

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/332814465>

Human Target Detection, Tracking, And Classification Using 24 GHz FMCW Radar

Article in IEEE Sensors Journal · May 2019

DOI: 10.1109/JSEN.2019.2914365

CITATIONS

10

READS

2,030

4 authors:



Christoph Will
Infineon Technologies
18 PUBLICATIONS 184 CITATIONS

[SEE PROFILE](#)



Prachi Vaishnav
Infineon Technologies
2 PUBLICATIONS 10 CITATIONS

[SEE PROFILE](#)



Abhiram Chakraborty
Infineon Technologies
10 PUBLICATIONS 56 CITATIONS

[SEE PROFILE](#)



Avik Santra
Infineon Technologies
43 PUBLICATIONS 108 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Deep Learning for Target Recognition in Synthetic Aperture Radar (SAR) images [View project](#)



Adaptive Waveform Design for MIMO Radar [View project](#)

Human Target Detection, Tracking, And Classification Using 24 GHz FMCW Radar

Christoph Will, *Student Member, IEEE*, Prachi Vaishnav, Abhiram Chakraborty, *Member, IEEE*,
and Avik Santra, *Senior Member, IEEE*

Abstract—This paper presents a millimeter wave radar in the 24-GHz ISM band for detection, tracking, and classification of human targets. Linear frequency modulation of the transmit signal and two receive antennas enable distance and angle measurements, respectively. Multiple consecutive frequency chirps are used for target velocity calculation. Hardware as well as firmware concepts of the proposed system are described in detail. Various algorithms for human detection and tracking are investigated and combined to a new signal processing routine optimized for compactness and low-power to run on a microcontroller. Additionally, a novel Doppler-compensated angle-of-arrival estimation method as well as a one-class support vector machine for human classification are proposed to further enhance the human detection and tracking performance. The achieved performances of the designed hardware and the implemented algorithm are verified in extensive measurements. The distance and angle errors of the realized radar sensor are at most 25 cm along a measurement range of 18 m and 10° for a two-sided angle sweep of 65°, respectively. The achieved range resolution is 0.9 m. Dedicated verifications of the most important signal processing routines are presented to verify their functionality and experiments with several human targets illustrate the performance and limits of the overall tracking algorithm. It is shown that range, velocity, and angle of up to five humans are correctly detected and tracked. The presented one-class classifier successfully distinguishes human targets from other quasi-static targets like trees and shadowing effects of human subjects on walls.

Index Terms—Millimeter wave radar, radar detection, radar tracking.

I. INTRODUCTION

AUTOMATIC detection of human targets offers a variety of applications in modern society. In smart homes, human target detection can be utilized in light sensors for instance to automatically switch the light on or off if a person enters or leaves the room, respectively. This idea can be enlarged to smart cities with smart street lighting. In the industrial area, fast detection of humans helps robots to avoid collision and security systems to find intruders. Also in the automotive area, there is always a need for enhanced safety and security, especially regarding autonomous cars. Human target detection in general is enabled by advanced sensor systems consisting of a hardware platform and a dedicated firmware, which adjusts the component settings and gathers raw sensor data. Depending on the system concept, either the firmware itself pre-processes these data and evaluates them by algorithms or it communicates with a host processor, that takes charge of the signal processing.

All authors are with Infineon Technologies AG, 85579 Neubiberg, Germany.
The authors declare no conflict of interest.

Various sensor concepts can be used for human target detection. The most obvious approach is to use a camera together with an image processing and feature detection algorithm [1], [2], [3]. Advanced algorithms like a combination of temporal differencing and template matching enable tracking of the classified targets [1]. Special RGB-D cameras improve the performance by utilizing the additional depth information [4]. Next to these single camera systems, also distributed camera systems for three dimensional tracking have been published [5]. All camera-based systems suffer from high costs and the lack of privacy. Alternative sensor concepts use acoustic waves [6], [7], millimeter waves [8] or non-visible light [9], [10] and therefore ensure more privacy. Apart from camera-based sensor fusion [9] and lidar systems [11], [12], which have a high accuracy and precision, these sensors are additionally considerably cheaper. This paper focuses on radar-based systems as their electromagnetic waves have some advantages compared to the other concepts. While ultrasonic systems as well as cameras require line-of-sight, pulsed, broadband electromagnetic waves can penetrate walls [13]. In contrast to infrared or lidar sensors, radar systems also work in harsh environments and are less affected by rain, dust or fog [14]. Besides, they can be aesthetically hidden without affecting the operating performance. Furthermore, radar systems can detect micro motions like vital signs by interferometric signal evaluation even in the micrometer range [15], [16]. Multiple targets can be detected by signal modulation [8] or multiple-input multiple-output (MIMO) systems [17].

Radar-based human tracking has been researched for several years meanwhile, using diverse radar concepts. Frequency-modulated continuous wave (FMCW) [18], [19], [20], [21] and ultra-wideband (UWB) [8], [22] radar sensors use multiple frequencies to obtain distance information of one or more targets. For short-range human tracking up to 5 m, a 24-GHz FMCW radar based on Six-Port interferometry [14] utilizing vital-Doppler standard deviation was published [18], [19]. Other FMCW radars use inverse synthetic aperture radar (ISAR) images for indoor tracking of a single person [20] or a Kalman filter with track management table and adaptive thresholding for multi-target tracking [21]. A Kalman filter is also used in UWB radar-based multi-target tracking combined with an expectation maximization algorithm [8]. The UWB radar in [22] uses a motion filter and tracking point cohesion to track one target even behind a wall. Another common approach is to combine several FMCW [23], [24] or UWB [25], [26], [27] radar sensors to a multistatic radar system. Target detection by motion filtering or an ordered statistics

constant false alarm rate (CFAR) detector each combined with tracking point cohesion enables the tracking of one person through a wall [23] or three humans in a cluttered environment [24], respectively. Two UWB receive (RX) radar sensors are used each for single target tracking in Cartesian coordinate system [25], indoor human tracking [26], and through-wall tracking of two persons [27]. Alternative approaches use pulse-Doppler radar together with an alpha-beta filter for real-time embedded tracking of two targets [28], synthetic aperture radar (SAR) with ground moving target indication for automatic target recognition from airplanes [29], MIMO radar with two-dimensional MUSIC (MULTiple SIgnal Classification) correlation for indoor multi-path environment [30], and passive acoustic radar with moving target detection and tracking to measure coordinates of a flying multicopter [31]. 2D tracking of two humans by an unmodulated CW radar sensor is enabled by three RX antennas and the evaluation of the direction of arrival (DOA) [32].

Radar-based classification is an arising topic, while most publications focus on specific human activities [33] or gestures [34] and on target recognition of pre-defined classes like small drones and birds [35]. A rather open research topic is human target classification, which means to distinguish humans from other targets with radar systems. Instantaneous presence detection using Doppler radar detects if a human is present or not [36]. Other publications limit the targets to pre-defined classes like "one person", "two persons", and "vehicle" [37]. In [38] radar classification of indoor targets to distinguish human targets from harsh clutter using through-the-wall radar systems is investigated. Doppler spectrogram-based features are utilized in [39] to distinguish humans from dogs, bicycles, and cars. All of these publications pre-define two or more specific target classes, some combine several targets in one non-human class to generate a two-class problem. Most of these classifiers are based on a support vector machine (SVM) [33], [35], [38], [36], [39], since an SVM is a convenient classifier for binary problems like human classification. The disadvantage of a two-class SVM with one class for humans and the other for non-human targets is the required pre-selection of the non-human targets, which are used to train the classifier. The classification for un-trained targets is not predictable and can easily result in false human target detection. This issue can be solved by one-class classifiers for outlier detection. A one-class SVM is trained only with data of the desired target class, in the presented human targets. With appropriate features, all non-human targets should be detected as outliers. Examples for one-class SVMs from literature are image-based handwritten signature verification [40], sheet metal forming limit detection with camera [41] or oil spill detection using airborne SAR [42].

This paper presents an FMCW radar sensor for multi-target tracking with all algorithms to detect and track up to five humans running on a microcontroller. In comparison to the state-of-the-art, the researched signal processing routine is tailored and optimized to detect and track more than three targets in an outdoor environment within an angular range of over 100° and a distance up to 20m. While some parts of the proposed signal processing routine are smartly

combined already-known procedures, two novel algorithms are introduced to ensure high robustness against losing tracks and assigning ghost targets. First, an advanced angle estimation algorithm for humans targets is proposed, since common angle estimation using the range-Doppler peak often results in wrong estimates as the individually moving parts of a human body induce different angles even and especially for (quasi-)static humans. The presented algorithm uses the centroid of the Doppler spread along the slow-time and utilizes a Doppler-compensation for static humans. Second, a one-class SVM classifier with handcrafted features is introduced to distinguish quasi-static humans from other static but slightly moving targets (like trees in outdoor scenarios) to mitigate false human target detection. To the author's knowledge, the proposed one-class classifier is the first radar-based approach to distinguish humans from any other target without using pre-selected target classes. The complete proposed algorithm pipeline is simultaneously optimized with respect to compactness and energy-efficiency to enable an implementation on an ARM M4 cortex real-time embedded platform. The implementation of all signal processing routines on the micro-controller along with measurements in complex outdoor scenarios shifts laboratory research to real-world scenarios and underlines the feasibility of the proposed system.

The paper is structured as follows. In Sec. II, the challenges of this research are described. The hardware and firmware concept for radar-based human target detection and tracking are designed on system level in Sec. III. The sections IV, V, and VI explain the algorithms for human detection, tracking, and classification, respectively. The measurement results are shown in Sec. VII and compared to the state-of-the-art in Sec. VIII. A conclusion is given in Sec. IX.

II. CHALLENGES OF HUMAN DETECTION & TRACKING

Human detection and tracking with radar systems is of high interest regarding consumer, industrial, and automotive applications. However, a robust detection of humans in a noisy clutter environment is a challenging topic. Major challenges are multi-path reflections especially in bounded environments and the differentiation of human-related movements from non-human movements like animals or trees. Since millimeter-wave radar systems are very sensitive to slight motions, multi-path reflections from static objects and walls induce ghost targets at their positions. It is a challenging task to distinguish human targets from these ghost targets.

Furthermore, in contrast to completely static targets like walls or clutter, multi-path effects and speckle lead to high variations of the received signal strength reflected from a human [43]. Destructive interferences can temporarily extinguish the backscattered signal from the human target. In this case, the signal strength of the human target can be way below the detection threshold for several consecutive measurements, whereby target information is lost. Thus, in order to be able to track a human target, it is required to account for this issue. The reflected signal from targets like walls always possess high signal strength as compared to human targets and hence, for a human target being very close to such targets, the signal

strength of a human target is easily overshadowed by such strong targets. Therefore, it is also necessary to diminish or extinguish the backscattered signals from large static targets by intelligent signal processing routines.

III. SYSTEM DESIGN

The hardware and firmware concepts of the proposed radar system are designed in this section.

A. Hardware Concept

Figure 1 shows the block diagram of the microwave multi-channel FMCW radar sensor designed to work in the 24-GHz ISM band. The sensor consists of a radio frequency (RF) unit formed by a 24-GHz monolithic microwave integrated circuit (MMIC) and microstrip patch antennas, a low noise fractional-N phase-locked loop (PLL) for frequency stabilization and FMCW ramp generation, a programmable gain low noise analog baseband (ABB) section and a digital part consisting of Cortex-M4-based signal processing and control unit. The sensor is powered by 5V direct current (DC) supply and on-board low noise low dropout regulators (LDOs) are used to supply each of the sensor sections. The complete radar system is fabricated on a 10-mil-thick *Rogers 4350B* substrate-based single printed circuit board (PCB) comprising six metal layers. The MCU on board is programmed via a serial-wire debug (SWD) interface using an external breakable debugger unit not shown in the block diagram. The sub-sections below describe each of the functional blocks in detail.

1) *24-GHz Front-end MMIC*: The RF unit consists of the 24-GHz multichannel transceiver MMIC *BGT24MTR12* with one transmit and two receive sections, impedance matching structures, harmonic filters, and microstrip patch antennas. The MMIC is designed in a Silicon-Germanium (SiGe) Bipolar technology [44] with f_t/f_{max} of 200/250GHz from *Infineon Technologies* and is enclosed in a small 32-pin VQFN package. The transceiver consists of a voltage controlled oscillator (VCO) with prescaler outputs for frequency generation and control, a transmitter (TX) chain including a power amplifier (PA) and buffer amplifiers for the TX and local oscillator (LO) sections, and two homodyne quadrature receiver sections with low noise amplifiers (LNA) and highly-linear modified gilbert type mixers. The multiple receive channel configuration enables angle-of-arrival (AoA) estimation of the signal reflected from the radar target.

