEEE Microwave Theory and Techniques (MTT)-S, the International Menopause Society, and the Circuits and Systems Society. He serves as the IEEE Distinguished Microwave Lecturer from 2014 to 2016. He has served as an Associate Editor of the IEEE Microwave and Wireless Components Letters. He currently serves as the Co-Chair of MTT-27 Wireless-Enabled Automotive and Vehicular Applications. He was a recipient of several awards, including the 2008 IEEE MTT Society Outstanding Young Engineer Award and the 2011 IEEE Microwave Prize. Furthermore, he was also a recipient of the 2012 European Conference on Antennas and Propagation Best Measurement Paper Prize, the 2012 Asia Pacific Conference on Antennas and Propagation Best Paper Award, the 2011 German Microwave Conference Best Paper Award, the EEEfCOM Innovation Award, and the 2003 European Microwave Association Radar Prize of the European Radar Conference.