

Human gait model individualized by low-cost radar measurements

Roy Veld

Faculty of Electrical Engineering, Mathematics and Computer Science

University of Twente

Enschede, The Netherlands

Abstract—In this paper, a methodology for developing a human walking model individualized by radar measurements was proposed. Fundamental parameters encompassed in the model are the gait parameters and physical dimension parameters. Empirical validation of the proposed methodology is undertaken, involving the acquisition of data using a 77-GHz FMCW radar. The data is collected from three distinct individuals walking five diverse trajectories with respect to the radar. Moreover, the gait parameter estimation accuracy is evaluated for the different walking trajectories of the target. The studied gait parameters are speed, step length and step frequency. These could be estimated with a mean error up to 0.077 m/s, 9.3 cm, and 0.128 Hz for all trajectories, respectively. Nevertheless, these errors diminish to 0.022 m/s, 2.2 cm and 0.03 Hz, respectively when the targets walk in a straight trajectory aligned with the radar beam. Moreover, the feasibility of estimating body part dimensions directly from the radar data is investigated. It was found that only the total human height could be directly estimated using the employed hardware. Except for the tallest participant of 2.01 m, the height could be estimated with a mean absolute error up to 10.9 cm. Enhanced hardware configurations or the integration of machine learning techniques may improve the accuracy of body part dimension estimations.

Index Terms—human gait, FMCW radar, Doppler signatures, walking model

I. INTRODUCTION

Human gait is a distinctive motion pattern that reflects unique characteristics of an individual. As a fundamental aspect of human locomotion, gait serves as a signature of the movement style and the physical condition of the person. The development of an accurate individualized model of human walking, capable of capturing these distinct signatures, holds great value for a variety of applications.

Significant applications where the gait model finds utility lie in healthcare and sports, specifically in gait monitoring for diagnostic and rehabilitative purposes. Medical professionals can analyze the gait model and its associated parameters to gain insights into medical conditions. As an example, several studies have demonstrated that self-selected gait speed and step length serve as reliable markers for diagnosing Parkinson’s disease [1]–[3]. On top of that, progression or degeneration in the human motor system and other change in gait features can be analyzed when comparing models captured at different time instances. Another field of application is found within computer animation. The use of gait models in computer animation makes virtual characters more realistic. Human gait representation is crucial for creating lifelike and appealing

characters in animated films, video games, and virtual reality (VR) experiences. The integration of an individualized gait model into animation enables persons to incorporate their own unique walking style into their virtual twins. Furthermore, having an accurate model on the past motion of a person may be used to predict the future motion. This prediction can be used in for example smart home systems. By utilizing gait models, smart home systems can employ predictive gait modelling to anticipate the needs of residents, such as adjusting lighting or room temperature based on their anticipated movements.

Gait measurements have been obtained through various sensing methods, such as making use of vision-based sensors [4] [5]. However, the use of vision-based sensors raises concerns about privacy, and their performance can be easily affected by varying light conditions, such as low target illumination due to darkness or smoke. Alternatives utilizing laser optics, as demonstrated in [6] using lidar devices, can generate high-resolution point clouds for gait analysis but come with a higher cost. Another approach involves the use of motion sensors that make contact with the body, e.g. utilizing accelerometers and gyroscopes [7]. These motion sensor-based methods exhibit high measurement accuracies. Nevertheless, one major drawback of wearables lies in the inconvenience of wearing the devices, as individuals may forget or feel uncomfortable with wearing them.

In contrast, radar-based measurements offer several advantages. Radar-based systems are low-cost, accurate, privacy-insensitive and contactless, making them suitable for gait monitoring applications. Consequently, there are plenty of studies that use radar for human gait measurements, e.g. [8]–[13]. These works have successfully extracted gait parameters such as step time, step length, cadence and walking speed using monostatic radars. In the cited studies, the gait parameters were determined by exploiting the range and Doppler information over time. The measurements were done with the target walking either on a treadmill, or on a straight trajectory walking inbound or outbound with respect to the radar. The Doppler signature is well-expressed in these cases, since high Doppler shifts can be measured. In a realistic scenario, however, the target is unlikely to always follow a path aligned with the radar beam. When the location and orientation of the human relative to the radar changes, the Doppler signatures are affected as well. Hence, the gait estimation performance is expected to deteriorate. This has not been investigated in

earlier work before. Therefore, in contrast to existing work on gait measurements with radar, the angular dependency on the accuracy of gait parameter estimation is studied in this work.

In addition to accurately determining gait parameters, an essential aspect in creating a lifelike human walking model involves estimating the size of body parts. Several studies exist in which the human body pose and/or keypoints are estimated, applying different methods [14]. In [15] and [16], ellipses are fitted on a 2D input image made by a camera. These ellipses represent the body parts, scaled and rotated dependent on size and location of the human. In [17], a depth camera is used to generate a dense point cloud for body pose and shape estimation using a mathematical method. The aforementioned techniques are not feasible for radar, because here detected data points in space are rather sparse compared to an image. The sparse nature of the spatial radar scatterers presents a difficulty when it comes to estimating the size of body parts. To realize pose estimation from radar point clouds, machine learning (ML) was utilized in [18]–[20]. Using this technology, locations of human keypoints could be estimated within a certain error margin. However, a huge amount of data and time is required to train a reliable neural network (NN). On top of that, training the NN requires a large variety of training data containing lots of different people and environments in order to make the estimation more versatile. Another rather simple method to deduce the human height from radar measurements has been presented before in both [21] and [22]. The gait parameter was substituted into a formula from [23] which was determined based on experimental data, in order to estimate the physical height of the human body. Thus, the presented method is highly dependent on the accuracy of the gait parameter estimation. Notably, the validation of the method in the former paper was limited to a single individual and was solely evaluated through simulations rather than empirical measurements. The method was tested using measurements on multiple targets in the latter paper, and the experiments had shown the formula can not be used directly for accurate body height estimation. Furthermore, in [22] a ML technique called random forest for regression was used to determine the height using gait parameters, resulting in slightly lower errors than the earlier method. In the current research, the feasibility of estimating the human body dimensions directly from the post-processed radar data is studied. The determination of the body height parameter is conducted without utilizing any ML technique. This omits the need for large datasets, time and costs accompanied by it. Furthermore, the body height is estimated independently, without any direct empirical reliance on the gait parameters. Therefore, the estimation is less generalized compared to earlier work using typical relations between gait parameters and height. Also, the body dimension estimation performance will be evaluated for different aspect angles.

The contributions of this work are threefold:

- 1) Study of the angular dependency on gait parameter estimation accuracy when utilizing a monostatic radar.
- 2) Study the feasibility of estimating body part dimensions from processed radar data and present a method of

estimating the target’s height directly from the radar data. The height estimation capability is evaluated under different target trajectories with respect to the radar.

- 3) Present a novel pipeline for generating an individualized spatio-temporal human walking model from monostatic radar measurements.

The remainder of this paper is organized as follows. In the next section, the related background information is discussed. Section III gives a description of the presented pipeline. In section V, results are presented and discussed, as well as the limitations of this work. Finally, conclusions are drawn and recommendations for future work are made in section VI.

II. THEORETICAL BACKGROUND

In radars, electromagnetic (EM) waves are emitted by an antenna, which are potentially reflected by objects in the environment. After reception of the reflected signal, the range, velocity and angle of the objects can be determined. The time needed for the wave to travel from the transceiver to an object at distance r and back is given by $\tau = 2r/c_0$, with c_0 the speed of light. Current highly accurate millimeter wave (mmWave) sensing devices employ frequency-modulated continuous-wave (FMCW) signals with multiple-input multiple-output (MIMO) technology [24]–[26]. An FMCW radar continuously transmits signals called chirps. The signal frequency of a chirp increases linearly over time with slope S . The chirp is characterized by the start frequency f_c , the bandwidth B and the chirp duration T_C (chirp cycle time).

A. Range estimation

Before being able to estimate the radar parameters (range, velocity, angle), the transmitter (TX) and receiver (RX) signals are mixed in a quadrature mixer and low-pass filtered. The output is an intermediate frequency (IF) signal, which is digitized by an analog-to-digital converter (ADC). The instantaneous frequency of the IF signal (also called beat frequency) is equal to the difference of the instantaneous frequencies of the two inputs of the mixer:

$$f_{IF} = f_{TX}(t) - f_{RX}(t) = S\tau = \frac{S2r}{c_0}. \quad (1)$$

This frequency is not a function of time and stays constant, as can be seen in Fig. 1.

When multiple objects are present at distinct distances from the radar, the range fast-Fourier transform (FFT) can be taken over the chirp samples (i.e. fast time) to separate them. When the chirp signal is transformed to the frequency domain, for every detected object a peak is visible in the spectrum. The frequency axis, obtained when the IF signal is sampled N_s times with a sampling frequency of F_s by the ADC, is directly related to the range information:

$$r_k = k \frac{c_0 F_s}{2SN_s} = k \frac{c_0}{2S\tilde{T}_c} = k \frac{c_0}{2\tilde{B}}, \quad (2)$$

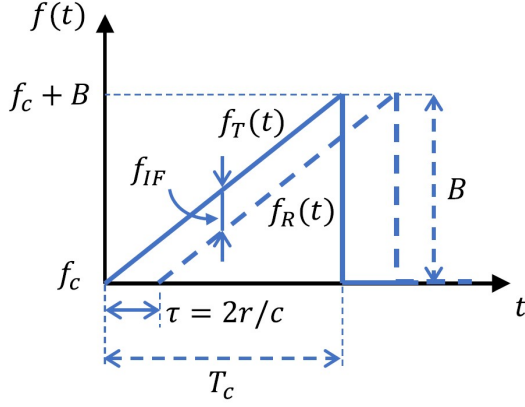


Fig. 1: IF signal derivation from the transmitted and received chirp signals. (From [25]).

with $0 \leq k < N_s$. $\tilde{T}_c = N_s/F_s$ is the length of the sampled IF signal in time, also known as the ADC sampling window length. $\tilde{B} = S\tilde{T}_c$ is the sweep bandwidth of the chirp signal.

The FMCW radar entails a resolution under which objects can not be differentiated anymore. The range resolution, which represents the minimum distance required to distinguish between two objects, can be determined by the following expression:

$$r_{res} = \frac{c_0}{2\tilde{B}}. \quad (3)$$

A limit exists on FMCW radar range, called the maximum unambiguous range. Above this limit, the time duration for the EM wave to fulfil a roundtrip is longer than the chirp duration. The radar cannot ascertain whether the echo originates from the most recently transmitted chirp or the preceding one. This situation arises when the range exceeds a certain value of

$$r_{max} = \frac{c_0\tilde{T}_c}{2}. \quad (4)$$

Note the dynamic range of Equation 2 is between zero and r_{max} and does not include negative values. This is because the complex IF signal has a one-sided positive frequency spectrum with no power on the negative side of the spectrum. Therefore, the frequencies $F_s/2 < f_{IF} \leq F_s$ which are aliased towards the negative frequencies (between $-F_s/2$ and 0) can be interpreted as the positive frequencies above the Nyquist sampling frequency. Thus, the spectrum is extended to F_s .