The VCO is a free-running, fundamental frequency oscillator with a low phase noise of -85 dBc/Hz at 100-kHz offset. It is controlled by two tuning inputs, one for coarse pre-adjustment (COARSE) and one for fine-tuning (FINE). The VCO is followed by two prescaler blocks. The first prescaler has an output frequency of 1.5 GHz and the second prescaler has a 23-kHz square-wave output. The prescalers enable to interface the VCO with an external frequency control circuit like a PLL or a digital-to-analog converter (DAC).

The TX section includes a power amplifier with differential output. Its typical output power is +11 dBm. A part of the TX signal is used as LO signal for the mixers in the receiver.

A passive poly-phase filter (PPF) generates quadrature signals for the mixers LO input. The receiver sections have a single-sideband noise figure (NF_{SSB}) of 12 dB and a voltage conversion gain (CG) of 26 dB. Figure 2 shows the performance characteristics of the transceiver over temperature at 24.125 GHz.

Depending on the target distance and radar cross section (RCS), the signal received at the sensor may completely saturate the analog-to-digital converters (ADCs) in the signal processing unit. Therefore, depending on the strength of the received signal, the transceiver needs to reconfigure the TX output power and the RX conversion gain. Integrated power sensors based on RF to DC converters at the TX ports enable to monitor the output power and adjust it via a 3-bit serial peripheral interface (SPI) setting. The transmitter output power can be reduced by 9 dB from its maximum value. On the receiver side the gain of the LNA can be reduced by a single gain-step of 5 dB. Further gain configuration at the receiver side is provided by programmable gain amplifiers (PGAs) in the intermediate frequency (IF) stage as described in Sec. III-A3. The built-in modified Gilbert-based quadrature downconversion mixers help to achieve a high input 1-dB compression point (IP1dB) of -12 dBm and convert the RF signal directly to zero-IF. An on-chip temperature sensor constantly monitors the temperature of the MMIC. The chip is controlled via a simple 3-wire SPI. When fully enabled, the chip consumes approximately 700 mW from a 3.3-V power supply. Since the MMIC does not support a power down mode, the transceiver is connected to the power supply over a PMOS switch, which enables to operate the sensor in duty-cycle mode for low power applications.

The transmitter output of the MMIC is differential. These differential outputs are first connected over matching structures followed by a Wilkinson power combiner on the PCB. The matching structures compensate for the bond wire inductance and other parasitic effects due to the VQFN package. Following the power combiner, a microstrip harmonic filter is used to attenuate the harmonics around 48 GHz. The harmonic filter provides an attenuation larger than 20 dB and has a simulated loss of approximately 0.5 dB. The filter path then goes over a DC block and a feedthrough via to the other side of the PCB to the antennas. The simulated loss for the entire RF section connecting the TX output from the MMIC to the antennas on the other side of the board including the vias is approximately 2 dB. There are DC shorts before the feedthrough vias for enhanced electrostatic discharge (ESD) protection. The receiver input of the MMIC is single-ended. The RX input is connected over a matching structure, a DC block and a feedthrough via to the antennas on the other side of the board. The simulated loss for the entire RF section connecting the RX input at the MMIC to the antennas on the other side of the board including the vias is approximately 1 dB. Also in this path, DC shorts before the feedthrough vias enhance ESD protection. All feedthrough vias for ESD protection go from top layer to the adjacent inner ground layer. They are visible in the photography of Fig. 3, which shows the top and bottom view of the sensor module. The size of the module is 45 mm by 50 mm in dimension.

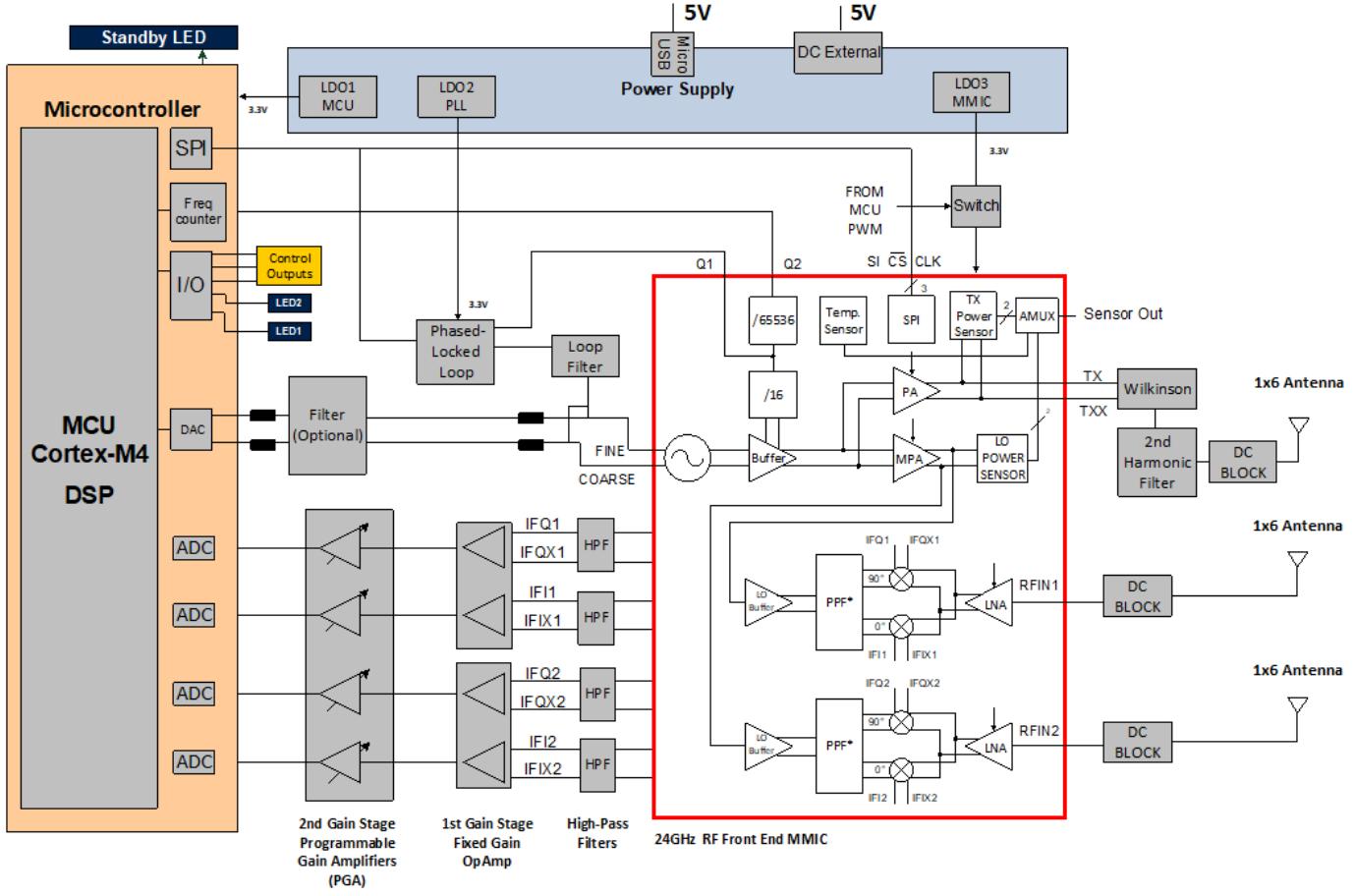


Fig. 1. 24-GHz multi-channel FMCW radar system.

A tapered series-fed 1x6 linear antenna array with low side lobe levels is used for the transmission and reception of the microwave signal. Figure 4 shows the measured transmit field of view of a sensor module. The sensor has a typical effective isotropic radiated power (EIRP) of +18 dBm and a 3-dB half power beam width (HPBW) of 19° and 76° in the E- and H-plane, respectively.

2) Phase-Locked Loop : A fractional-N PLL is used to stabilize the VCO on the 24-GHz MMIC and generate the frequency ramps for target tracking. The PLL consists of a phase detector with programmable frequency dividers for both the high frequency RF input signal and the low frequency reference clock signal. The 1.5-GHz prescaler output from the MMIC is connected to the high frequency RF input of the PLL and an external low phase noise 40-MHz reference oscillator clock is connected to the reference input. The phase detector compares the phase differences between the two input signals and generates a voltage proportional to the phase error which is converted to a correction current by a charge pump. The output of the charge pump is filtered with a third order passive loop filter formed by discrete components and given to the tuning port of the VCO. For the aimed application, the coarse and fine tuning pins of the MMIC are tied together to minimize the tuning sensitivity. The loop filter is designed to support frequency ramps up to 300 us and at the same time fulfill the

ETSI and FCC specifications for unwanted emissions.

3) Analog Baseband Section: Depending on the target in front of the antennas, the analog IF output signal from the MMIC can be very low in amplitude (μ V to mV range). To process these low amplitude signals it is necessary to amplify the IF signals with analog amplifiers. Another commonly occurring issue in a homodyne architecture-based FMCW radar system is the feedthrough of the TX signal into the RX part due to limited isolation on the PCB, also called TX-to-RX leakage. Consequently, there is always a dominating low frequency component at the receiver output of the radar IC. The value of this low frequency component depends on the value of the frequency ramp settings. This low frequency signal is further amplified by the baseband amplifiers and may completely saturate the radar IF chain. This effect is inherent to all FMCW radar systems and cannot be eliminated completely in the analog domain. The TX-to-RX leakage also limits the minimum distance that can be measured by the radar. However, the effect of this effect can be minimized using appropriate filtering. This requires the implementation of filtering stages prior to the amplification of the IF signal in the baseband section. Therefore, the baseband section of this sensor is designed to have a bandpass characteristic.

The MMIC provides both in-phase and quadrature (I/Q) phase IF signals from each of its receive channels. Since

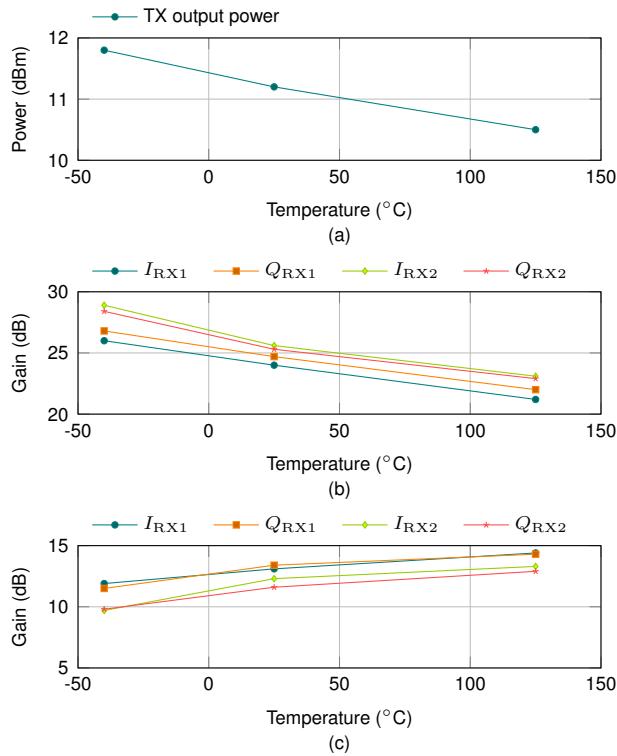


Fig. 2. (a) Transmitter output power, (b) CG, and (c) NFSSB of the BGT24MTRJ2 over temperature at 24.125 GHz.

the I/Q signals are provided differentially from the MMIC, there are four IF output signals per receive channel: IFI, IFIx, IFQ, and IFQx. Each IF path comprises two stages of low noise analog baseband amplifiers with a bandpass characteristic. The first stage consists of two dual-channel low noise operational amplifiers with a gain of 23.5 dB, which also perform the differential to single-ended conversion of each IF signal pair, thereby reducing the eight IF signal lines from the MMIC to four. The second amplification stage consists of four programmable gain amplifiers for each of the IF signals output by the first amplification stage. The PGAs in the IF section

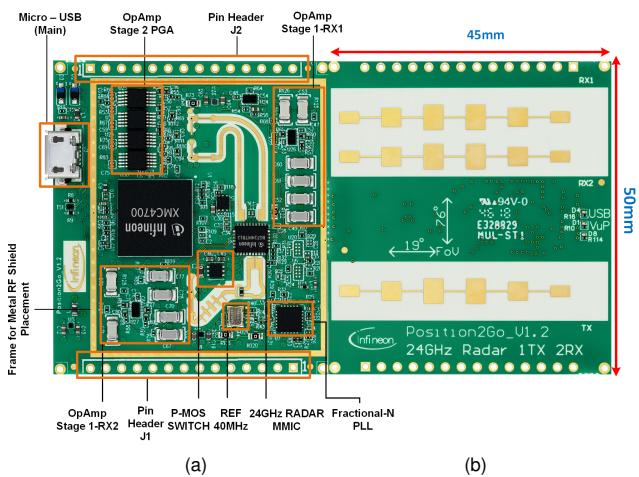


Fig. 3. (a) Top and (b) bottom view of the fabricated sensor module.

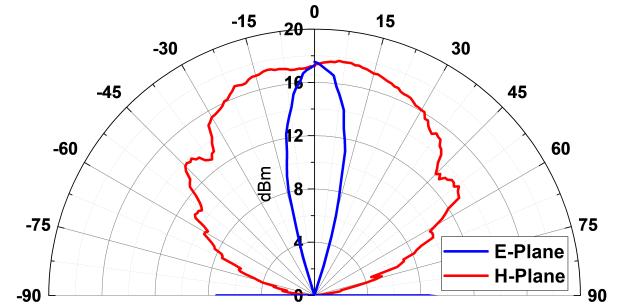


Fig. 4. Measured radiation pattern in E- and H-plane of the sensor.

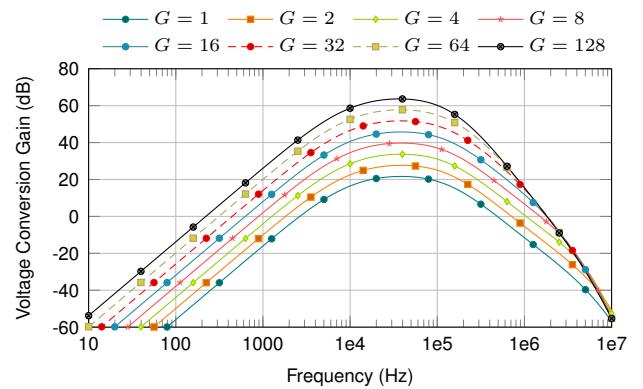


Fig. 5. Simulated transfer characteristics of the analog baseband for various PGA gains.

along with the configurable gain of the MMIC LNAs enable an automatic gain control of the receive section based on the received signal strength from a target. The second IF stage in combination with the first IF stage provide a total IF gain of 65.5 dB with a 3-dB bandwidth from 13 kHz to 105 kHz. Figure 5 shows the simulated transfer characteristics of the analog baseband section for different PGA gain settings. For the current sensor implementation, the desired IF frequencies appear between 10 ... 100 kHz depending on the target distance from the sensor.

4) Signal Processing and Control Unit: The sensor platform uses the 32-bit ARM Cortex-M4 144-MHz microcontroller unit (MCU) XMC4700 in a 194-pin BGA package from *Infineon Technologies*, to perform the radar signal processing. The MCU takes care of communication with all sub-systems on the sensor module, enables data acquisition, performs the complete radar signal processing including sampling and FFT, and sends the results via its universal asynchronous receiver transmitter (UART) or universal serial bus (USB) interface to an external device. Four 12-bit versatile ADCs (VADC) in the MCU implement the radar signal sampling on the IF signals from the PGAs and also acquire various sensor data from the MMIC. With a 2048-kB flash and 352-kB RAM, the MCU enables highly accurate range-Doppler processing. Several general purpose input/output (GPIO) pins from the MCU are available on the header pins of the sensor module to interface with external circuits.

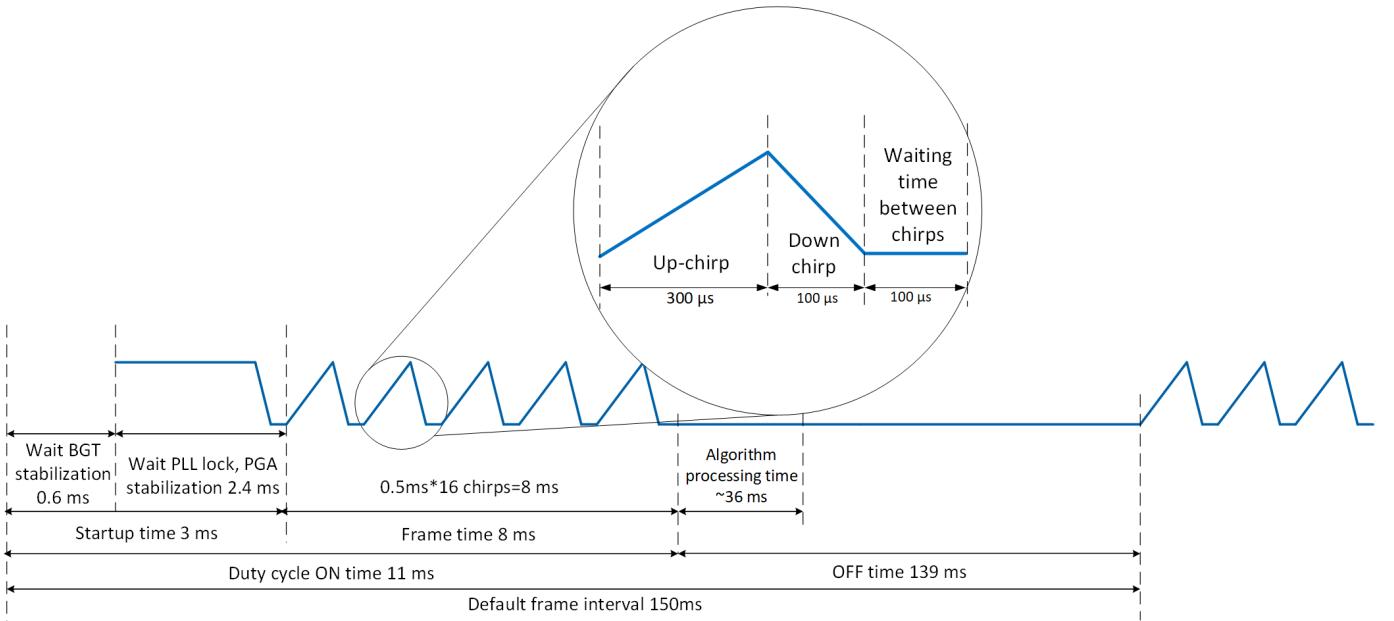


Fig. 6. Concept of the FMCW frame format.

TABLE I
FRAME CONFIGURATION PARAMETERS.

Number of chirps	16
Pulse repetition time	500 μ s
Up-chirp time	300 μ s
Samples per chirp	128
Range FFT size	256
Doppler FFT size	32

B. Firmware Concept

The radar firmware is based on a frame concept, whose configuration is illustrated in Fig. 6 and whose parameters are listed in Tab. I. Each frame has a duration of 150 ms and starts with enabling the PLL for 3 ms, followed by $N_C = 16$ identical frequency chirps with a repetition time of 500 μ s. After this 11-ms on-time, the BGT and PLL are disabled for power saving. The single chirps consist of an up-chirp of 300 μ s, a down-chirp of 100 μ s, and a waiting time of 100 μ s. Although only the up-chirps are sampled for signal processing, a down-chirp instead of a frequency jump is used to prevent illegal spurs in the adjacent frequency bands. The waiting time before a new up-chirp starts ensures a correct settling of the start frequency. All frequency changes are automatically controlled by the pre-set PLL and all frequency ramps are in the range of 24.025 ... 24.225 GHz. The closest 25-MHz bands at the ISM band limits are not used to avoid regulatory violations in the adjacent frequency bands. The used bandwidth $B = 200$ MHz results in a theoretical range resolution $\Delta R = \frac{c}{2B} \approx 60$ cm, whereas the real resolution is typically degraded due to applied windowing. During the PLL off-time, the algorithm is processed on the microcontroller within the first approximately 36 ms and the remaining time serves for host communication.

A flow chart of the complete algorithm is shown in Fig. 7. Correct raw data acquisition as the first step is ensured by

setting the start time and sampling frequency of the ADC so that the required $N_S = 128$ samples are equally distributed within ramp start and end. For each frame, a three-dimensional data cube $\Phi \in \mathcal{C}^{N_S \times N_C \times N_{RX}}$ containing the complex-valued baseband signals is obtained. The first dimension contains all samples per chirp (fast-time) for range estimation, the second dimension belongs to the different chirps per frame (slow-time) for velocity estimation, and the third dimension corresponds to the $N_{RX} = 2$ RX antennas for AoA estimation.

In the second step, a fast-Fourier transform (FFT) is applied to the fast-time dimension with a zero padding of factor two. The resulting FFT is called range FFT and has a size of 256 values. Subsequently, coherent pulse integration along the slow-time dimension aims to improve the signal-to-noise ratio and moving target indication (MTI) filtering is used for clutter suppression. The loop expresses the mutual influences over consecutive frames. Both signal processing routines as well as all other steps mentioned in the following are explained in detail in Sec. IV. In the next step, up to five human targets are detected in the range dimension by peak detection utilizing an adaptive threshold. The targets' velocities are estimated by the so-called Doppler FFT along the slow-time dimension for the corresponding range bins. Zero padding with a factor of two is also applied to this FFT resulting in a size of 32 values. Computing the Doppler FFT solely for the target indices reduces the computational complexity with regard to the implementation on the microcontroller. This range-Doppler processing is done for both antennas, for which reason the corresponding peak pairs are used for subsequent AoA estimation to determine the angle of each target to the radar sensor. In the last step of each frame, the calculated target parameters are fed to the tracking algorithm, which takes care of track association, track management, and track filtering. Due to its low complexity, the chosen tracking algorithm is based on alpha-beta filtering, which is explained in Sec. V.

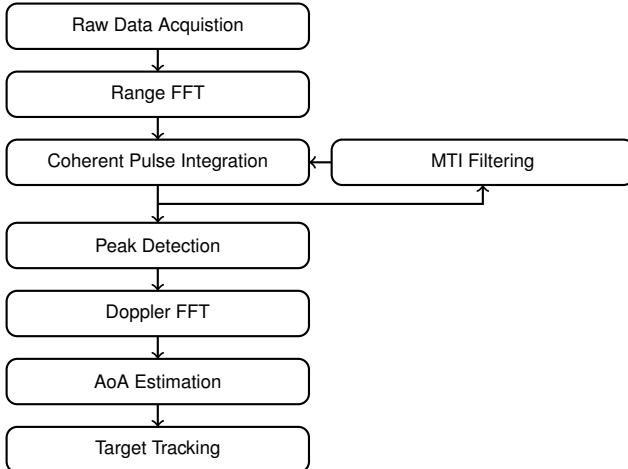


Fig. 7. Concept of the tracking algorithm.

IV. HUMAN DETECTION

This section investigates three signal processing routines for human target detection to be implemented on the microcontroller. While moving target indication filtering and chirp data integration are well-known concepts from literature, Doppler-compensated angle-of-arrival estimation is a newly proposed algorithm for enhanced angle estimation of quasi-static human targets.

A. Moving Target Indication Filtering

TX-to-RX leakage is a common challenge in radar systems and is also present in the *Infineon BGT24MTR12* radar chip. Due to the bandpass filtering in the baseband, this leakage leads to a low-frequency component, which results in high amplitudes in the first range FFT bins. Next to extensive hardware-related cancelation methods [45], [46], there are also software-based approaches to remove this leakage. One possible software solution is to calibrate the system before usage measuring the system itself by placing an absorber in front of the antennas to generate a reflection-free environment. At runtime these calibration values are subtracted from the measurement data. On the one hand, this method simultaneously extinguishes other impairments of the radar hardware, on the other hand, the calibration values are highly temperature dependent. Another software solution is MTI filtering [47], which is a type of an in-situ calibration and therefore temperature independent.

MTI filters in principle are low-order, simple finite impulse response (FIR) designs. At each time stamp the absolute maximum value over slow time of each range bin is denoted by $r_{i,\max}$. The MTI filter value t_i is the weighted average of this maximum value and the previous MTI filter value t_{i-1} with a weight of α :

$$t_i = \alpha \cdot r_{i,\max} + (1 - \alpha) \cdot t_{i-1}. \quad (1)$$

In the first time stamp t_0 is initialized with Zero. For each range bin, the previous MTI filter value is subtracted from $r_{i,\max}$ to obtain the filtered range FFT value $r_{i,\text{filt}}$:

$$r_{i,\text{filt}} = \text{abs}(r_{i,\max} - t_{i-1}). \quad (2)$$

This filtered value is then utilized for the subsequent target detection. MTI processing performs a linear filtering which leads to a diminished signal strength for static targets while maintaining the signal strength of the moving targets. Therefore, MTI processing helps to remove completely static targets, while non-static targets like humans are retained.