B. Velocity estimation

To determine the velocity v of an object, it is necessary to transmit at least two chirps. A sequence of equally-spaced chirps is called a chirp frame. The range-FFT may have peaks in the same location, but with different phases. Depending on the phase difference between two chirps $\Delta\phi = 4\pi vT_c/\lambda$, the velocity can be determined because the phase changes linearly due to equal spacing of the chirps in the frame:

$$v = \frac{\lambda\Delta\phi}{4\pi T_c}. \quad (5)$$

Here, λ is the wavelength of the chirp signal. By performing the Doppler-FFT over the chirps (i.e. slow-time), the velocities of different objects can be determined. For two objects at the same distance but different velocities, the phase between the received chirps will vary, enabling the differentiation and separation of the two objects. For a set of N_c chirps spaced by T_c , the velocity is given by

$$v_l = \omega_l \frac{\lambda}{4\pi T_c} = l \frac{\lambda}{2N_c T_c} = l \frac{\lambda}{2T_f}, \quad (6)$$

with discrete-time angular frequency $\omega_l = 2\pi l/N_c$ and index $-N_c/2 \leq l < N_c/2$. $T_f = N_c T_c$ is the frame duration time.

Analogous to the range resolution, the velocity resolution is also constrained. The velocity resolution is given by

$$v_{res} = \frac{\lambda}{2T_f} = \frac{\lambda}{2N_c T_c}, \quad (7)$$

Furthermore, a detectable velocity range exists. This is given by

$$-\frac{\lambda}{4T_c} \leq v < \frac{\lambda}{4T_c}. \quad (8)$$

A positive velocity means the object is moving away from the radar, and a negative velocity indicates the object is approaching the radar.

C. Angle estimation

When the radar antenna consists of at least two elements, it becomes feasible to perform direction of arrival (DOA) estimation. Due to the distinct locations of these elements, the measured distances from the same reflected signal exhibit slight differences. These differences arise from the time delay $\Delta\tau = d \sin(\theta)/c_0$ associated with the DOA angle θ , where d represents the antenna spacing. Therefore, the received phase depends on the element location in space. The phase difference between two elements is denoted as $\Delta\phi$. It is assumed the DOA is equal to the direction of departure (DOD), i.e. the antenna aperture dimensions are small with respect to the distance to the reflecting object. In other words, a far-field assumption is made here, with the measured target residing at a distance larger than the Fraunhofer distance $d_{Fraun} = 2D^2/\lambda$, in which D is the maximum target dimension [27]. The DOA can be computed as

$$\theta = \sin^{-1} \left(\frac{\lambda\Delta\phi}{2\pi d} \right). \quad (9)$$

The angular resolution at boresight view with half wavelength antenna spacing is given by

$$\theta_{res} = 2/N_R \quad (10)$$

in unit radians. N_R is the number of receive antennas here.

1) *TDM MIMO radar*: As shown in Equation 10, the angular resolution depends on the number of receive antennas. By increasing this number, the angular resolution is enhanced accordingly. This improvement can be effectively achieved by employing MIMO with N_T transmit antennas. When utilizing an $N_T \times N_R$ MIMO array antenna, a virtual array of $N_T \times N_R$ antennas can be synthesized, increasing angular resolution. For instance, in Fig. 2 the array has four RX elements with spacing d and two TX elements with spacing $4d$. Due to the difference in travel distances of the signal from each transmit element, which is affected by the angle θ as $4d \sin(\theta)$, eight distinct channels are created, each with a unique time of flight. Therefore, the MIMO array can be treated as an eight-element uniform linear array (ULA).

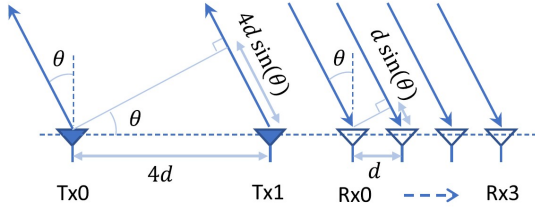


Fig. 2: Angle estimation using MIMO radar. (From [25]).

The multiple transmit antennas are transmitting the chirp signal sequentially using time-division multiplexing (TDM). Within one frame, each TX transmits N_c chirps. The chirp time duration is extended by a factor N_T , leading to an increased chirp time of $T_{c,TDM} = T_c N_T$. The improved angular resolution comes at the expense of reduced maximum velocity because the chirp time increases for the same frame duration. Alternatively, the velocity resolution is deteriorated if the chirp time is kept constant, resulting in a higher frame duration.

2) *MVDR beamforming*: When the DOA is estimated by taking the FFT over the spatial samples, the resolution is constrained by the number of elements. To overcome this limitation and achieve higher resolution, various methods can be employed including beamforming methods, subspace methods and parametric methods [28]. In this work, a beamforming method is adopted because it is sufficiently accurate for the intended purpose while having a relatively low computational complexity. Hence, only this method will be explained hereafter. A beamformer is a spatial filter which combines weighted element outputs linearly using a weight vector \mathbf{w} . For an array with M elements, the weighted signal is given by

$$y(n) = \sum_{m=1}^M w_m^* x_m(n) = \mathbf{w}^H \mathbf{x}(n). \quad (11)$$

With N_C time samples, the antenna observation signal of element m is given by $x_m(n)$ with $n = 1, \dots, N_C$. This signal represents the slow-time signal (i.e. signal consisting of chirp samples in a frame), which is assumed to be narrowband.

The narrowband assumption holds if the product of the signal bandwidth B_s and the maximal travel time ΔT between two array elements for an impinging plane wave $B_s \Delta T \ll 1$. In the system being utilized, this condition is met as the obtained IF signal, which is fed to the ADC, is narrowband. When all M elements are included, the observation signal is written as a vector $\mathbf{x}(n)$. The power spectrum for angle θ is computed by

$$P(\theta) = \frac{1}{N} \sum_{n=1}^N |y(n)|^2 = \mathbf{w}^H(\theta) \hat{\mathbf{R}}_x \mathbf{w}(\theta). \quad (12)$$

The beampattern displays a prominent peak in the spatial power spectrum, indicating the angular location where the highest power is concentrated. This corresponds to the direction from which most of the reflected power originates due to the presence of an object at that specific angle. The conventional beamformer, also known as the delay-and-sum beamformer, combines the output of each element coherently to obtain an enhanced signal. The downside of this method is the limited resolution when the number of antenna elements is small. The resolution of the minimum variance distortionless response (MVDR) beamformer is less limited by the number of elements, creating a sharp peak at the target DOA [29]. This method was chosen because it has a higher spatial resolution than a conventional beamformer. On top of that, contrary to e.g. the MUSIC algorithm, the number of sources (i.e. scattering directions) does not have to be known and can vary arbitrarily. The MVDR beamformer is based on the following constrained optimization problem:

$$\min \mathbf{w}^H(\theta) \hat{\mathbf{R}}_x \mathbf{w}(\theta) \quad (13)$$

$$\text{subject to } \mathbf{w}^H(\theta) \mathbf{a}(\theta) = 1, \quad (14)$$

with $\mathbf{a}(\theta)$ the steering vector and $\hat{\mathbf{R}}_x$ the sample correlation matrix. The output power is minimized (first function), while ensuring that the signals from the desired direction θ remain undistorted (second function). So, power from all other directions is minimized while the beamformer concentrates only in one direction.

The sample correlation matrix is as follows:

$$\hat{\mathbf{R}}_x = \frac{1}{N} \sum_{n=1}^N \mathbf{x}(n) \mathbf{x}(n)^H. \quad (15)$$

The solution of the optimization problem in Equations 13 and 14 gives an expression for the weight vector:

$$\mathbf{w}_{MVDR}(\theta) = \frac{\hat{\mathbf{R}}_x^{-1} \mathbf{a}(\theta)}{\mathbf{a}^H(\theta) \hat{\mathbf{R}}_x^{-1} \mathbf{a}(\theta)} \quad (16)$$

The power function is then obtained by substituting Equation 16 into Equation 12:

$$P_{MVDR}(\theta) = \frac{1}{\mathbf{a}^H(\theta) \hat{\mathbf{R}}_x^{-1} \mathbf{a}(\theta)} \quad (17)$$

It is important to note that a matrix inversion of the sample correlation matrix is carried out, so this matrix needs to be

full rank. In other words, the number of temporal samples must be larger than the number of elements ($N > M$). The power function is computed for every range value, such that a range-azimuth plot is created.

D. Signal model

1) *Wavefield description*: Due to the far-field assumption, it is known the wave impinging on the antenna from every scattering point is a plane wave. The plane wave propagating and the coordinate system with respect to the array antenna can be seen in Fig. 3. θ is the azimuth angle, which is the angle between the x-axis and the orthogonal projection of the signal vector on the x-y plane. ϕ denotes the elevation angle, which is the angle between the signal vector and the z-axis. The Cartesian coordinates of an arbitrary antenna element are given by $\mathbf{r} = [x, y, z]$. The spatial frequency of the incoming wave is

$$\mathbf{k} = -\frac{\omega}{c_0} \begin{bmatrix} \cos(\theta) \sin(\phi) \\ \sin(\theta) \sin(\phi) \\ \cos(\phi) \end{bmatrix}. \quad (18)$$

The spatial frequency is related to the temporal frequency through $\|\mathbf{k}\|^2 = \omega^2/c_0^2$. The steering vector for a wave coming in on an array of M elements is given by

$$\mathbf{a}(\theta, \phi) = \begin{bmatrix} e^{-j\mathbf{k} \cdot \mathbf{r}_1} \\ \vdots \\ e^{-j\mathbf{k} \cdot \mathbf{r}_M} \end{bmatrix}. \quad (19)$$

This expression for the steering vector can be substituted into Equation 17 to find the power function of the MVDR beamformer.

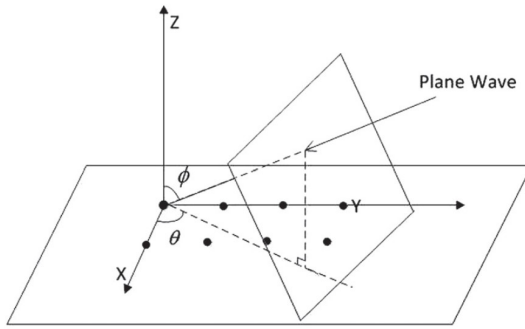


Fig. 3: Coordinate system for impinging plane wave on antenna elements. (From [28]).

2) *Backscattering*: The FMCW emits a chirp of finite length. The signal propagates to the target and its energy is reflected by the target in K scattering points. Then, the signal is received by the radar again. If the operating frequency f_0 is high enough (i.e. the wavelength small compared to target dimensions), the scattering behaviour can be modelled as a summation of the different scattering responses [30]. The signal model employed in this context [31] provides

the received baseband signal, which can be mathematically represented as

$$s(t) = \sum_{k=1}^K \rho_k \exp\left(j2\pi f_0 \frac{2R_k(t)}{c_0}\right) = \sum_{k=1}^K \rho_k \exp(j\Phi_k(t)), \quad (20)$$

with ρ_k the reflectivity function, R_k the scalar range between the radar and the scattering point and Φ_k the phase of the baseband signal of the scattering point k . The reflectivity function contains factors like path loss, antenna gain, radar cross section (RCS), transmit power and other losses like atmospheric loss and system loss [32]. So, the target is regarded as a number of point scatterers moving in space over time. Each scatterer induces a phase delay proportional to the range and a frequency shift proportional to the scatterer velocity. The phase of the signal reflected by scatterer k is

$$\Phi_k(t) = 4\pi f_0 R_k(t)/c_0. \quad (21)$$

The Doppler frequency shift induced by a single point scatterer moving over distance, varying the scalar range R_k over time, is obtained by taking the time derivative of the phase:

$$f_{D,k} = -\frac{1}{2\pi} \frac{d\Phi_k(t)}{dt} = -\frac{2f}{c_0} \frac{d}{dt} R_k(t) = -\frac{2f}{c_0} v_k(t) = -\frac{2}{\lambda} v_k(t), \quad (22)$$

in which $v_k(t) = \frac{dR_k(t)}{dt}$ is the radial velocity of the point k in the moving object [33]. So, if the range with respect to the radar stays constant over time, no Doppler shift will be visible. Contrary, if the object is moving directly towards or away from the radar, the highest Doppler shift is visible since the range changes the most over time. A positive Doppler shift means the target is moving towards the radar, and a negative Doppler shift indicates the target moves away from the radar.