B. Chirp Data Combination

After computing the range FFT for each chirp and applying the MTI filter it is required to combine the data over all chirps of a frame for subsequent target detection. One possible data combination strategy is coherent integration [47], which combines the phase and magnitude of the range FFT data coherently over all chirps of a frame. Coherent integration is based on the mean R_{mean} of the range FFT data r_i over N_C chirps:

$$R_{\text{mean}} = \frac{1}{N_C} \cdot \sum_{i=1}^{N_C} r_i \quad (3)$$

$$r_i = a_i \cdot e^{j\phi_i} \quad (4)$$

Here, amplitude and phase of the range FFT value for chirp i are denoted by a_i and ϕ_i , respectively. This method enhances the signal-to-noise ratio (SNR), but it has to be considered that all phase values of all chirps have to be aligned, since phase misalignment leads to signal distortion. Another approach is to find and use the chirp i with the maximum absolute range FFT value R_{\max} over all N_C chirps:

$$R_{\max} = \max_{i=1}^{N_C} (y_i) \quad (5)$$

This approach assures that the FFT data with highest amplitudes is chosen for target detection and prevents phase misalignments. Both approaches are evaluated in Sec. VII, but only the variant using the maximum absolute value is implemented in firmware due to the enhanced signal quality for target detection and the necessity of phase alignment.

C. Doppler-Compensated Angle-Of-Arrival

In case of human targets, the individually moving body parts like torso or shoulders induce different velocities and (wrong) angles in the Doppler FFT. To overcome this issue, the proposed procedure uses the centroid of the Doppler spread along the slow-time for angle estimation. Furthermore, the vital signals of static and quasi-static humans modulates the data along slow-time, which results in high variations in the angle estimation. In the presented approach these modulations are compensated to stabilize the angle estimations.

The basic approach to determine the AoA is to evaluate the phase difference between both RX antenna beams, also referred to as phase-comparison monopulse. The AoA θ is calculated by geometrical considerations based on an incoming plane wavefront with

$$\theta = \arcsin\left(\frac{\lambda}{2\pi} \cdot \frac{\Delta\phi}{d}\right), \quad (6)$$

where d is the geometric distance between two antennas, $\Delta\phi$ is the observed phase difference between both RX signals, and

λ represents their wavelength. The received signal at antenna m is given as

$$x_m(t) = s(t)e^{j2\pi\tau(m-1)} + z_m(t), \quad (7)$$

where $s(t)$ is the source signal, $\tau = \frac{d\cos(\theta)}{\lambda}$ which is the phase shift per antenna and $z_m(t)$ is the interference-noise term. In matrix-vector form this is represented as

$$\begin{bmatrix} x_1(t)^T \\ x_2(t)^T \end{bmatrix} = \underline{a}(\theta) [s(t)], \quad (8)$$

where $\underline{a}(\theta)$ is the steering vector $[1 \ e^{j2\pi\tau(\theta)}]^T$ and $s(t)$ is the range-Doppler signal peak extracted from the fast-time and slow-time processed signal. However, in case of a human, $s(t)$ is not constant due to breathing, but fluctuating as $Ae^{j2\pi f_v t}$. Here, A is the representation of the human radar cross section (RCS) and f_v is the vital-Doppler frequency due to breathing and heartbeat. In compact notation, this signal is represented as $X(t) = \underline{a}(\theta)s(t) + Z$, where $X(t) \in \mathcal{C}^{2 \times N_V}$, $\underline{a}(\theta) \in \mathcal{C}^{2 \times 1}$, and $s(t) \in \mathcal{C}^{1 \times N_V}$, where Z represents noise and N_V is the number of data points considered to represent a full-cycle of breathing, typically three seconds.

To improve the calculated AoA, this paper therefore proposes following estimation:

$$XX^H = \underline{a}(\theta)s s^H \underline{a}(\theta)^H + ZZ^H \quad (9)$$

$$XX^H = N_v |A|^2 \begin{bmatrix} 1 & e^{j2\pi\tau(\theta)} \\ e^{-j2\pi\tau(\theta)} & 1 \end{bmatrix} \quad (10)$$

Here, the AoA θ is estimated by considering the off-diagonal elements in the matrix $XX^H \in \mathcal{C}^{2 \times 2}$. This approach increases the SNR of the received signal for estimating the AoA and additionally eliminates fluctuations due to breathing, which result in high variance in the estimated angle of the human target.

V. HUMAN TRACKING

Target detection provides an estimate of target's location in the radar's field of view. Target tracking maintains the state and identity of the detected target over time accounting for detection errors (false negatives, false alarms), target occlusions, measurement noise, ghost targets, interference sources, and the presence of other objects.

A. Track Management

Track management consists of several parts, which are described in the following.

1) *Measurement to Track Association*: The task of the measurement to track association is to assign the current measurement to an existing track. This step is also called gating in literature. In the proposed approach, deterministic sub-optimal nearest neighbor (SNN) is used, whereby each measurement is assigned to a track based on the closest Euclidean distance to that track's state.

2) *Track Initiation*: Track initiation is the process of creating a new radar track from an unassociated radar measurement. During track initialization, when a new unassociated radar measurement is encountered, the measurement is assigned to a new track, but the status of the track is kept tentative. The status of these tracks are changed to confirmed only after $N = 5$ out of $M = 8$ subsequent measurements have a positive measurement of the target. This reduces false positives from ghost targets entering into the tracks.

3) *Track Maintenance*: Track maintenance is the process, in which a decision is made if a track has to be terminated. If a track was not associated with a measurement during the association phase for consecutive five frames, there is a high likelihood that the target no longer exists in the radar's field of view.

B. Track Filtering

An alpha-beta filter is used for human target tracking due to its simplicity. It does not require a detailed system model and its low computational complexity ensures code realization on a microcontroller. The alpha-beta filter predicts position and velocity assuming zero acceleration of a moving target. It iterates the prediction as well as updates and smooths processes. The prediction process is expressed as follows:

$$x(k) = \bar{x}(k-1) + \delta T v(k-1), \quad (11)$$

$$v(k) = \bar{v}(k-1), \quad (12)$$

where δT is the measurement update interval, $x(k)$ and $\bar{x}(k)$ are the predicted and smoothed target position at time $k\delta T$, respectively, and $v(k)$ and $\bar{v}(k)$ are the predicted and smoothed target velocities at time $k\delta T$, respectively. The parameter $k = 1$ denotes the time instance, when a track is initialized and the counter is incremented henceforth until the track is cleared.

The update and smoothing process is defined as

$$\bar{x}(k) = x(k) + \alpha(\hat{x}(k) - x(k)), \quad (13)$$

$$\bar{v}(k) = v(k) + \beta_1(\hat{v}(k) - v(k)) + \beta_2/\delta T(\hat{x}(k) - x(k)), \quad (14)$$

where $\hat{x}(k)$ and $\hat{v}(k)$ are the measured position and velocity of the target at $k\delta T$ refresh time, respectively. Equation (14) has a deviation from the conventional alpha-beta filter, since the contribution from the measured target's velocity $\hat{v}(k)$ is added. The values of the correction gains α , β_1 , and β_2 are chosen empirically, however, $\beta_1 + \beta_2 \leq 1$ is ensured. Additionally, the value of α is chosen adaptively as a function of target state's velocity. If the velocity increases, the α value is also adaptively increased.

A further modification to common alpha-beta filtering is the introduction of the target's angle value. The target's angle value is fed into a median filter followed by an exponentially weighted moving average filter given as

$$\Theta(k) = \gamma\Theta(k-1) + (1-\gamma)\theta(k), \quad (15)$$

$$\Theta(k) = \frac{\Theta(k)}{1-\gamma^k}, \quad (16)$$

where $\Theta(k)$ is the exponentially averaged angle and $\theta(k)$ is the measured angle at time instant $k\delta T$ after median filtering.

The parameter γ denotes the exponential weighted coefficient, whose value is chosen empirically in the range $0 < \gamma < 1$. The factor $\frac{1}{1-\gamma^k}$ is the bias correction term used for correcting initial bias, when the filter starts. The proposed alpha-beta filter with exponentially weighted angle average filter is different than the conventional alpha-beta-gamma filter, which assumes acceleration as the third state variable.

VI. HUMAN CLASSIFICATION

Any movements from quasi-static targets like trees in outdoor scenarios and multi-path reflections on strong-reflecting targets like walls can induce a false human target detection. To mitigate such false alarms, the usage of a one-class SVM classifier with hand-crafted features extracted from the detected target is proposed. The advantage of a one-class classifier in comparison to a two- or multi-class classifier is the independence from all non-human targets, since only human data are used for training. The SVM is a convenient classifier for binary problems, which uses a soft margin to enable an optimized separation of overlapping classes. In the addressed human classification as one-class problem, this margin implicates two positive aspects. First, an overfitting regarding the human subjects used for training is prevented and hereby, the detection rate for other human subjects is higher. Second, non-human targets are easier detected as outliers.

In the proposed human classification, ten handcrafted features are extracted from the measurement data. All features of a detected target are based on its complex micro-Doppler data at the corresponding range bin. In a pre-processing step, the IQ values of all 16 chirps j per frame and along N consecutive frames i are stored in a matrix W with dimensions $i \times j$, which is updated every frame like a sliding window. The higher the number of utilized frames N , the larger the delay, but also the higher the accuracy. For every matrix of this sliding window, the extracted features k are stored in the feature vector F . Four features utilize averaging (mean) and the standard deviation (std) in both dimensions to exploit the micro-motions of quasi-static targets in comparison to completely static targets:

- $F_1 = \text{mean}(\text{std}(W, 2))$
- $F_2 = \text{std}(\text{mean}(W, 2))$
- $F_3 = \text{std}(\text{std}(W))$
- $F_4 = \text{std}(\text{std}(W, 2))$

In these equations, the second argument of a function denotes the dimension of the operation, no argument denotes operation on the first dimension. The other six features require further pre-processing as each three of them utilize breathing and heartbeat information, which are included in the micro-Doppler signal. Therefore, the argument (arg) of the first chirp has to be extracted along all N frames, since the relative micro-distance d_i within one range bin is directly proportional to the unwrapped phase information:

$$d_i = \text{unwrap}(\text{angle}(W_{i,1})). \quad (17)$$

The vector d equals the displacements of the target along all N frames of the sliding window. Since all micro-motions of the target are overlapped in this signal, the vital signals due to breathing d_{br} and heartbeat d_{hb} have to be extracted

by filtering. Here, forth-order Butterworth filters are used with passbands of 0.1 ... 0.5 Hz and 0.75 ... 3.0 Hz, respectively. Subsequently, a common peak search is applied on both filtered signals. The heights of the N_{br} detected peaks in the breathing signal are stored in the vector p_{br} , the heights of the N_{hb} detected peaks in the heartbeat signal are stored in the vector p_{hb} . The number of peaks and the sums of the heights denote four other features:

- $F_5 = N_{br}$
- $F_6 = N_{hb}$
- $F_7 = \text{sum}(p_{br})$
- $F_8 = \text{sum}(p_{hb})$

The last two features represent the peak-to-peak amplitude of the filtered signals:

- $F_9 = \max(d_{br}) - \min(d_{br})$
- $F_{10} = \max(d_{hb}) - \min(d_{hb})$

All ten features are scaled to be within a range of 0 ... 1. To ensure an appropriate quality of the vital signals, the proposed length of the sliding window N is 56, which implies a time window of 8.4 s for a frame time of 150 ms. Another sliding window with length L subsequently applied on the classification results further enhances the final classification result, which is defined by the majority within the sliding window. Using a window length $L = 11$ implies an additional delay of ten frames and a total delay 66 frames (9.9 s).

The open-source library *LIBSVM* is chosen as SVM implementation, as it is based on a C interface and additionally provides a *MATLAB* interface for preliminary investigations [48]. Except for one parameter, all default settings are used to create the one-class SVM model, which implies a kernel with a radial basis function and a degree of three. Since γ as one of the two most important optimization parameters by default is one divided by the number of features, a value of 0.1 is used. The second important optimization parameter is ν , which defines the outlier fraction. With a default value of 0.5 optimizes the SVM by classifying each half of the training data as in-class and outlier. This would imply an automatic misclassification of at least 50% for human targets, which is why a smaller value has to be chosen for ν . On the contrary, a too small value results in too many false human detection, for which reason a value of 0.2 is proposed. The unpreventable misclassifications for human and non-human targets can subsequently be decreased by the proposed sliding windowing of the classification results.

VII. RESULTS

This section presents the results of the measurements to evaluate the performance of the radar system and to verify the implemented algorithms for detection and tracking of up to five humans.

A. System Performance Evaluation

The performance parameters of the created radar system are evaluated in this sub-section.

1) Measurement Scenarios: Figure 8 gives an overview of the measurement scenarios for system performance evaluation. A block diagram of the accuracy measurements regarding distance and angle are depicted in Fig. 8a and Fig. 8b, respectively. The measurement setup for both tests is an open space outdoor environment with a wall at a distance of approximately 20 m to the radar, as shown in Fig. 8c. The radar sensor is mounted on a tripod at a height of 1.5 m facing the open space. For distance measurements, a corner reflector with an RCS of 1 m^2 is placed at different distances d from the radar sensor with an angle of 0° . The step size is 0.1 m for distances in the range of 0.5 ... 5.5 m and 0.5 m for larger distances up to 18 m. The distances between radar and target were endured by a laser-based triangulation sensor. Angle measurements were done with the same setup, but the corner reflector positioned fixed at a distance of $d = 3.0 \text{ m}$. The angle between sensor and target is varied from -65° to $+65^\circ$ with a step size of 5.0° by turning the sensor on the tripod. The markers on the tripod serve as reference values. Larger angles than $\pm 65^\circ$ are not measured since the received signal strength is mostly below the threshold, which implies an unstable target detection.

The range resolution was tested using two corner reflectors each with a radar cross section of 1 m^2 , as shown in Fig. 9. One corner reflector (CR1) was positioned at a distance of 2.0 m and an angle of -2° . The second reflector (CR2) was moved radially at an angle of $+5^\circ$ starting at 2.5 m with increasing distance until the system detected two targets.

2) Measurement Results: The measurement results for distance and angle accuracy tests are depicted in Fig. 10. Figure 10a shows the delta between measured and actual distance values along the actual distances. These value correspond to the systematic distance measurement error, whose absolute value is below 20 cm for distances up to 17 m and maximum 25 cm along the complete measurement range. The angle accuracy is shown in Fig. 10b depicting the delta between measured and actual angle values. The delta is higher for larger angles with a maximum value of 10° at an actual angle of 60° . In the range of $-25 \dots +55^\circ$ the delta is equal or less than 4° . The result of the range resolution test is 0.9 m.

B. Human Detection

For human detection, the measurement scenario was similar to the one for system performance evaluation in Fig. 8, but without the corner reflector.

1) Moving Target Indication Filtering: MTI filtering is applied only on the first range bins, which correspond to the closest 5 m, in order to remove the high amplitude low frequency components due to radar hardware impairments. Larger ranges are not filtered, since it occasionally also extinguishes static humans due to the decreasing received signal strength of human targets for increasing distances. The impacts of the TX leakage and the MTI filter are investigated by measurements in a small antenna chamber fully covered with RF absorbers. The results are illustrated in Fig. 11 for both disabled and enabled MTI filtering for the measurement values of RX1. The measurement values of RX2 show equal results, for which reason these plots are omitted. In both sub-figures,

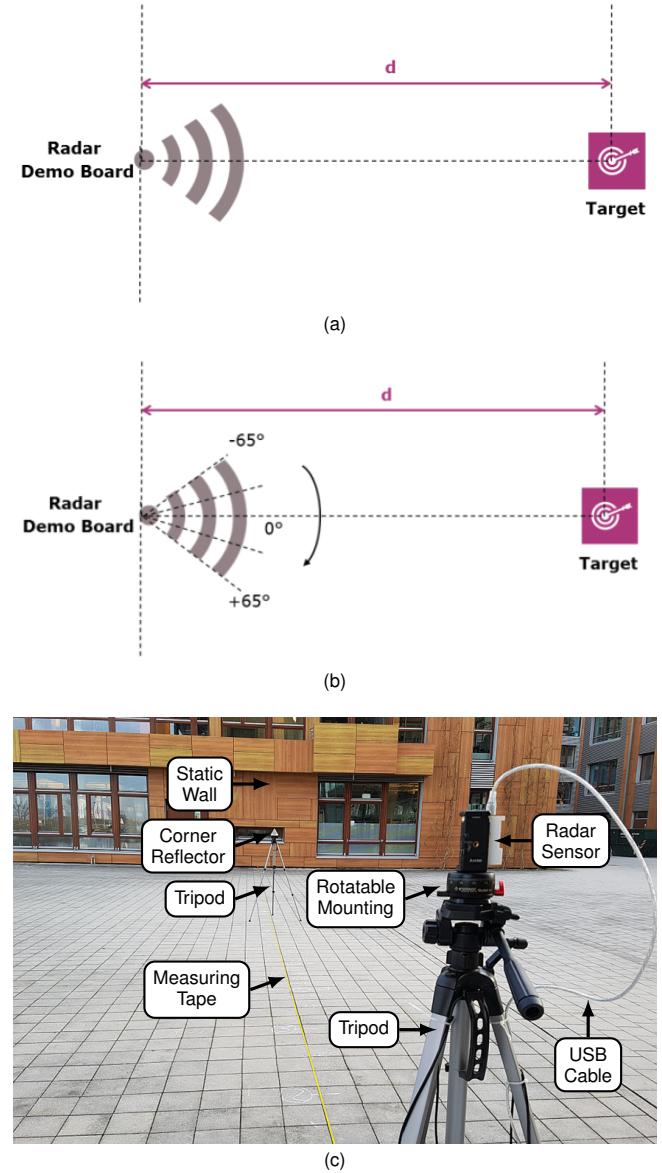


Fig. 8. (a) Block diagram for distance and (b) angle accuracy measurements. (c) Photograph of the measurement scenario for both accuracy tests.

the maximum absolute range FFT values over all chirps per frame are shown for the first 25 frames after start-up. With MTI filtering disabled in the Fig. 11a, high amplitudes are observed in the first range bins. Applying MTI filtering with a filter value of $\alpha = 0.2$ rapidly diminishes this effect until almost complete removal in frame 17, as shown in Fig. 11b.

2) Chirp Data Combination: The impacts of coherent integration and usage of maximum absolute value are compared in Fig. 12. In the presented measurement scenario, a person is walking from a distance of 1.0 m to 15 m with respect to the radar at an angle of 0° . Both sub-figures show the amplitudes of the resulting range FFT data utilized for subsequent target detection over all measurement frames. Especially for larger ranges in the later frames, the signal quality for maximum absolute value in the lower sub-plot is way higher than for coherent integration in the upper sub-plot.

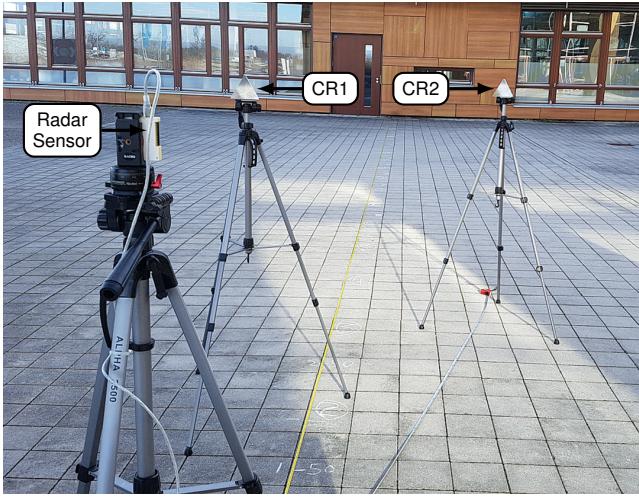


Fig. 9. Photograph of the measurement scenario for range resolution test.

3) *Doppler-Compensated Angle-Of-Arrival:* The improvement of the Doppler-compensated AoA in comparison to the non-compensated measurement result is shown in Fig. 13. In this experiment, one person was breathing normally while standing still at a distance of 3 m and an angle of 5° with respect to the radar sensor. Without compensation, there are high angle variations over time with a false mean value of approximately 8° . The Doppler-compensation algorithm detects the wrong estimation after 20 frames and correctly calculates the actual angle.

C. Human Tracking

The measurement setup for all human tracking tests is this sub-section is same open space outdoor environment with a wall at approximately 20 m as for the system performance evaluation.

1) *Track Management:* The functionality of the track management is investigated for a scenario, in which a single human walks tangentially from the left to the right side and back

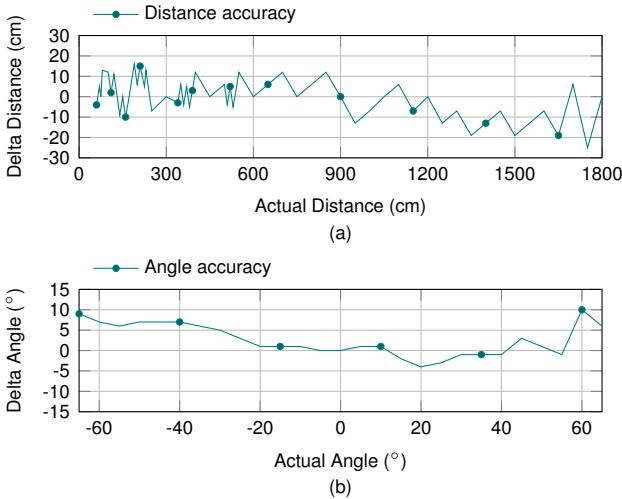


Fig. 10. Measurement results of (b) distance and (c) angle accuracy tests.

at a radius of 5.0 m with respect to the sensor. Figure 14 presents the intermediate target values in terms of range FFT, Doppler FFT amplitude, Doppler FFT value, and angle over all frames obtained by range-Doppler processing and AoA estimation. In the topmost sub-plot, it can be seen that the range FFT strength of the wall is mostly much higher than of the human target. Two exceptions are the time instances around frame 60 and 175, in which the human is directly between sensor and wall. The FFT thresholds for target detection are indicated by green lines. The Doppler FFT amplitude threshold prevents the wall to be assigned as target despite of the high fluctuations regarding its velocity, which can be observed in the two corresponding Doppler FFT sub-plots. The phenomenon of high velocity fluctuations for a static target is explained by the mean removal in Doppler FFT processing. The IQ representation of a static target over time, interferences like shading effects or multi-path reflections excluded, is a tiny point cloud with certain amplitude. Removing its mean results in noise around the origin and applying the Doppler FFT on these processed data results in large velocity fluctuations but with a very small amplitude.

The intermediate values are subsequently used by the tracking algorithm for target assignment, track update and track clearance. The tracking results in terms of the three target properties range, velocity, and angle are shown in Fig. 15. The human target is tracked correctly, while the wall is not deduced as target. Track clearance can be observed between both tangentially walks as the human target intermittently left the sensor's field of view. Due to median filtering with a length of five, the target properties are smoother compared to the intermediate values.

2) *Track Filtering:* The functionality of the track management is investigated for a scenario, in which two people walk in front of the radar within a distance of maximum 15 m and an angle of $\pm 50^\circ$ with respect to the sensor. Figure 16 shows

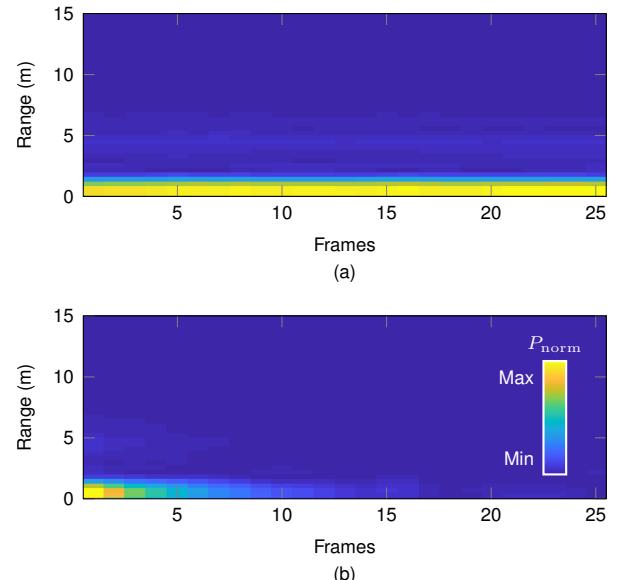


Fig. 11. (a) Remaining TX leakage in the first range bins and (b) vanishing TX leakage by moving target indicator filtering.

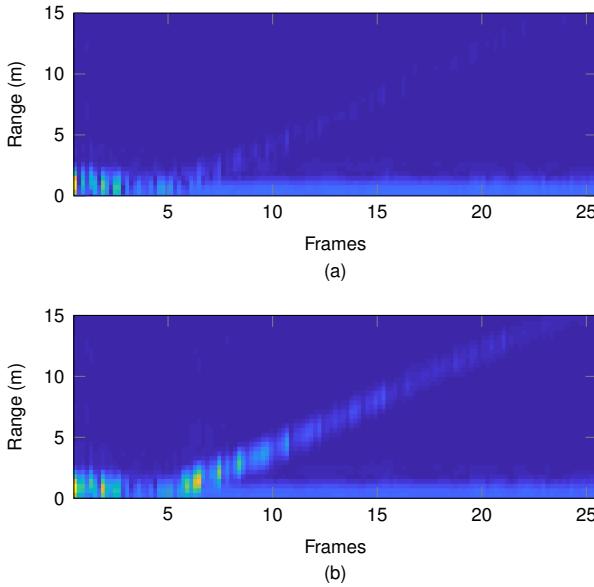


Fig. 12. (a) Mean and (b) maximum absolute range FFT value over all chirps per frame are shown for all measured frames. In the presented scenario, a human walked radially away from sensor.

the measured range FFT amplitudes for both antennas RX1 and RX2, which are highly similar.