The unit vector of the radar line of sight (LOS) direction is given by

$$\mathbf{n} = \mathbf{R}_0 / \|\mathbf{R}_0\| = [\cos(\theta) \sin(\phi), \sin(\theta) \sin(\phi), \cos(\phi)]^T, \quad (23)$$

with \mathbf{R}_0 the distance vector from the target to the radar, θ and ϕ the azimuth and elevation angles in the radar coordinate system, respectively. $\|\cdot\|$ denotes the Euclidean norm. When the object is moving in space over a distance r towards a direction given by the unit vector \mathbf{n}_m , the observed micro-Doppler shift induced by the movement given in Equation 22 can also be written as

$$f_{D,k} = -\frac{2f}{c_0} \left(\frac{d}{dt} r(t) \mathbf{n}_m \right)^T \cdot \mathbf{n}. \quad (24)$$

So, when the directions of \mathbf{n} and \mathbf{n}_m are either equal or equal with opposite sign, the observed Doppler shift is maximum. When the movement unit vector is perpendicular to the radar LOS direction, the observed Doppler shift is minimum.

A distinct object is characterized by multiple scattering points. The different velocities of different moving objects

result in different Doppler signatures, which can be used to distinguish between them. For example, the lower legs of a human body have a higher Doppler shift than the torso, as they move with greater velocity in the direction of the radar. However, the separation of velocity profiles for different body parts presents a challenge [34]. This difficulty arises due to the received signal being a superposition of signals from multiple body parts, each with different Doppler shifts and RCSs. Additionally, multipath fading can introduce additional components to the received signal, making the task even more complex. Adding to the challenge, the radar can only measure the radial velocity, making it unable to directly measure the instantaneous velocity of individual body parts, as illustrated by Equation 24.

III. PROPOSED PIPELINE

The overall pipeline describing the methodology is illustrated in Fig. 4. The FMCW radar acquires measurements from the human subject. Subsequently, the acquired signal undergoes post-processing, where various radar processing techniques are employed, as elaborated in the succeeding paragraphs. After post-processing, the extracted parameters are the human height dimension and the gait parameters. These parameters are then used to individualize the spatio-temporal walking model. In the present study, only the human height dimension is estimated; nevertheless, the proposed pipeline remains applicable when estimating dimensions of multiple human body parts, such as leg lengths, torso length and width, and so on.

A. Post-processing pipeline

Post-processing is a crucial part within the overall pipeline. Following the radar measurement, the raw radar data undergoes post-processing to extract valuable information. This involves a series of steps, as depicted in Fig. 5. In the subsequent paragraphs, each block of the post-processing pipeline will be explained in detail.

1) *Radar data representation*: The output data from the ADC is initially recorded in a binary format during the measurement process. After parsing, the data is represented comprehensively in three dimensions. This three-dimensional representation takes the form of a cube, as depicted in Fig. 6. The first dimension entails the N_s samples of each chirp IF signal, or fast-time. The second dimension involves the N_c chirps in one frame, referred to as slow-time. Lastly, the third dimension represents the $N_R \times N_T$ virtual receivers. As previously mentioned in Section II, the three dimensions can be converted to range, Doppler and angle, respectively.

2) *Static clutter removal*: The radar system may encounter undesired echoes, referred to as clutter, which arise from contributions of objects other than the target. To ensure that only the behavior of the target is prominently visible in the acquired data, it is crucial to minimize the effects of clutter. Static clutter from direct reflections from surfaces or objects such as walls and furniture have zero velocity and are therefore described by a DC component in the Doppler-FFT plot. A

straightforward yet effective approach to eliminate this DC component is by subtracting the average value from the signal values [34]. The average signal value over slow-time in a single frame is computed and subtracted from the received signal in the frame. The number of chirps is given by N_c and the signal is given by $y[n]$. The computation of the averaged signal is as follows:

$$\hat{y}(n) = y(n) - \frac{1}{N_c} \sum_{n=1}^{N_c} y(n). \quad (25)$$

3) *Range-azimuth estimation*: Following static clutter removal, the resulting signal mainly consists of contributions from the non-static target of interest. The next step is to estimate the range r between the radar and the target. This is done by taking the FFT over the fast-time samples. The result is a range spectrum in which the peaks correspond to the principal wave reflections. Exploiting the signals of the antenna elements in the horizontal plane (x-y plane), the azimuth angle θ with respect to the radar can be estimated using the theory described in Section II-C. After doing this for every range value, a range-azimuth plot is created. An example of a range-azimuth plot is shown in Fig. 7. This plot characterizes the range and azimuth DOA of the reflected waves. Subsequently, the spherical coordinates are converted into Cartesian coordinates to obtain the locations in the x-y plane (top view) for each frame:

$$x = r \cos(\theta) \quad (26)$$

$$y = r \sin(\theta) \quad (27)$$

The center location of the target in the x-y plane is computed by taking the mean location of the detected target points.

4) *Target detection*: When the range and azimuth angle of an object with respect to the radar are determined, its location can be estimated as described above. The location is expressed from a top view perspective (i.e. range-azimuth plot). The range-azimuth plot contains the targets, but is also likely to be surrounded by clutter and noise which are time- and position-variant. To effectively distinguish the target within this dynamic environment, a detector with a variable threshold is required, adapted to the background noise. The detector which is used is the constant false alarm rate (CFAR) detector. It is designed to maintain a predefined false alarm rate, ensuring that detections above a certain threshold are labelled as part of potential targets. A false alarm occurs when an object is erroneously detected as a target when it is, in fact, clutter or noise. This method is the most common algorithm for target detection, and it can be readily applied to the range-azimuth (or range-elevation) plot. Hence, it will be used in this work. In a cell-averaging constant false alarm rate (CA-CFAR) detector the noise level is estimated using the samples of N_{cells} reference cells surrounding the cell under test (CUT) [35] [36]. The one-dimensional CA-CFAR detector can be seen in Fig. 8. The noise level is the average value of the reference cells. The scaling factor B is determined using a desired probability of false alarm P_{FA} :

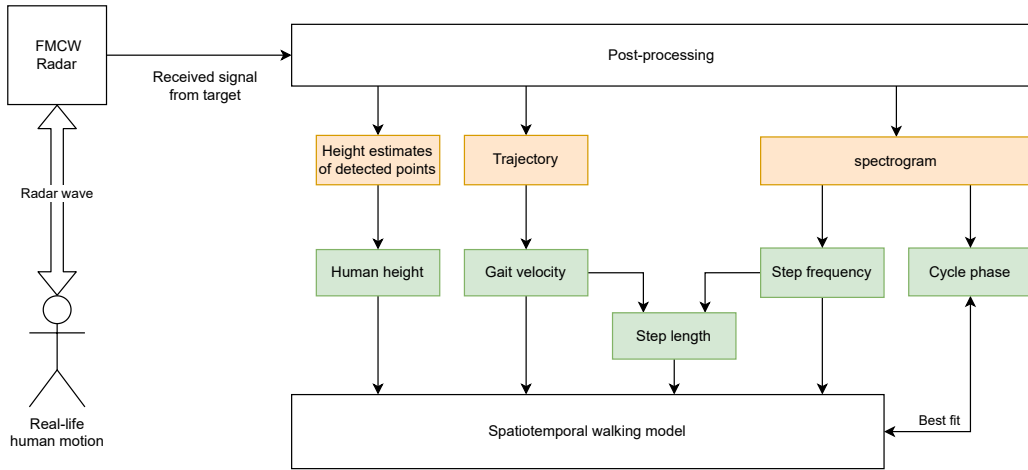


Fig. 4: Proposed individualized gait model pipeline.

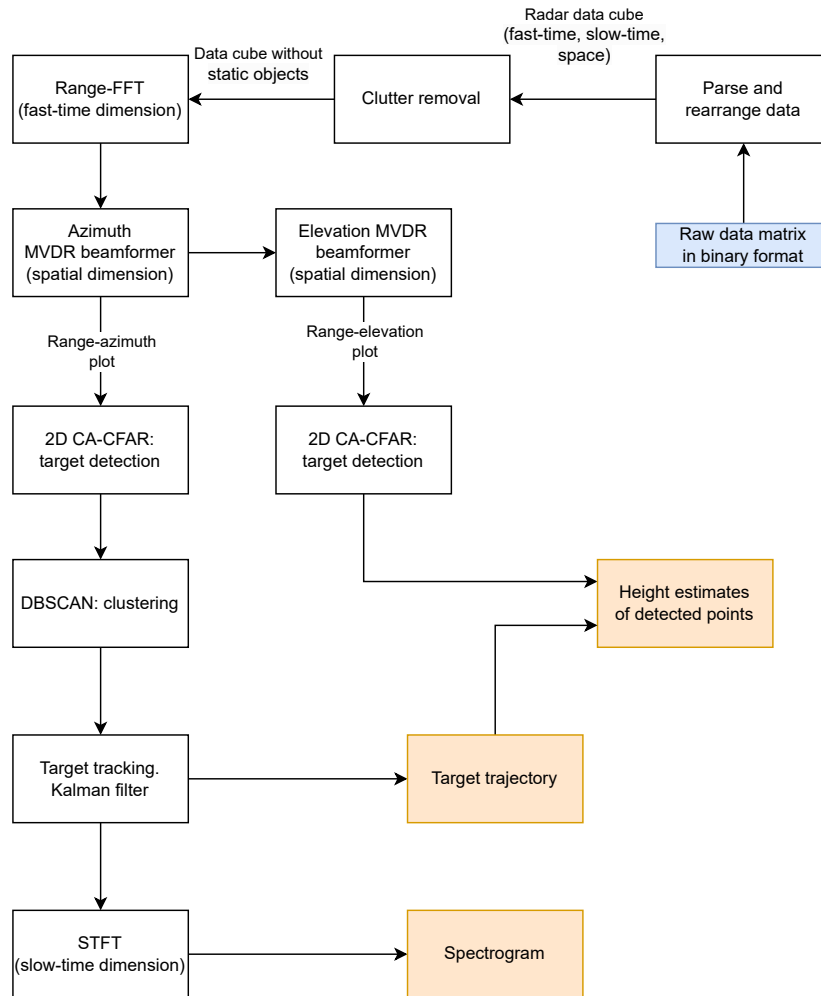


Fig. 5: Post-processing pipeline.