The results of the applied tracking algorithm are shown in Fig. 17. It can be observed that infrequently one of the two target disappears. Comparing the sub-plots for range and angle reveals that in these situations the disappearing target has a high angle and is therefore outside the sensor's field of view. The coherency between range and velocity substantiates the functionality of the track management. For decreasing target range, which represents an approaching person, the target velocity is deduced as constant and positive. For a departing target, its velocity is deduced as constant and negative. Around frame 650, a third target is false detected, a so-called ghost target. At this specific scenario, the closer target (Track 2) is solely around one meter to the sensor, for which reason multiple reflections induce a ghost target at approximately 4 m. In general, ghost targets sporadically appear, when noise in both range FFT as well as Doppler FFT are above the thresholds. Since this happens only for few consecutive frames, ghost targets are automatically cleared by a counter in the track management after same frames. Another limitation of the track filtering can be observed near frames 600 and 700, when the

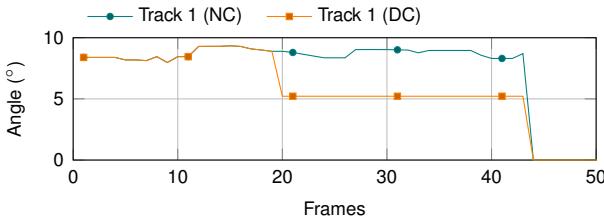


Fig. 13. Estimated target angle for non-compensated (NC) as well as Doppler-compensated (DC) AoA.

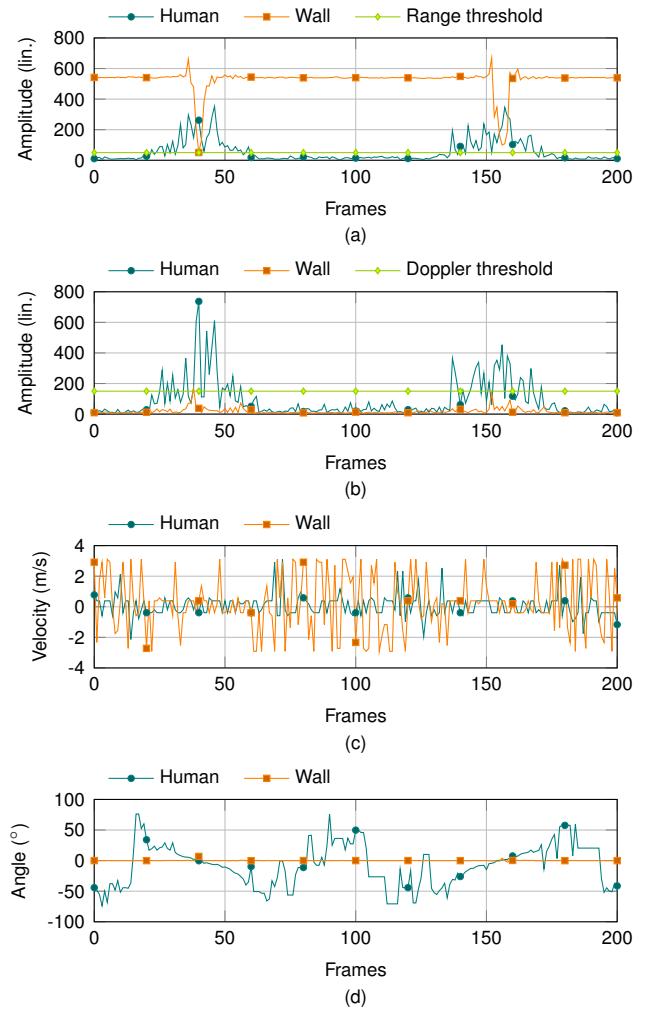


Fig. 14. (a) Linear Range FFT amplitude, (b) linear Doppler FFT amplitude, (c) corresponding Doppler FFT value, and (d) angle as intermediate measurement values for a tangentially moving human with a wall as static object at 20 m.

two targets cross their paths and appear in the same range bins. Instead of correctly continuing each track, both tracks are exchanged due to the range-based gating in the applied algorithm.

Figure 18 shows the tracking results for five human targets, each walking from behind the sensor to the wall and back with a distance of approximately 4 m among each other. In this measurement scenario, the sensor was placed at height of 2.5 m to prevent shadowing effects. All targets are correctly tracked in a range of 2.5 ... 20 m. At around 20 m, the range FFT signal strengths for the human tracks are too low and only the wall is detected as static target. Therefore, all tracks are cleared after a couple of frames and re-assigned shortly after approaching the sensor again.

D. Human Classification

To investigate the proposed human classification, measurement data were gathered for different outdoor scenarios:

- 1) One person at 4 m (for three different persons)
- 2) One person at 6 m (for three different persons)

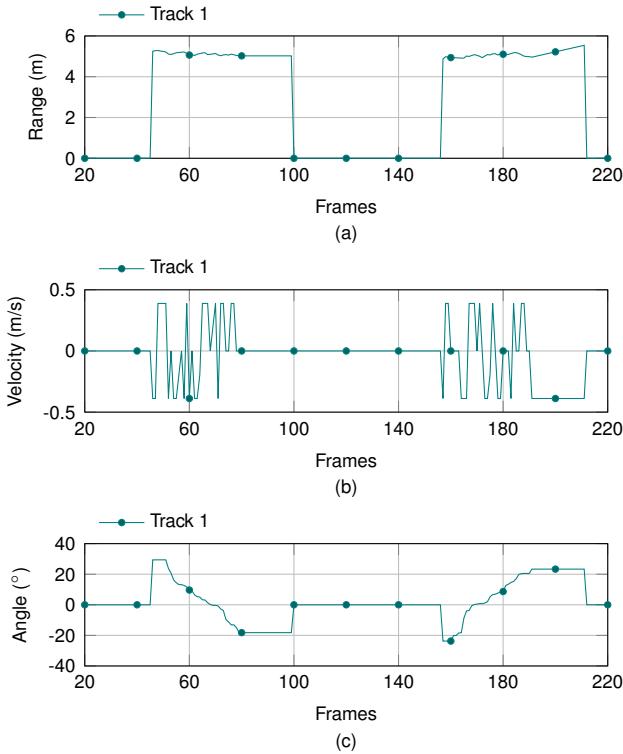


Fig. 15. Tracking results (range, velocity, and angle) for a tangentially moving human with a wall as static object at 20 m.

- 3) One person at 4 m and one person at 6 m (for six different person constellations)
- 4) A wall at approximately 21 m
- 5) A tree at around 4.5 ... 9.0 m in a almost non-windy environment (four different trees)
- 6) A tree at around 4.5 ... 9.0 m in moderate wind (five different trees)

For each scenario the detected target peaks in the range FFT were utilized to investigate the proposed classification. In the third scenario, both persons were assigned as a human-target and investigated. In the first three scenarios the wall was additionally assigned and labeled as a non-human target to investigate shadowing effects by human subjects. This results in 40 different data sets each with a length of around 20 s.

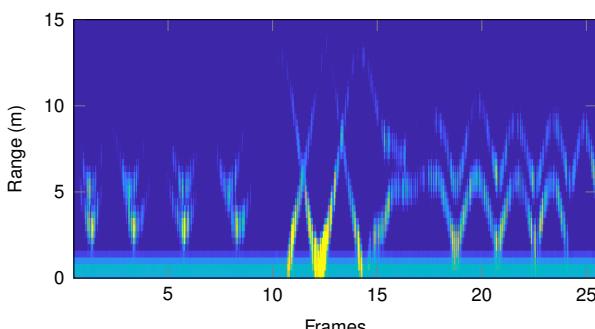


Fig. 16. Range FFT amplitudes over several frames (slow time) for two moving persons.

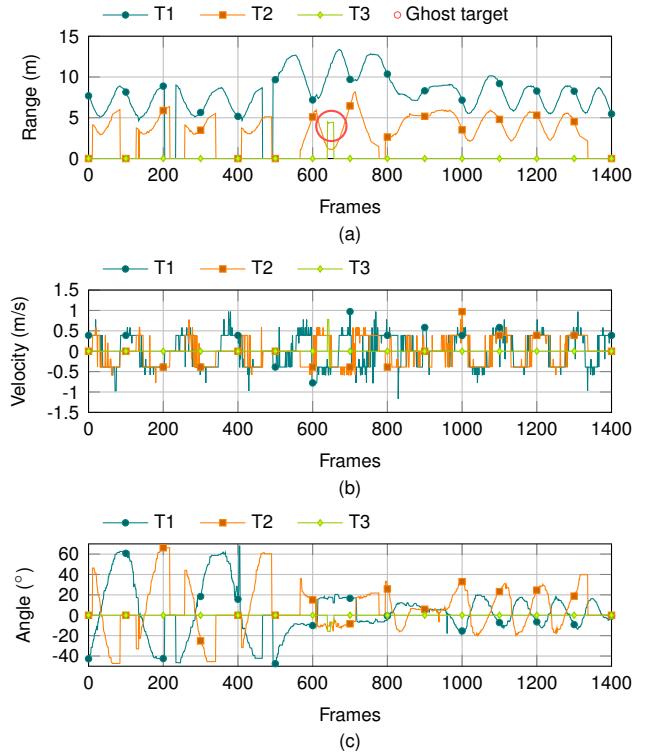


Fig. 17. (a) Range, (b) velocity, and (c) angle as tracking results for two moving persons.

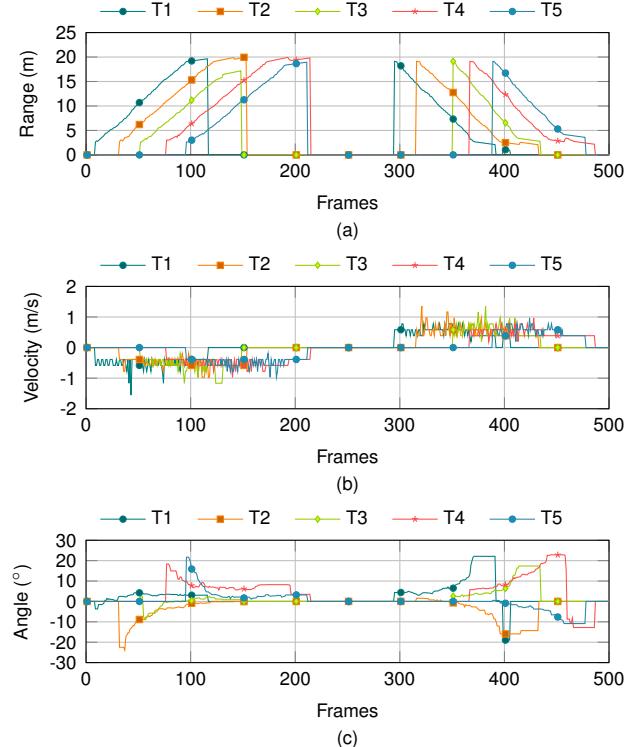


Fig. 18. (a) Range, (b) velocity, and (c) angle as tracking results for five moving persons.

The IQ data for all chirps of all frames are illustrated for four exemplary data sets in Fig. 19, in which each colorful

line or point cloud represents the micro-Doppler movement along all chirps of a single frame. The shadowed wall in subplot 19a clearly shows a different behavior, but the IQ signatures for a static human in subplot 19b, a tree in an almost non-windy environment in subplot 19c, and a tree in moderate wind in subplot 19d are rather similar. While the IQ signatures of the trees look rather random, a more uniformly and concentric signature can be observed for the human subject. This corresponds to the vital signals of a human target, which are utilized by some of the proposed features. It has to be emphasized that these are only four exemplary data sets out of forty and that there are high variations in the IQ signatures within all four presented measurements rows.