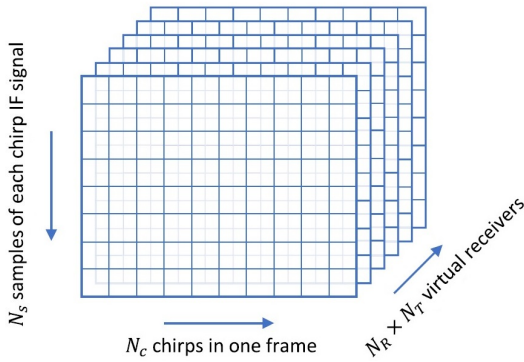


Fig. 6: Radar cube. (From [25]).

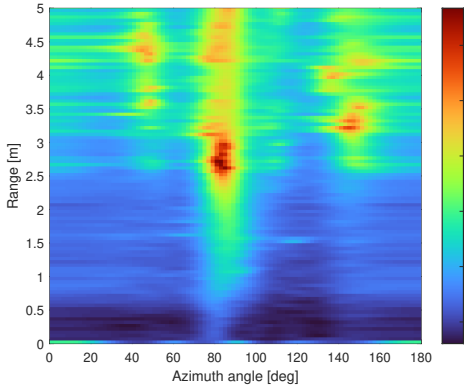


Fig. 7: Range-azimuth plot.

$$B = N_{cells}(P_{FA}^{-1/N_{cells}} - 1) \quad (28)$$

The CUT is surrounded by a designated number of guard cells. These cells are intentionally excluded from the averaging process to ensure that the noise level remains uninfluenced by the presence of the target. Without the inclusion of guard cells, the computed noise level would include some of the target power as well because the target is usually spread over multiple cells. Instead of one dimension, the reference cells and guard cells are taken in two dimensions within this study. The guard cells encircle the CUT, while the reference cells surround the guard cells once more. This two-dimensional arrangement allows for an effective estimation of the noise level while minimizing the impact of the target's signal on the computation.

5) *Clustering*: The process of identifying potential target cells is followed by clustering these cells to group together points belonging to the same object, enabling the differentiation of multiple objects while marking outliers. To achieve this, the density-based spatial clustering of applications and noise (DBSCAN) algorithm is utilized [38]. This algorithm is employed due to its ability to mitigate the presence of non-static clutter induced by multipath reflections, which could not be omitted before by the clutter removal. The clustering

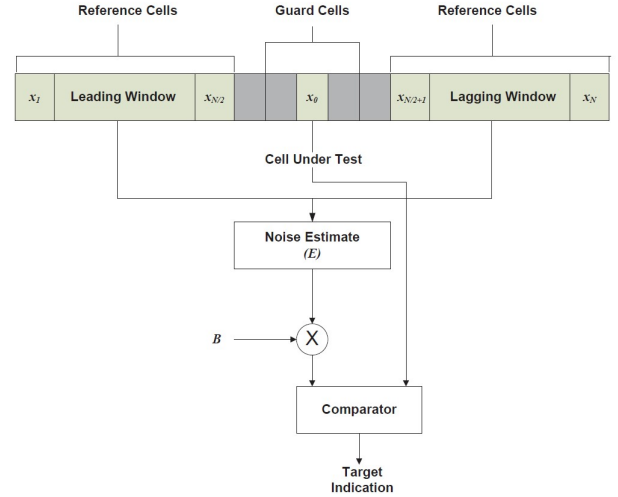


Fig. 8: CA-CFAR architecture. (From [37]).

process relies on two input parameters, namely the neighbour search radius (ϵ) and the minimum number of points to identify a core point ($MinPts$). The algorithm starts with an arbitrary point p and finds all points within reach with respect to ϵ and $MinPts$. If the minimum number of points is not found within the search radius, p is a noise point (outlier). However, if the condition is met, p is a core point and a cluster is established. The algorithm iterates over the neighbouring points and retrieves all core points and border points belonging to the detected cluster. The algorithm repeats the search for all points (detected cells) until every point is labelled either as part of a cluster or as an outlier.

The detected and clustered points in the range-azimuth plot of Fig. 7 are highlighted in Fig. 9. The main cluster depicted in green at the center represents the target's location, containing P detected points, each characterized by unique values for r and θ . The smaller cluster on the right (colored in blue with red outliers) corresponds to clutter and is neglected automatically by the algorithm as it does not contain relevant information on the target of interest. This is done knowing only a single target resides in the room, corresponding to the bigger cluster.

After applying the DBSCAN algorithm to the points detected by CFAR, it is utilized once again on the center locations of the target. The center locations were obtained by taking the mean in both the x and y direction (i.e. dimensions of x-y plane shown in Fig. 3), representing the trajectory samples over the time frames. The clustering operation of the center locations makes sure (non-static) clutter is not part of the target trajectory and can be ignored. Furthermore, if a situation occurs in which multiple targets are present in the scene, these targets can be separated when they are not in proximity of each other. Nevertheless, this research only focuses on single targets, hence this situation is not taken into account.

6) *Trajectory filtering*: From the previous steps, the trajectory samples (center locations) of the target are obtained.

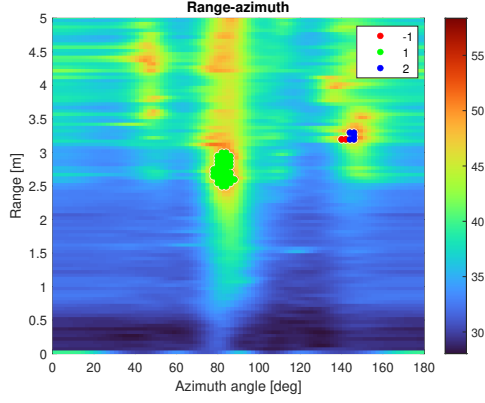


Fig. 9: Detected and clustered points in the range-azimuth plot of Fig. 7.

However, these samples may not all be accurately representing the torso location of the human due to measurement uncertainty. To make the temporal estimates more robust, filtering is applied to the system state \mathbf{x}_n , which contains both the location (x_n, y_n) and velocity (\dot{x}_n, \dot{y}_n) of the target in two dimensions:

$$\mathbf{x}_n = [x_n \quad \dot{x}_n \quad y_n \quad \dot{y}_n]^T \quad (29)$$

First, the future target position is predicted using Newton's motion equations, relying on knowledge of the current state. The applied model assumes the velocity between two adjacent samples in the trajectory to be constant. This motion model is applicable because it is presumed the human walks with approximately the same gait speed during the whole trajectory. Nonetheless, the actual motion might deviate from the model, leading to a dynamic model uncertainty known as process noise. So, a prediction algorithm is required which takes into account both the measurement noise and the process noise. Both the measurement noise and the process noise are assumed to be independent zero-mean Gaussian white noise. The algorithm used to do this is the linear Kalman filter [39]–[41]. This filter consists of two parts.

The first part is the time update, in which the human location for the next frame is predicted by looking forward in time. This is done by using the state extrapolation equation, with the motion equations embedded in the state transition matrix \mathbf{F} :

$$\hat{\mathbf{x}}_{n+1}^- = \mathbf{F}\hat{\mathbf{x}}_n, \quad (30)$$

with $\hat{\mathbf{x}}_n$ the estimated system state vector at time step n and $\hat{\mathbf{x}}_{n+1}^-$ the predicted system state vector for time step $n+1$. Also, the uncertainty of the prediction is provided using the covariance extrapolation equation:

$$\mathbf{P}_{n+1}^- = \mathbf{F}\mathbf{P}_n\mathbf{F}^T + \mathbf{Q}, \quad (31)$$

with \mathbf{P}_n the covariance matrix of the state estimation at time step n , \mathbf{P}_{n+1}^- the predicted covariance matrix of the next state estimation and $\mathbf{Q} = E\{\mathbf{w}_n\mathbf{w}_n^T\}$ the process noise matrix,

with \mathbf{w} the process noise vector. If no previous iterations are available, an initial estimate of both $\hat{\mathbf{x}}_0$ and \mathbf{P}_0 has to be made. This initial estimate of the location and velocity is based on the non-filtered samples. The initial location guess is taken to be the first sample of the non-filtered trajectory. The initial velocity guess is based on the first five non-filtered samples of the trajectory. To get this initial velocity estimate, the covered distance after five samples is divided by the elapsed time while covering this distance.

The second part of the Kalman filter is the measurement update, which combines this prediction with the actual measurement to correct its location. The first step is to compute the Kalman gain \mathbf{K}_n from the previously predicted covariance matrix and the observation matrix \mathbf{H} :

$$\mathbf{K}_n = \mathbf{P}_n^- \mathbf{H}_n^T (\mathbf{H}_n \mathbf{P}_n^- \mathbf{H}_n^T + \mathbf{R}_n)^{-1}, \quad (32)$$

in which $\mathbf{R}_n = E\{\mathbf{v}_n\mathbf{v}_n^T\}$ is the measurement error covariance matrix, with \mathbf{v} the measurement noise vector. The observation matrix relates the measurement vector \mathbf{z}_n to the state vector, i.e. $\mathbf{z}_n = \mathbf{H}\mathbf{x}_n$. The next step of the measurement update is updating the estimate with the measurement:

$$\hat{\mathbf{x}}_n = \mathbf{x}_n^- + \mathbf{K}(\mathbf{z}_n - \mathbf{H}_n\hat{\mathbf{x}}_n^-). \quad (33)$$

Finally, the third and last step of the measurement update is updating the measurement uncertainty matrix

$$\mathbf{P}_n = (\mathbf{I} - \mathbf{K}_n\mathbf{H}_n)\mathbf{P}_n^-. \quad (34)$$

After both the time and measurement are updated, the process is repeated iteratively in order to predict and update the system state vectors for all time samples. By performing these iterations, both the measurement and process noise are reduced.

7) *Gait feature time-frequency representation*: The subsequent step involves performing the Doppler-FFT such that for multiple time instants, the Doppler spectrum is generated. Due to target tracking, the range bins in which the target resides are known. Consequently, relevant range bins are selected to analyze the Doppler frequency of the body. The advantage of having knowledge of the target's range bins is that the Doppler signatures of objects in other range bins are not visible in the Doppler spectrum. This results in a more pure Doppler spectrum in which the main contribution comes from the target of interest. To illustrate the change of (Doppler) frequency over time, the spectrogram is frequently used as a method of visualization. It is computed using the short-time Fourier transform (STFT). The STFT is a moving window Fourier transform [42], in which the Fourier transform is taken over the samples inside the window which consists of a short segment of a longer time signal. The spectrogram is expressed as

$$P(f_D, t) = |STFT\{y(t)\}|^2 \quad (35)$$

$$= \left| \int y(\tau)w(\tau - t) \exp(-j2\pi f_D\tau) d\tau \right|^2,$$

in which $y(t)$ is the analyzed signal and $w(t)$ the sliding window [31]. The time-frequency representation is generated for the slow-time signal $y(t)$ for every relevant range bin. This window is taken to be a Hamming window of length 256 with an overlap length of 194 and a FFT length of 256 as well. As the body size exceeds the range resolution, it occupies multiple range bins. Therefore, the spectrograms of multiple range bins wherein the body potentially resides are summed to obtain the overall spectrogram of the body. The drawback of the STFT is the limited resolution trade-off. A shorter time window creates a better time resolution, but results in a worse frequency resolution and vice versa.