Applying a sliding window with a length of 56 frames as proposed in Sec. VI yields 2473 windows of human data, 2393 windows of tree data, and 1780 windows of wall data. For each of these windows the ten proposed features are extracted. The windows of human data were randomly split with a common ratio of 80% for training and 20% for test. The one-class SVM model was trained with the parameters proposed in Sec. VI, thus, default values for all parameters except ν with a value of 0.2. For 100 random slits of the human data, the proposed classifier achieves the following accuracies:

- Human (Training): $80.0\% \pm 0.9\%$
- Human (Test): $80.1\% \pm 9.7\%$
- Tree: $81.7\% \pm 14.8\%$
- Wall: $100.0\% \pm 0.0\%$

The accuracies of human targets correspond with a value of around 80% to the preset outlier fraction ν . Detecting trees as outliers has a similar accuracy of around 82%. Applying a subsequent slighting windowing on the classification results further enhances these accuracies. An accuracy of 100% for walls implies that the proposed classifier can clearly distinguish shadowing or multi-path effects of human subjects from real quasi-static human targets. It has to be mentioned that using only 50% or 20% of the human data for training results in similar accuracies. Furthermore, additional investigations were done with applying a window length of $N = 5$ frames and using only the first four features $f_{1..4}$, which do not evaluate vital sign information. Here, γ is adjusted to a value of 0.25. While the accuracies for human and wall test data remained almost constant with $80.0\% \pm 11.8\%$ and $98.4\% \pm 1.4\%$, respectively, the accuracy for trees decreased to $17.2\% \pm 14.4\%$, which implies that humans and trees cannot be distinguished by the classifier at all. This emphasizes the necessity of advanced features like such based on breathing and heartbeat signals as proposed in this paper.

VIII. DISCUSSION

This paper aimed for a linear FMCW radar system, which can detect and track up to five humans, while all signal processing routines run on an embedded microcontroller. Regarding the measurement results in Sec. VII, this aim has been achieved. Within an angular range of 100° and a distance up to 20 m, five radially persons were correctly tracked. The successful tracking of two persons moving freely in the valid area emphasizes the performances of the algorithm. Track

exchanges for crossing targets due to the simple alpha-beta filter is a limit of the presented system, which is also shown in this experiment. Sporadic wrong angle estimations of static humans due to human micro-motions and ghost targets due to multi-path reflections or non-human quasi-static objects like trees are challenges of alternative systems, which are addressed in this paper by a novel Doppler-compensated AoA estimation algorithm and a novel human classification method based on a one-class SVM. Other challenges like TX-to-RX leakage, chirp data combination, and signal fluctuations due to breathing are solved by a smart combination of known algorithm. All signal processing routines run on a microcontroller to provide tracking results in real-time.

In comparison to alternative radar systems for human tracking, the proposed approach offers various improvements. One major advantage is the easy portability due to its small size and low power consumption. While many approaches are based on multistatic radar systems [23], [24], [26], [27], [25], the proposed system is a single sensor with a size of 45×55 mm. It is solely powered by USB and does not require an external voltage source, although the complete signal processing runs on the embedded microcontroller. This is facilitated by simple linear frequency modulation and alpha-beta track filtering instead of UWB pulses [8], [22], [25], [26], [27] and Kalman filtering [21], [8], respectively, which would considerably enlarge the system complexity. Despite of the lower complexity compared to other systems, the presented radar sensor successfully tracks up to five persons instead of only one or two persons [18], [20], [19], [22], [25], [27], [28], [32]. While many systems are only used for indoor tracking due to their short detection range [20], [26], [30], the presented system tracked targets up to 20 m in the verification measurements, which allows for usage also in outdoor environment.

Furthermore, this paper firstly introduces a one-class classifier to distinguish humans from other quasi-static targets like trees in outdoor scenarios or shadowing effects on walls without using non-human target information in the training stage. While other publications only detect if a person is present or not [36], specify several target classes for a multi-class classifier [37] or preselect specific targets for the non-human class and use them to train a two-class SVM [39], the presented one-class classifier is exclusively trained with human data. Shadowed walls as well as trees in differently windy scenarios depict typical non-human targets, which degrade the detection performance in practical outdoor measurements. With an accuracy of 100%, shadowing effects on walls can clearly be distinguished from real human targets. The achieved accuracy of around 82% for trees is above the predefined detectability of 80% for human targets and an adequate high value for a target with highly similar micro-movements. Further investigation on the feature set revealed that the usage of advanced features, which are for example based on vital sign information, are definitely required to successfully distinguish human targets from other quasi-static targets.

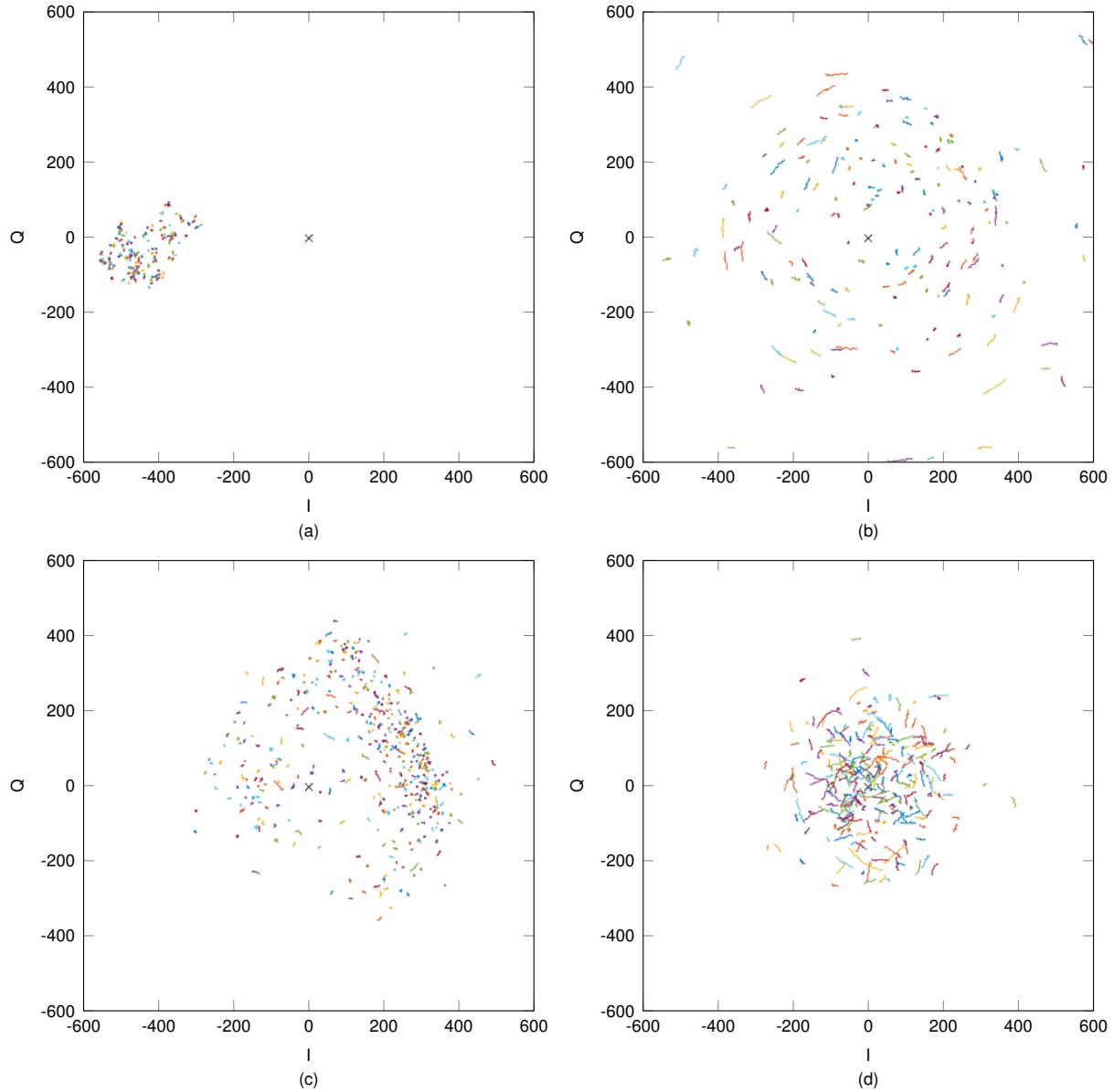


Fig. 19. Exemplary IQ data for (a) a wall behind a person at a distance of 21 m (b) a static human at a distance of 4 m, (c) a tree in an almost non-windy environment, and (d) a tree in moderate wind. Each line or point cloud illustrates the micro-Doppler movement within a single frame, the black cross in the mid of each subplot marks the complex origin.

IX. CONCLUSION

This paper presented an FMCW radar system for human target detection, tracking, and classification utilizing the 24-GHz ISM band. After investigating the state-of-the-art and the challenges of the contemplated application, a dedicated hardware-firmware system concept was designed. The researched algorithm consists of various signal processing routines, which cancel non-idealities in the hardware, improve the SNR, and compensate signal instabilities to ensure proper detection of human targets. The proposed signal processing procedure smartly combines known algorithms with two novel algorithms firstly introduced in this paper, an angle-of-arrival estimation method for human targets as well as a one-class SVM for human classification. The alpha-beta filter as tracking routine and all other chosen algorithms allow for

an implementation on a microcontroller by simultaneously implicating good performance. The created radar system along with the implemented algorithm showed convincing results in extensive verification measurements. Distance and angle accuracies, resolution, and maximum distance are competitive to comparable radar systems. It was shown that the proposed system correctly detects and tracks up to five moving persons in an angular range of over 100° and a distance up to 20 m. The proposed one-class classifier is the first approach to distinguish human targets from other quasi-static targets without utilizing any previous knowledge of non-human targets. The presented classifier achieves the preset outlier fraction for human targets, an accuracy of around 82% for trees in different windy and non-windy environments, and an accuracy of 100% for shadowing effects of a wall behind human subjects.

ACKNOWLEDGMENT

The authors would like to thank Andreas Dorfner and Thomas Finke for supervising the project and reviewing this paper. Furthermore, the authors would like to thank Alejandro El Hares, Lukas Rauh, and Michael Kießling for supporting the measurements, as well as Alexander Glas for supporting the hardware development.

AUTHOR CONTRIBUTIONS

CW and PV implemented and tested parts of the algorithm, verified the proposed systems in extensive measurements, and evaluated the results. CW wrote most part of the paper, while PV contributed parts to the sections challenges, human detection, and results. AC designed, created, and tested the hardware as well as contributed the hardware-related parts to the paper. AS designed the system concept, researched and implemented parts of the algorithm, and wrote the parts regarding firmware concept and track filtering.