Upon computing the spectrogram, the subsequent task is to extract a gait feature from it. Each body part exhibits a distinct micro-Doppler signature. As previously mentioned in Section II-D, separating Doppler profiles of different body parts within the spectrogram is challenging. Despite this, extracting the Doppler signature of the torso is relatively straightforward. This is primarily due to the torso's higher RCS compared to other body parts (e.g., limbs), resulting in the highest power Doppler frequency component representation. Additionally, the radar's approximate location is at torso height, causing the majority of the power to be directed towards the torso, given the directional antenna's orientation. To obtain the radial velocity (or Doppler frequency) of the torso from the measurements, the weighted mean of the spectrogram is taken [34]. The power-weighted average frequency for time sample t is given by

$$f_{torso}(t) = \frac{\sum_{k=1}^N f_D(k) \cdot P(k, t)}{\sum_{k=1}^N P(k, t)}, \quad (36)$$

in which N is the size of the FFT during the STFT operation. So, k represents the k th frequency component.

8) *Elevation estimation*: Analogous to the range-azimuth estimation, a range-elevation plot can be generated as well. Instead of using array elements in the horizontal direction, this is done using a vertically orientated ULA in the z -direction. As for the range-azimuth points, CFAR detection is executed on the range-elevation plot. The z -coordinate of a detected point with range r and elevation angle ϕ in a coordinate system with the radar in the origin is given by

$$z = r \cos(\phi). \quad (37)$$

B. Individualized walking model

The spatio-temporal walking model used in the pipeline is the well-known global human walking model presented by Boulic et al. [23]. This kinematic model is based on experimental biomechanical data. The model describes the spatio-temporal motions of 12 human body parts, depending on a normalized gait velocity v_g . The gait speed s_g is normalized by the leg length L_{leg} , which is the distance from the floor to the thigh in a static pose. The model assumes the walking motion is periodic, with phases having either double support of both legs or single support of one of the legs. One full gait cycle, also called a stride, is illustrated in Fig. 10. The cycle

is expressed by the relative time is a value between zero and one and is given by

$$t_{rel,l} \equiv t f_g + \phi_g \pmod{1} \quad (38)$$

and

$$t_{rel,r} \equiv t f_g + \phi_g + 0.5 \pmod{1}, \quad (39)$$

for the left limbs and the right limbs, respectively. Here, t is the time instant, f_g the step frequency and ϕ_g the cycle phase. The phase is a value between zero and one, hence the modulo 1 included in the formula. Note the relative time of the limbs on the right side of the body has an offset of 0.5 compared to the left side, corresponding to half a gait cycle.

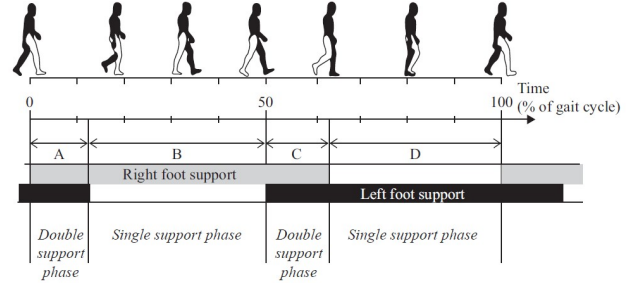


Fig. 10: Human walking gait cycle. (From [31]).

Positions of the body parts are computed relative to a human reference point. This is a point in the human body which has a certain translation and rotation with respect to the radar position. Translations and rotations of body parts are described as a function of the velocity and the relative cycle time. The rotations and translations with respect to the body can be seen in Fig. 11. Translations include vertical T_V , lateral T_L and horizontal T_H translations of the body around the human reference point. Rotations include forward/backward ϕ_{FB} , left/right ϕ_{LR} and torsion ϕ_{TO} rotations of the body around the human reference point. The whole body also has a certain thorax angle ϕ_T , representing the direction in which the human is moving. Furthermore, rotations of joints on the left and right side of the body are described. These joint rotations include flexing at the hip with angle ϕ_H , knee with angle ϕ_K , ankle with angle ϕ_A , shoulder with angle ϕ_S and elbow with angle ϕ_E . For example, the rotation of the shoulder in degrees is given by

$$\phi_S(t_{rel}, v_g) = -3 - 9.88v_g (0.5 + \cos(2\pi t_{rel})). \quad (40)$$

For all expressions of the rotations and translations of the body joints, the reader is referred to the original paper by Boulic et al. The human walking model was realized from scratch in MATLAB using the described expressions. An image of the skeleton of a walking human representing the model can be seen in Fig. 12, in which the 12 body parts are visible. The body parts are listed in Table I, together with their default length. Note the head is modelled as a sphere. The default dimensions are based on the anthropometric data described

in [43]. These sizes can be scaled using parameters extracted from the radar measurement data. In this work, this is done using the total human height as a scalar.

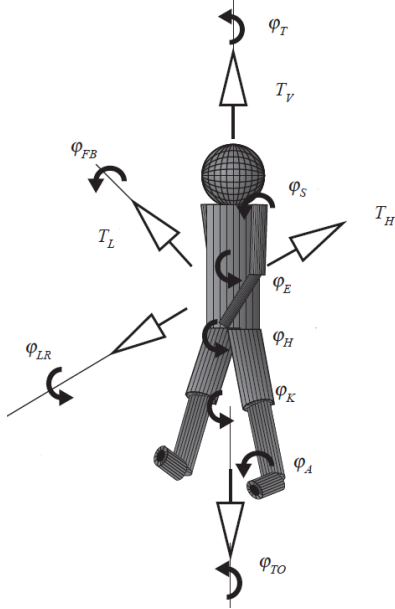


Fig. 11: Human with twelve body parts and its rotations and translations. (From [31]).

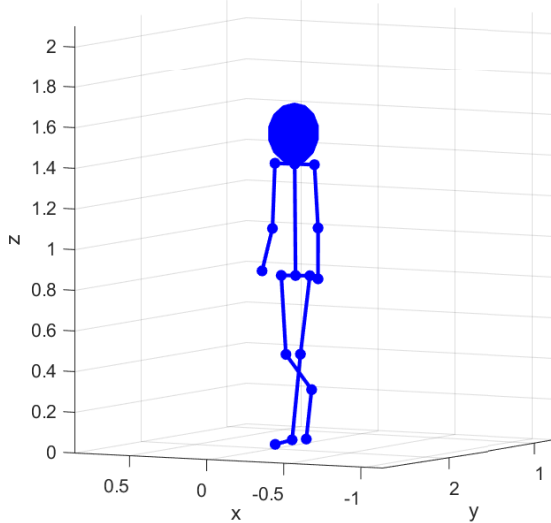


Fig. 12: Human model figure.

The general walking model is individualized using the gait parameters, the height parameter and the cycle phase. These are determined as described below.

1) *Gait parameter estimation*: Biomechanical walking parameters are extracted using information from both the estimated target trajectory and the micro-Doppler signature of the torso. The two gait parameters which are estimated and fed to

TABLE I: Default model dimensions.

Human element	Length [m]
Head	0.30
Torso	0.55
Upper legs	0.40
Lower legs	0.43
Feet	0.25
Upper arms	0.32
Lower arms	0.25

the walking model are the normalized gait speed and the step frequency (convertible to cadence). Note the step length can be computed from these two parameters.

As explained before in the post-processing pipeline, the Kalman-filtered trajectory is expressed by the system state vector \mathbf{x}_n for every frame. The system state vector contains the velocity in both the x and y direction. The gait velocity for frame n is derived from these values as

$$s_{g,n} = \sqrt{\dot{x}_n^2 + \dot{y}_n^2}. \quad (41)$$

Before it is fed to the model, the gait velocity needs to be normalized to a leg length of one meter such that the expressions in provided by the model are correct:

$$v_g = \frac{1}{L_{leg}} s_g. \quad (42)$$

The step frequency is directly determined using the torso Doppler signature obtained from the spectrogram. First, the moving average is subtracted from this signal such that the DC component is removed. Then, by taking the FFT of the resulting waveform, the step frequency is obtained by determining the highest peak in the frequency spectrum. Subsequently, the step length is readily computed using

$$l_g = \frac{v_g}{f_g}. \quad (43)$$

The walking model assumes a continuous motion. Therefore, not all combinations of the step frequency and velocity are realistic. If the parameters do not match to a certain extent, an unlife-like step length is computed and skidding of the feet will occur.

2) *Human height estimation*: The next step is to estimate the human body height from the radar data. To achieve this, the detected points in the range-elevation plot are used. A detected point with range r and elevation ϕ with respect to the radar is converted to the height h in Cartesian coordinates using

$$h = z + h_r = r \cos(\phi) + h_r, \quad (44)$$

with h_r the radar height with respect to the floor. After this conversion, an arbitrary number of scattering points are expressed within a certain height range. These points also include multipath reflections via the floor, the ceiling and the wall. These contributions are not of interest because these are not direct reflections from the human body to the radar. Therefore, these contributions will be omitted before making

a height estimate. This is done by specifying upper and lower limits for both the range and the elevation angle, with $r \in [R_l, R_u]$ and $\phi \in [\Phi_l, \Phi_u]$. These bounds are based on the estimated ranges of the scattering points in the x-y plane, obtained from the horizontal radar. For P detected points in range-azimuth, a range vector $\mathbf{r}_{p,az} \in \mathbb{R}^{P \times 1}$ and azimuth angle vector $\theta_p \in \mathbb{R}^{P \times 1}$ can be constructed which describe the detected points. The limits for the points in range-elevation are then given by

$$R_l = \min\{\mathbf{r}_{p,az}\} - 0.1, \quad (45)$$

$$R_u = \max\{\sqrt{\mathbf{r}_{p,az}^2 + (2.6 - h_r)^2}\}, \quad (46)$$

$$\Phi_l = 90^\circ - \sin^{-1}(-h_r/R_u), \quad (47)$$

$$\Phi_u = 90^\circ - \sin^{-1}((2.6 - h_r)/R_u). \quad (48)$$

The factor of 0.1 m in equation 45 is included to account for the distance between the two radars and for possible errors in range estimation. The maximum height which can be estimated is approximately 2.6 m, as shown in equation 46. This value is chosen because people are generally shorter than this length.

Due to the target's motion, the measured DOA ϕ of the backscattered EM waves exhibits fluctuations. These angle variations arise from phase changes caused by the target's movement [44] [27]. As previously discussed, the target comprises multiple scattering points, each with changing aspect angles, leading to fluctuations in the relative amplitude and phase of the received echoes at the radar receiver. Consequently, the amplitude and phase of the signals received by the antenna also experience fluctuations. As a result, the signal power of some scatterers may be too low to be detected by CFAR, and the estimated DOAs may contain glint noise components, resulting in inconsistent values across different frames. Due to these uncertainties, it is not always expected that the maximum value of the height corresponds to the actual human height. In some frames, the maximum value may be too high, while in others, it may be too low. To obtain a reliable estimate, the height is determined based on the maximum height values from multiple frames. Firstly, the height maximum of each frame is calculated. Subsequently, a moving average of these maxima over 10 frames is computed. The highest value of this moving average is taken as the estimate of the human height. The default dimensions in Table I are then scaled using the estimated human height to achieve a more accurate model representing the target's true physical dimensions.

3) *Cycle phase estimation*: The final component of the methodology involves estimating the cycle phase ϕ_g , which attains values between zero and one. This parameter is determined through a method based on [45]. At this stage of the pipeline, a model can be constructed with an arbitrary value of t_{rel} by selecting different values of ϕ_g , as indicated by equation 38. The model encompasses the trajectory, gait parameters, and human size parameters. Utilizing the model, the observed Doppler frequency can be computed over time using equation 22, given that the body parts' movements in 3D space are known with respect to the radar. These

movements are described by the generated spatio-temporal model, enabling to determine the expressions for the micro-Doppler signatures of all body parts. The scattering point of a body part is taken to be in its center. Subsequently, the Doppler frequency of the torso, as simulated by the model $f_{D,sim}(t)$, is fitted to the Doppler frequency of the torso obtained from measurements $f_{D,meas}(t)$. A fit function is generated, computing the mean squared absolute error between the measured and simulated Doppler frequencies over time with T time samples. Then, the best value for ϕ_g is selected corresponding to the minimum value of the fit function. This is described mathematically as

$$\phi_{g,fit} = \arg \min \frac{1}{T} \sum |f_{D,meas}(t) - f_{D,sim}(t, \phi_g)|^2. \quad (49)$$

IV. MEASUREMENT APPROACH

A. Measurement scenario

Measurements were conducted to evaluate the performance of the presented pipeline. The measurement scenario is illustrated in Fig. 13. In the actual setup shown in Fig. 14, the trajectories were marked on the floor using tape, with blue tape indicating distances of 0.5 m. The radar was positioned on a table at a height of 1 meter. Three volunteers participated in the experiment, walking along five different predetermined trajectories. Each trajectory had a distinct angle with respect to the antenna beam. The volunteers moved outward with respect to the radar and repeated this motion five times for each angle, walking at their self-chosen pace, step length, and step frequency. Upon reaching the end point, the target remained stationary until the radar measurement was completed. The walking angles indicated in Fig. 13 refer to the angles with respect to the y-axis, representing the broadside beam direction of the radar. However, the aspect angle varies depending on the human's location in relation to the radar. The aspect angle denotes the angle between the vector from the radar to the target and the vector in the direction of the target's movement. For each trajectory, the aspect angles are depicted as a function of the covered distance in Fig. 15.

B. Equipment and experimental setup

The measurements were performed using two Texas Instruments (TI) mmWave radars (two IWR1443 with IWR1443BOOST EVM board) [46]. Each MIMO radar consists of 2 TX elements and 4 RX elements, realizing two virtual ULAs, each of 8 elements. One board is solely used for azimuth estimation, while the other is solely used for elevation estimation. Both boards operate simultaneously, with the total available bandwidth divided over the two boards. The radar boards are aligned and connected rigidly using two metal plates, as depicted in Fig. 16. The vertical displacement between the antennas of the two boards is 7.2 cm. Having two boards, from which one is 90° tilted with respect to the other, allows the target to be measured in both azimuth and elevation with a good angular resolution. This comes at the cost of an

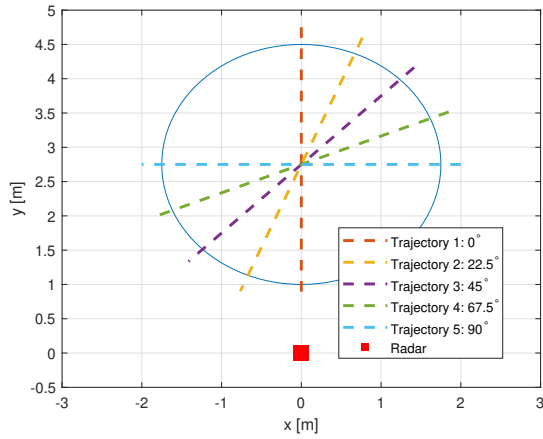


Fig. 13: Measurement scenario.

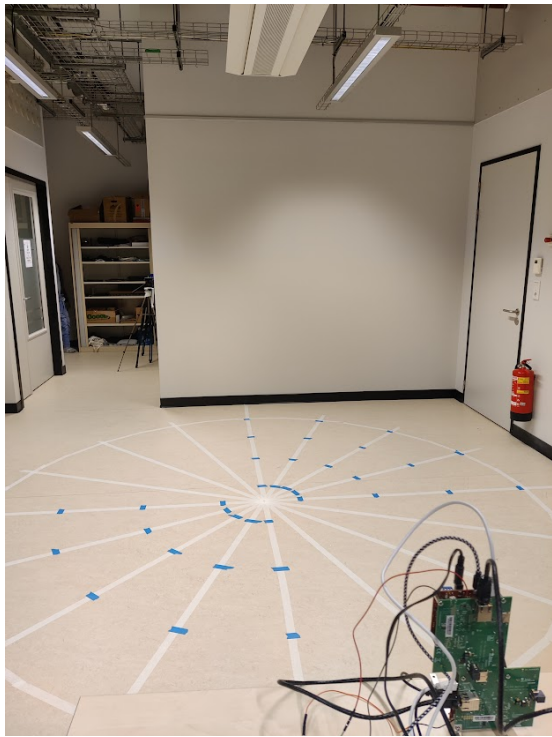


Fig. 14: Realized setup.

increase in hardware and an increase of captured data which needs to be stored.

The chirp emitted by the TI radar board and its configuration parameters is visualized in Fig. 17. The chirp is configured using the programmed parameters in Table II. The chirp configuration results in the properties listed in Table III. Several design choices were considered while configuring the chirp. The range resolution should be maximized, such that the human body scatterer locations can be localized as accurate as possible. Therefore, the sweep bandwidth should be as high as possible. In order to realize this, the available bandwidth should be divided efficiently over the two boards. The two

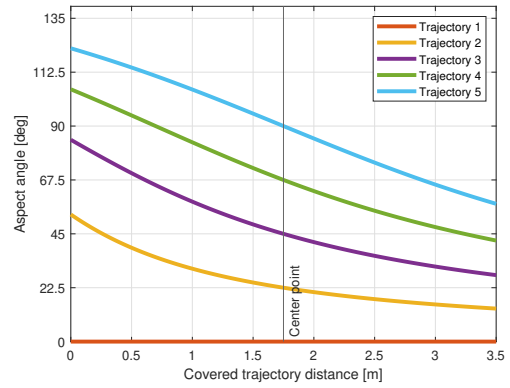


Fig. 15: Aspect angles as a function of the covered distance.

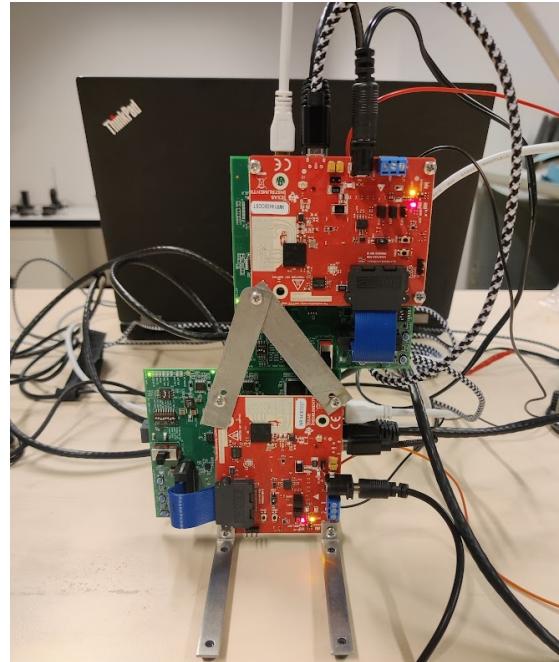


Fig. 16: Radar boards setup.

radars used different start frequencies to prevent interference. The frequency separation (500 MHz) is high enough to not interfere with each other. Using Equation 1, the computed range corresponding to the intermediate frequency between the two radar boards is over 3 km, which is much higher than the maximum unambiguous range. On top of that, such a high IF will be low-pass filtered after the quadrature mixer. Therefore, the radars do not interfere with each other. The resulting sweep bandwidth is 3001.6 MHz, creating a range resolution of approximately 5 cm with a maximum unambiguous range of over 10 m. The maximum range is considered sufficient since the application is indoor, and the dimensions of the room in which measurements were performed are smaller than 10 m.

A trade-off had to be made between the number of antenna elements and the velocity parameters. The aim is to measure the torso velocity of a walking person. Therefore, the maxi-

imum unambiguous velocity should not be too low, otherwise the velocity of a fast-walking person cannot be measured. The chosen chirp parameter results in a maximum unambiguous velocity of approximately 3.2 m/s. This is enough to capture the macro-Doppler gait velocity with micro-Doppler velocity variations of the torso. It is desired to have a virtual array which is as large as possible, such that the angular resolution is as good as possible. However, the number of feasible elements is limited by the velocity parameters due to TDM. Two virtual ULAs are realized, of which one array is horizontally orientated and the other is vertically orientated, enabling estimation of the azimuth and elevation angles, respectively. Each board uses four RX elements and two TX elements, creating a virtual ULA of eight elements.

The measurement trajectory distance is around 3.5 m, so a measurement time of 7 seconds was chosen. This is enough to capture the entire motion on every repetition. The periodicity is 0.1 seconds, so 70 frames were recorded.

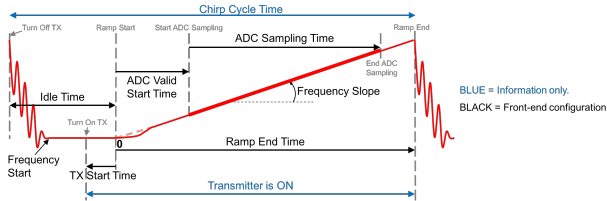


Fig. 17: Chirp diagram and chirps parameters (From [47]).

TABLE II: Programmed chirp parameters.

Parameter	Value
Start frequency	Radar 1: 77.00 GHz Radar 2: 77.50 GHz
Frequency slope	23.45 MHz/ μ s
ADC sampling rate	2 Ms/s
Number of ADC samples	256
Number of chirp loops	255
Idle time	7 μ s
ADC valid start time	6.4 μ s
Ramp end time	140 μ s
Frame periodicity	100 ms
Number of frames	70
Number of RX antennas per radar	4
Number of TX antennas per radar	2

Synchronous triggering of the two radar boards is crucial, such that the frames of data are aligned. The two radar boards are controlled by a Cmod S7 FPGA board [48]. In order to trigger a frame on the IWR1443, the FPGA gives a pulse to a specific pin. The active frame time matches the frame periodicity of the chirp configuration, and the active trigger time is chosen to be at least 25 ns, as indicated in the user’s guide of the device [46]. The trigger moment of the two boards should match exactly, such that no synchronization errors occur. Therefore, the physical length of trigger signal wiring is taken to be equal for both. The two radar boards are each connected to a PC using several provided cables.

TABLE III: Calculated chirp parameters.

Parameter	Value
Maximum unambiguous range	10.235 m
Range resolution	0.050 m
Maximum unambiguous velocity	3.243 m/s
Velocity resolution	0.025 m/s
Azimuth angular resolution	14.3°
Elevation angular resolution	14.3°
Sampling window length (chirp time)	128 μ s
Maximum beat frequency	1.6 MHz
Radar cube size	2040 kB
Valid sweep bandwidth	3001.6 MHz
End frequency	Radar 1: 80.28 GHz Radar 2: 80.78 GHz

Capturing and saving measurement data is done using TI’s mmWave Studio software.