REFERENCES

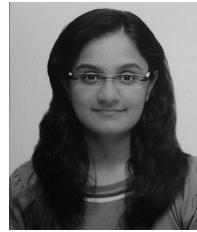
- [1] A. J. Lipton, H. Fujiyoshi, and R. S. Patil, "Moving target classification and tracking from real-time video," in *Proceedings Fourth IEEE Workshop on Applications of Computer Vision. WACV'98 (Cat. No.98EX201)*, Oct 1998, pp. 8–14.
- [2] D. Comaniciu and V. Ramesh, "Robust detection and tracking of human faces with an active camera," in *Proceedings Third IEEE International Workshop on Visual Surveillance*, July 2000, pp. 11–18.
- [3] M. Hussein, W. Abd-Almageed, Y. Ran, and L. Davis, "Real-time human detection, tracking, and verification in uncontrolled camera motion environments," in *Fourth IEEE International Conference on Computer Vision Systems (ICVS'06)*, Jan 2006, pp. 41–41.
- [4] J. Liu, Y. Liu, Y. Cui, and Y. Q. Chen, "Real-time human detection and tracking in complex environments using single RGBD camera," in *2013 IEEE International Conference on Image Processing*, Sept 2013, pp. 3088–3092.
- [5] Q. Cai and J. K. Aggarwal, "Tracking human motion in structured environments using a distributed-camera system," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 21, no. 11, pp. 1241–1247, Nov 1999.
- [6] A. Carullo and M. Parvis, "An ultrasonic sensor for distance measurement in automotive applications," *IEEE Sensors journal*, vol. 1, no. 2, p. 143, 2001.
- [7] C. G. Raghavendra, S. Akshay, P. Bharath, M. Santosh, and D. Vishwas, "Object tracking and detection for short range surveillance using 2D ultrasonic sensor array," in *2016 International Conference on Circuits, Controls, Communications and Computing (I4C)*, Oct 2016, pp. 1–4.
- [8] S. Chang, N. Mitsumoto, and J. W. Burdick, "An algorithm for UWB radar-based human detection," in *2009 IEEE Radar Conference*, May 2009, pp. 1–6.
- [9] J. Han and B. Bhanu, "Fusion of color and infrared video for moving human detection," *Pattern Recognition*, vol. 40, no. 6, pp. 1771–1784, 2007.
- [10] B. Yang, Y. Lei, and B. Yan, "Distributed multi-human location algorithm using naive Bayes classifier for a binary pyroelectric infrared sensor tracking system," *IEEE Sensors journal*, vol. 16, no. 1, pp. 216–223, 2016.
- [11] C. Premebida, O. Ludwig, and U. Nunes, "LIDAR and vision-based pedestrian detection system," *Journal of Field Robotics*, vol. 26, no. 9, pp. 696–711, 2009.
- [12] J. Shackleton, B. VanVoorst, and J. Hesch, "Tracking people with a 360-degree lidar," in *Advanced Video and Signal Based Surveillance (AVSS), 2010 Seventh IEEE International Conference on*. IEEE, 2010, pp. 420–426.
- [13] T. S. Ralston, G. L. Charvat, and J. E. Peabody, "Real-time through-wall imaging using an ultrawideband multiple-input multiple-output (MIMO) phased array radar system," in *2010 IEEE International Symposium on Phased Array Systems and Technology*, Oct 2010, pp. 551–558.
- [14] A. Koelpin, F. Lurz, S. Linz, S. Mann, C. Will, and S. Lindner, "Six-port based interferometry for precise radar and sensing applications," *Sensors*, vol. 16, no. 10, p. 1556, 2016.
- [15] A. Santra, R. V. Ulaganathan, and T. Finke, "Short-range millimetric-wave radar system for occupancy sensing application," *IEEE Sensors Letters*, vol. 2, no. 3, pp. 1–4, Sept 2018.
- [16] C. Will, K. Shi, S. Schellenberger, T. Steigleider, F. Michler, J. Fuchs, R. Weigel, C. Ostgathe, and A. Koelpin, "Radar-based heart sound detection," *Scientific reports*, vol. 8, no. 1, p. 11551, 2018.
- [17] E. Fishler, A. Haimovich, R. Blum, D. Chizik, L. Cimini, and R. Valenzuela, "MIMO radar: An idea whose time has come," in *Radar Conference, 2004. Proceedings of the IEEE*. IEEE, 2004, pp. 71–78.
- [18] Z. Peng and C. Li, "A portable 24-GHz FMCW radar based on six-port for short-range human tracking," in *2015 IEEE MTT-S 2015 International Microwave Workshop Series on RF and Wireless Technologies for Biomedical and Healthcare Applications (IMWS-BIO)*, Sept 2015, pp. 81–82.
- [19] Z. Peng, L. Ran, and C. Li, "A K-band portable FMCW radar with beamforming array for short-range localization and vital-doppler targets discrimination," *IEEE Transactions on Microwave Theory and Techniques*, vol. 65, no. 9, pp. 3443–3452, 2017.
- [20] Z. Peng, J. Muñoz-Ferreras, Y. Tang, R. Gómez-García, and C. Li, "Portable coherent frequency-modulated continuous-wave radar for indoor human tracking," in *2016 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems (BioWireless)*, Jan 2016, pp. 36–38.
- [21] E. Hyun and J. Lee, "Multi-target tracking scheme using a track management table for automotive radar systems," in *2016 17th International Radar Symposium (IRS)*, May 2016, pp. 1–5.
- [22] Q. Fu and J. Zhang, "Moving target detection based on motion filter and tracking point cohesion for through wall radar," in *2016 CIE International Conference on Radar (RADAR)*, Oct 2016, pp. 1–4.
- [23] J. Zhang, T. Jin, Y. He, L. Qiu, and Z. Zhou, "Foliage-penetration human tracking by multistatic radar," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, July 2016, pp. 6102–6105.
- [24] ———, "Human tracking using range and velocity measurements by multistatic radar," in *2016 Progress in Electromagnetic Research Symposium (PIERS)*, Aug 2016, pp. 520–525.
- [25] A. E. Mitchell, G. E. Smith, K. L. Bell, and M. Rangaswamy, "Single target tracking with distributed cognitive radar," in *2017 IEEE Radar Conference (RadarConf)*, May 2017, pp. 0285–0288.
- [26] Y. He, F. L. Chevalier, and A. G. Yarovoy, "Range-Doppler processing for indoor human tracking by multistatic ultra-wideband radar," in *2012 13th International Radar Symposium*, May 2012, pp. 250–253.
- [27] Y. He, P. Aubry, and F. L. Chevalier, "Ultra-wideband multistatic tracking of human targets," in *IET International Radar Conference 2013*, April 2013, pp. 1–5.
- [28] P. Thomson and B. Almutery, "A case study: The development of an embedded realtime tracking and control application for a tracking radar," in *2011 Saudi International Electronics, Communications and Photonics Conference (SIECPC)*, April 2011, pp. 1–5.
- [29] R. K. Hersey and E. Culpepper, "Radar processing architecture for simultaneous SAR, GMTI, ATR, and tracking," in *2016 IEEE Radar Conference (RadarConf)*, May 2016, pp. 1–5.
- [30] D. Sasakawa, N. Honma, K. Nishimori, T. Nakayama, and I. Shiochi, "Evaluation of fast human localization and tracking using MIMO radar in multi-path environment," in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, Sept 2016, pp. 1–6.
- [31] F. J. Yanovsky, R. B. Sinitsyn, Y. Chernovik, V. Makarenko, V. Tokarev, and O. Zaporozhets, "Moving target detection and tracking using passive acoustic radar," in *2016 IEEE Radar Methods and Systems Workshop (RMSW)*, Sept 2016, pp. 87–90.
- [32] A. Lin and H. Ling, "Two-dimensional human tracking using a three-element Doppler and direction-of-arrival (DDOA) radar," in *2006 IEEE Conference on Radar*, April 2006, pp. 4 pp.–.
- [33] Y. Kim and H. Ling, "Human activity classification based on micro-doppler signatures using a support vector machine," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1328–1337, May 2009.
- [34] S. Hazra and A. Santra, "Robust gesture recognition using millimetric-wave radar system," *IEEE Sensors Letters*, vol. 2, no. 4, pp. 1–4, Dec 2018.
- [35] S. Björklund, "Target detection and classification of small drones by boosting on radar micro-doppler," in *2018 15th European Radar Conference (EuRAD)*, Sep. 2018, pp. 182–185.

- [36] S. Schellenberger, K. Shi, T. Steigleder, F. Michler, F. Lurz, R. Weigel, and A. Koelpin, "Support vector machine-based instantaneous presence detection for continuous wave radar systems," in *2018 Asia-Pacific Microwave Conference (APMC)*, Nov 2018, pp. 1465–1467.
- [37] I. Bilik and J. Tabrikian, "Radar target classification using doppler signatures of human locomotion models," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 43, no. 4, pp. 1510–1522, October 2007.
- [38] T. D. Bufler and R. M. Narayanan, "Radar classification of indoor targets using support vector machines," *IET Radar, Sonar & Navigation*, vol. 10, no. 8, pp. 1468–1476, 2016.
- [39] Y. Kim, S. Ha, and J. Kwon, "Human detection using doppler radar based on physical characteristics of targets," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 2, pp. 289–293, Feb 2015.
- [40] Y. Guerbai, Y. Chibani, and B. Hadjadj, "The effective use of the one-class svm classifier for handwritten signature verification based on writer-independent parameters," *Pattern Recognition*, vol. 48, no. 1, pp. 103–113, 2015.
- [41] C. Jaremenko, E. Affronti, A. Maier, and M. Merklein, "Analysis of forming limits in sheet metal forming with pattern recognition methods. part 2: Unsupervised methodology and application," *Materials*, vol. 11, no. 10, p. 1892, 2018.
- [42] A. Gambardella, G. Giacinto, M. Migliaccio, and A. Montali, "One-class classification for oil spill detection," *Pattern Analysis and Applications*, vol. 13, no. 3, pp. 349–366, 2010.
- [43] A. Santra, R. Santhanakumar, K. Jadia, and R. Srinivasan, "SINR performance of matched illumination signals with dynamic target models," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 4890–4894.
- [44] J. Bock, H. Schafer, K. Aufinger, R. Stengl, S. Boguth, R. Schreiter, M. Rest, H. Knapp, M. Wurzer, W. Perndl *et al.*, "Sige bipolar technology for automotive radar applications," in *Bipolar/BiCMOS Circuits and Technology, 2004. Proceedings of the 2004 Meeting*. IEEE, 2004, pp. 84–87.
- [45] C.-Y. Kim, J.-G. Kim, and S. Hong, "A quadrature radar topology with Tx leakage canceller for 24-GHz radar applications," *IEEE transactions on microwave theory and techniques*, vol. 55, no. 7, pp. 1438–1444, 2007.
- [46] C. Will, S. Linz, S. Mann, F. Lurz, S. Lindner, R. Weigel, and A. Koelpin, "A 24 GHz waveguide based radar system using an advanced algorithm for I/Q offset cancelation," *Advances in Radio Science*, vol. 15, pp. 249–258, 2017.
- [47] M. A. Richards, J. Scheer, W. A. Holm, and W. L. Melvin, *Principles of modern radar*. Citeseer, 2010.
- [48] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Transactions on Intelligent Systems and Technology*, vol. 2, pp. 27:1–27:27, 2011, software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.



Christoph Will (S'15) was born in Mellrichstadt, Germany, in 1988. He received the M.S. degree (with distinction) in information and communication technology from the Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany, in April 2014. From June 2014 till March 2018, he was with the Institute for Electronics Engineering, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany, where he was working as a Research Assistant toward the Ph.D. degree. His research topics were related to innovative radar technology based on Six-Port interferometry for industrial and biomedical applications. Since July 2018 he is working as engineering consultant for Infineon Technology AG, Neubiberg, Germany, regarding baseband optimization of radar systems, radar signal processing, and algorithms for human tracking and vital sign monitoring.

Mr. Will is a member of the IEEE Microwave Theory and Techniques Society (IEEE MTT-S), the IEEE Engineering in Medicine and Biology Society (EMBS), and the German Association for Electrical, Electronic and Information Technologies (VDE). He received the IEEE 2016 Radio Wireless Week Excellent Demo Track Award, Austin, TX, 2016, the Second Prize at the IEEE 2017 Radio Wireless Week Student Paper Competition, Phoenix, AZ, 2017, and the Preis der ITG 2017, Berlin, Germany, 2017. Mr. Will is a reviewer for various IEEE journals.



Prachi Vaishnav from Ahmedabad, India, completed her Bachelors in Engineering from National Institute of Technology, Surat, India, in 2012. After her master thesis at Airbus, Ulm, Germany, in the field of digital beamforming algorithms for airborne radar systems, she received her M.Sc. degree in Communications Technology from the University of Ulm, Germany, in 2015.

Mrs. Vaishnav has worked in the Department of Array based radar Imaging at Fraunhofer Institute for High Frequency Physics and Radar Techniques (FHR) Bonn, Germany, from November 2015 till June 2018. Her focus was developing signal processing algorithms for space surveillance. Since July 2018 she has been working with Infineon Technology AG, Neubiberg, Germany, as a radar software development engineer on topics of human detection and tracking.



Abhiram Chakraborty (M'07) was born in Calcutta, India, in 1984. He received the Bachelor degree in Electronics and Communication Engineering from Anna University, Chennai, India, in 2006 and the M.Sc. degree in Information and Communication Engineering (specializing in Microwaves and High Frequency Circuit theory and design) from the Technical University of Darmstadt, Darmstadt, Germany, in 2009.

In 2009 he joined Infineon Technologies AG as product development engineer for GPS, LNAs, and Modules. From 2011 to 2013 he worked on the design and characterization of V-Band and D-Band millimeter-wave MMICs for wireless backhaul and radar applications in several Silicon Germanium technologies. Since 2014 he is working as a system application engineer for millimeter-wave MMICs at Infineon Technologies AG. During this time he was involved in the development of several V-band, E-band, 24-GHz and 60-GHz radar products at Infineon. His research interests include low power, low noise, and broadband millimeter-wave integrated circuits in SiGe and CMOS technologies for radar and communication application, highly integrated radar systems for industrial and consumer applications, and artificial intelligence and machine learning for radar applications.



Avik Santra (S'09-M'10-SM'18) received his B.E. degree from West Bengal University of Technology and his M.E. degree in Signal Processing with first class distinction from Indian Institute of Science, Bangalore, India, in 2010. He is currently working as Radar Algorithm developer developing signal processing and machine learning algorithms for industrial and consumer radars and depth sensors at Infineon, Neubiberg, Germany. Earlier in his career, he has worked as system engineer for LTE/4G modem at Broadcom Communications, and also has worked as research engineer developing fully adaptive phased array and MIMO Radars at Airbus Group. He is a reviewer at various IEEE and Elsevier journals, and is recipient of several outstanding reviewer awards. He has filed over 35 patents and published 15 research papers related to various topics of radar waveform design, radar signal processing, and radar machine/deep learning topics.