C. Groundtruth capturing

A groundtruth of the gait parameters is required to validate the measurements. The groundtruth data of the step length was created by fastening an erasable marker on the shoe of the volunteer, leaving a dot on the ground after every step. After walking the designated trajectories, the stride lengths (i.e. distance between consecutive dots) were measured using a measuring tape. To measure the distance, the dots were projected onto the straight white trajectory line, and the distance from the starting point towards the projected point was measured. By measuring in this manner, the step width is omitted from the measurement, and the distance solely consists of the step length. The marked dots have a certain mean and standard deviation (SD) with respect to the initial location of the individual (starting point of trajectory). Consequently, the step lengths can be described by a mean and standard deviation. Moreover, the groundtruth of the step frequency was captured using the inertial measurement unit (IMU) of a smartphone, which was attached to the ankle of the volunteer. The smartphone device is an OnePlus 8T, which contains a Bosch BMI260 series IMU [49], using a sampling rate of 200 S/s. The IMU measures the linear acceleration of the ankle, from which its frequency corresponds to the step frequency. Again, there is a certain variability and mean over the five repetitions in which the step frequency will be expressed. The ground truth velocity was computed by taking the product of the frequency and the step length (Eq. 43). For the step lengths the mean and variability is known, but no information was saved on which dot belongs to which repetition. Therefore, the standard deviations of the ground truth velocities could not be computed. Thus, the gait velocity will only be expressed by a mean value.

The groundtruth properties of the three volunteers are listed in Table IV. The distribution of the parameter values is assumed to be normal. Person 1 has a similar step length compared to person 2, but with a slightly lower total standard deviation. Person 3 has a significantly longer step length. Furthermore, person 1 and person 3 have comparable mean

TABLE IV: Ground truth parameters of the volunteers.

Person	Trajectory	Velocity mean [m/s]	Step frequency mean \pm SD [Hz]	Step length mean \pm SD [cm]
1	1	0.779	1.462 \pm 0.061	53.300 \pm 4.735
	2	0.840	1.464 \pm 0.020	57.380 \pm 2.693
	3	0.885	1.520 \pm 0.022	58.187 \pm 1.620
	4	0.832	1.498 \pm 0.023	55.510 \pm 2.150
	5	0.816	1.554 \pm 0.025	52.520 \pm 2.356
	Average of 5 trajectories	0.830	1.501 \pm 0.046	53.380 \pm 3.603
2	1	0.990	1.917 \pm 0.049	51.627 \pm 3.729
	2	1.029	1.946 \pm 0.024	52.857 \pm 3.480
	3	1.074	1.965 \pm 0.025	54.643 \pm 4.277
	4	1.036	1.922 \pm 0.024	53.910 \pm 4.708
	5	1.018	1.905 \pm 0.019	53.443 \pm 2.605
	Average of 5 trajectories	1.029	1.931 \pm 0.035	53.296 \pm 5.303
3	1	0.984	1.432 \pm 0.070	68.710 \pm 2.950
	2	1.030	1.511 \pm 0.052	68.175 \pm 4.274
	3	1.050	1.579 \pm 0.026	66.520 \pm 2.148
	4	1.074	1.613 \pm 0.035	66.600 \pm 2.137
	5	1.129	1.630 \pm 0.030	69.245 \pm 3.329
	Average of 5 trajectories	1.056	1.558 \pm 0.082	67.850 \pm 4.227

step frequencies, with person 3 having almost double the total standard deviation in step frequency. Person 2 walks with a higher cadence compared to the others. The average speeds of person 2 and person 3 are similar, while person 1 walks slower. The groundtruth heights of the volunteers are 1.82 m, 1.61 m and 2.01 m, respectively.

V. RESULTS AND DISCUSSION

A. Experimental pipeline validation

First, a validation process is conducted to ensure the functionality of various crucial components within the proposed pipeline. These components are the trajectory estimation, step frequency estimation, height estimation and the estimation of the cycle phase.

1) *Trajectory estimation*: Fig. 18 illustrates five instances of localization samples before and after Kalman filtering for various walking angles. The observed positions of the trajectories are prone to errors, but the implementation of the Kalman filter reduces these errors, resulting in continuous trajectories without abrupt outliers. It should be noted that only a fraction of the initial trajectory was extracted, excluding the acceleration and deceleration phases. The resultant trajectories align closely with the expected walking paths (indicated by striped lines), validating the goal of the approach. The filtered trajectory still has a certain deviation from the straight line. Note that the participants do not precisely follow the striped line either, since it was an indicative line to follow rather than a strict trajectory. It is impossible to walk straightly without having minor deviations from the line.

2) *Step frequency estimation*: For the frequency estimation, the initial step involved computing the spectrogram of the extracted trajectory, from which the step frequency was then estimated by examining this Doppler signature. Fig. 19 displays the spectrograms for various measurements taken at different angles, along with the corresponding Fourier transforms of the Doppler signatures. Also, a spectrogram without clutter removal is added and shown in Fig. 19a. The spectrograms with clutter removal show low backscattering at zero Doppler frequency even though there were multiple cluttering objects in the room, such as tables and lab equipment. This means the static clutter removal works, and it is further supported by the influential zero-Doppler component evident in Fig. 19a, which impacts the weighted mean.

An observable sinusoidal pattern is evident in the spectrograms depicted in Figs. 19b and 19d. This oscillating behavior represents the cyclic micro-Doppler gait movement, wherein the body undergoes slight back-and-forth motion, thereby introducing an additional Doppler shift on top of the macro-Doppler shift resulting from the individual's coarse velocity. The step frequency, as depicted in Figs. 19c and 19e as the main frequency component, corresponds to the micro-Doppler frequency of the torso. Notably, when the aspect angle between the radar and the target is low (approaching 0°), a distinct Doppler signature is observed due to the high radial velocity detected by the radar. However, with an increase in aspect angle, both the radial velocity component and the Doppler frequency decrease. As shown in Figs. 19f, 19h, and 19j, the spectrograms indicate a weighted mean approaching the zero Doppler line when the aspect angle approaches 90° . Nevertheless, the micro-Doppler frequency

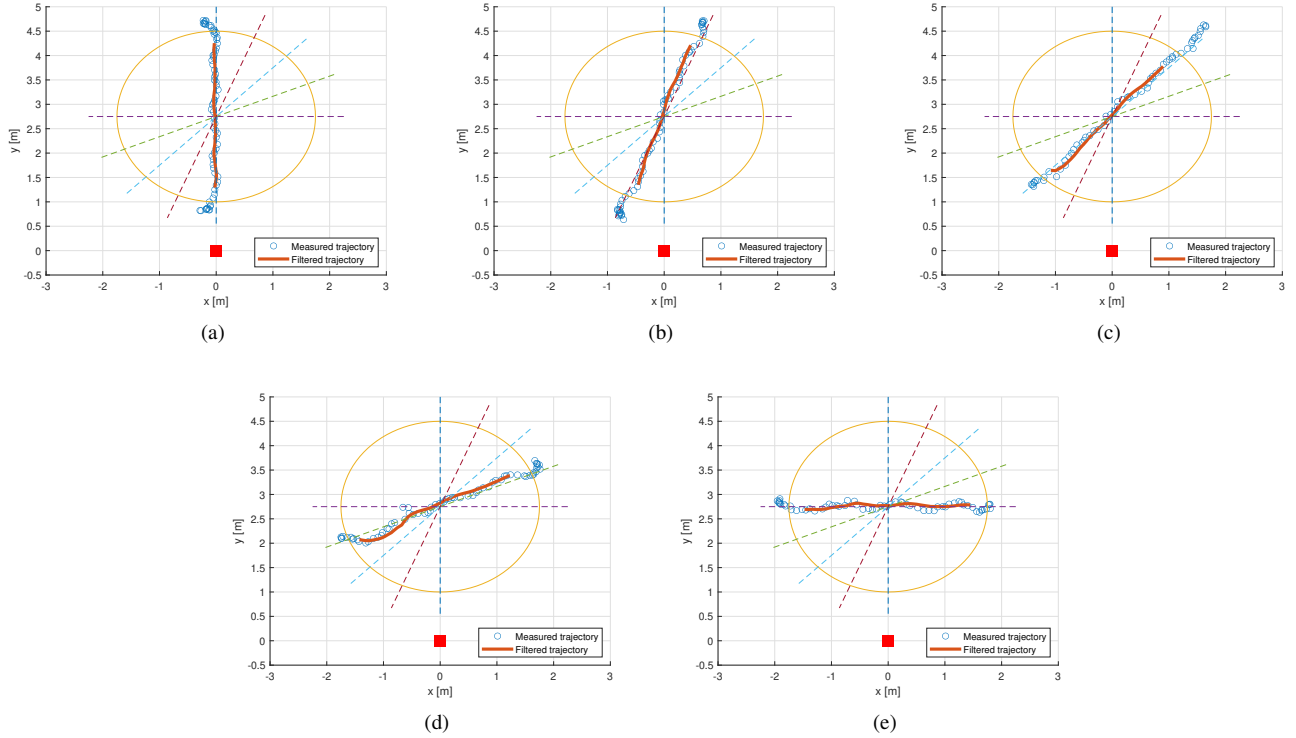


Fig. 18: Examples of measured trajectories and filtered trajectories of person 1. (a) Trajectory 1, (b) Trajectory 2, (c) Trajectory 3, (d) Trajectory 4 and (e) Trajectory 5.

swing remains present due to the radar beam never being continuously perpendicular to the movement normal vector. As such, the step frequency can still be derived from the frequency spectra displayed in Figs. 19g, 19i, and 19k. As the angle approaches 90° (lower plots), the prominence of the stride frequency becomes more pronounced, and the step frequency becomes less noticeable. The stride frequency, which is half the step frequency, characterizes the frequency of an entire gait cycle involving a step with the right leg and a step with the left leg. The frequency of movement in the lateral plane is half that of the movement frequency observed in the sagittal plane. Consequently, the spectrogram exhibits lateral plane motion in contrast to the previous sagittal plane motion.

3) *Height estimation:* The computation of the heights of detected points in the range-elevation plot was carried out following the previously described methodology. Given that the volunteers exhibit varying body heights, it is expected that the heights of detected points would also vary accordingly. In Fig. 20, histograms depicting the body heights of three individuals walking in trajectory 1 are presented. The histograms correspond to human height dimensions of 1.82 m, 1.61 m, and 2.01 m for the three persons, respectively. The lower part of the histogram does not display any detected points for all three individuals. This absence of detection is likely attributed to the fact that the feet have a low RCS and may move out of the range limits set for detection due to their continuous motion during walking. Moreover, the limited field of view

(FOV) of the radar in the elevation plane results in lower received power from the feet. Another property for all three graphs which stands out is the high backscattering at the radar height, which is approximately 1.07 m. This is due to the high RCS of the torso and the antenna gain being the highest at an elevation angle of 90° .

The height points corresponding to person 1 are presented in Fig. 21 as a function of the frame number, where low frame numbers indicate the beginning of the trajectory and high frame numbers indicate the end of the trajectory. Fig. 21a displays the outcome for trajectory 1, wherein the human walks with an aspect angle of 0° in front of the radar. At the start and the end of the trajectory, the number of detected points on head height is lower compared to the middle part. The difference in detected points can be attributed to the limited FOV at the beginning of the trajectory, leading to decreased backscattering when the target is in close proximity to the radar. For the end of the trajectory, the low amount of detected points on head height can be attributed to the lower backscattering power due to increased path loss. For trajectory 5 (Fig. 21b), the first and last parts of the trajectory show no detected points at all. The cause of this is again the limited FOV, but now in the horizontal plane. At the start and end of the target's trajectory, the azimuth angle θ for the vertically orientated array is too high (or too low) to be able to realize sufficient backscattered power for a successful detection.

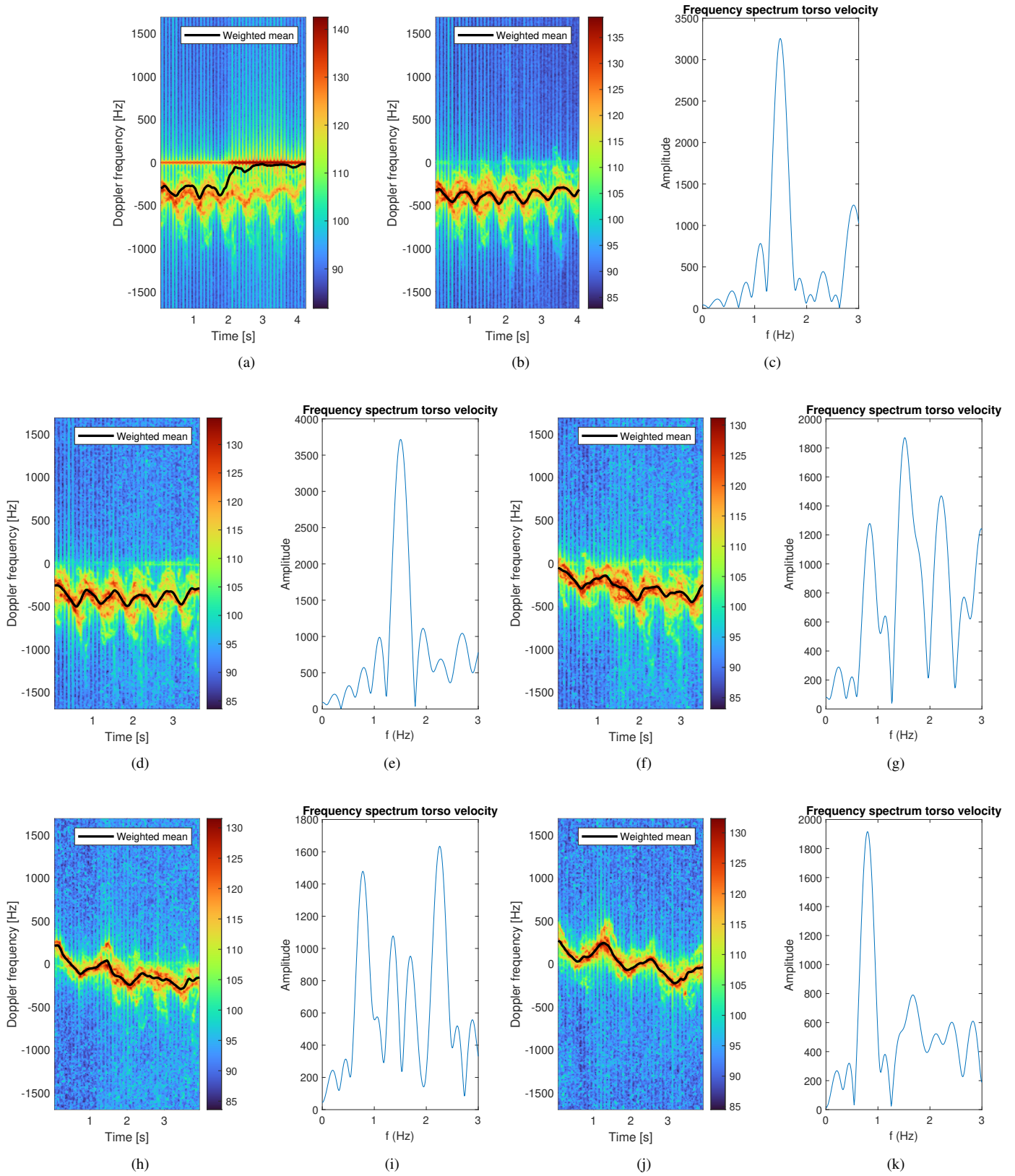


Fig. 19: Spectrograms (a) without and (b,d,f,h,j) with static clutter removal of torso Doppler signatures of person 1. Fourier transforms of weighted mean of corresponding spectrograms are depicted in (c,e,g,i,k). The figures correspond to measurements for incrementing angles of 0° , 22.5° , 45° , 67.5° and 90° , respectively.

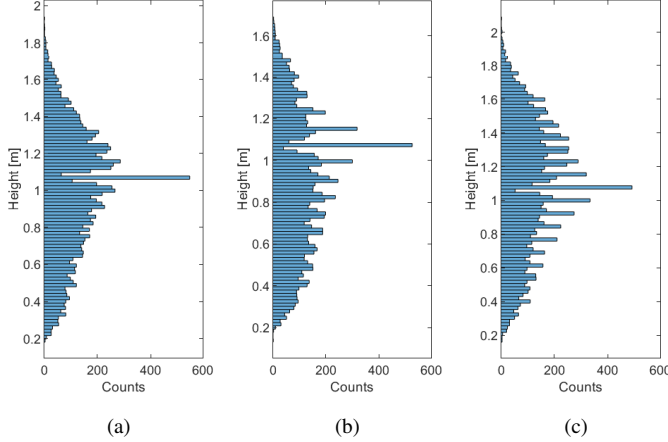


Fig. 20: Examples of histograms of heights aggregated over all frames for (a) person 1, (b) person 2 and (c) person 3 walking on trajectory 1.

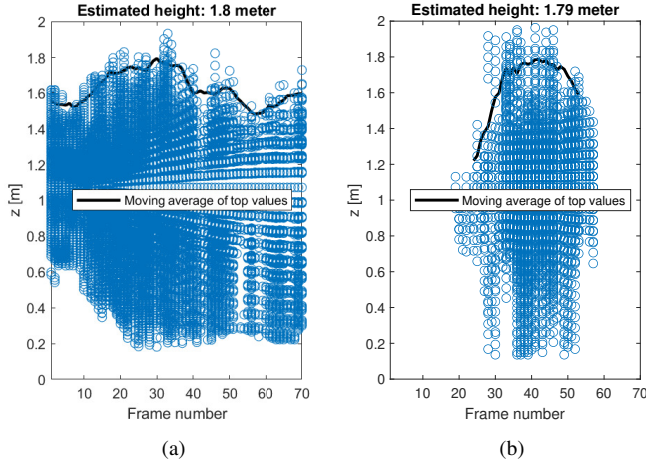


Fig. 21: Detected points of person 1 and the corresponding estimated height versus the frames for (a) trajectory 1 and (b) trajectory 5.

4) *Cycle phase estimation:* The cycle phase is estimated after computing the mean squared error between the simulated and measured Doppler frequencies of the human torso. The best fit and the measured Doppler frequency are shown in Fig. 22 for trajectory 1, trajectory 3 and trajectory 5, with measurements performed on person 1. Across all three trajectories, the measured Doppler frequency exhibits a consistent temporal behavior when compared to the Doppler frequency obtained from the fitted simulation. This confirms the velocity, step frequency and cycle phase estimations give a well-approximated simulated Doppler signature. The achieved accuracy is considered sufficient for estimating the cycle phase utilizing the model-fed gait parameters. This claim is based on the observation that the shapes of the frequency components

of both the simulated and measured Doppler curves remain consistent across all presented results.

B. Experimental results of performance

1) *Gait parameter estimation:* The gait parameters were estimated by applying the proposed algorithm on the experimental data. The computed values were compared to the ground truth data of the corresponding trajectory, ranging from number 1 to 5. The resulting estimation errors for velocity, step frequency and step length are presented in Figs. 23, 24 and 25, respectively.

Fig. 23 shows that the velocity was assessed with an overall error margin of less than 0.17 m/s for all single repetitions (red circles). The accuracy of the velocity estimation is not evidently dependent on the angle. The estimation for persons 1 and 2 was carried out with an overall maximum error of 0.12 m/s. So, person 3 displayed a relatively higher error with more spread compared to other participants. This is due to the fact that person 3 has a taller body with longer legs, resulting in long step lengths. Consequently, this individual was unable to fit as many steps within a single trajectory as the others before hitting the wall due to the confined space. As a result, this person had to end the walk earlier, thus covered a shorter distance. Due to the reduced covered distance, the accuracy of velocity estimation was lowered since the performance of the Kalman filter is dependent upon the number of available samples. In other words, the distance between acceleration (start of the trajectory) and deceleration (end of the trajectory) is shorter for person 3. Observed over five repetitions, the velocity is estimated with an error lower than 0.077 m/s for all angles for the three volunteers.

The depicted results in Fig. 24 generally show an increase in the error of step frequency estimation as the angle relative to the radar increases. This trend is observed across all three individuals. For person 1 and 2, the step frequency at angle 0° , 22.5° and 45° could be estimated with an error of less than 0.066 Hz. However, for the last two trajectories which have higher aspect angles, the step frequency error increases. This is in line with the results shown in Fig. 19, because here it is visible that the step frequency component is harder to derive for higher aspect angles as well. Extreme values of errors for person 1 and 2 go up to 0.175 Hz and 0.204 Hz, respectively. This shows that for some repetitions, the spectrogram is changed in such a way that the step frequency is hard to extract. Nevertheless, the mean error can be reduced by taking the mean over the five repetitions (blue asterisks), since the extreme outliers are averaged out. The step frequency error of person 3 is below 0.063 Hz at an angle of 0° or 22.5° and increases towards a maximum of 0.177 Hz for higher aspect angles. Overall, the error shows similar behaviour as a function of the angle compared to person 1 and 2. However, a remarkable outlier is visible at the second trajectory (22.5°). This error was caused by an inaccurate trajectory estimate in the x-y plane. Therefore, the selected range bins of the spectrogram were not the range bins where the target actually resided. Thus, a part of the spectrogram was generated for

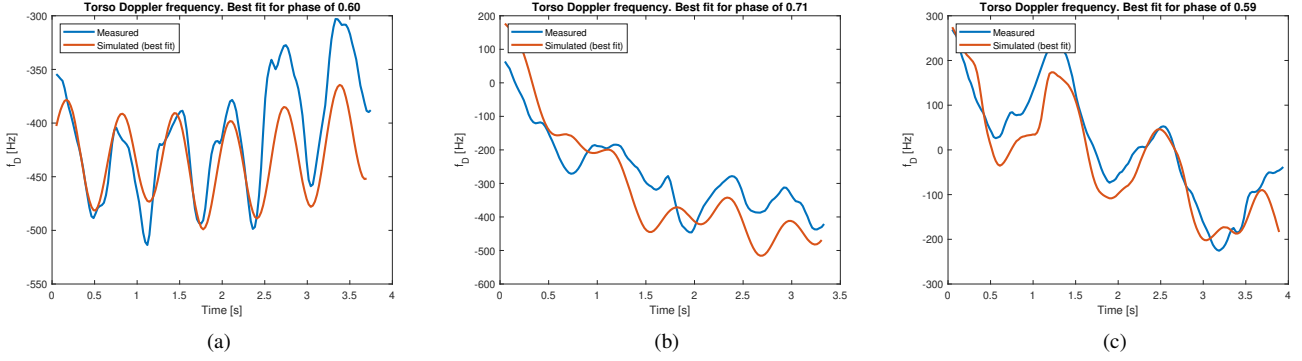


Fig. 22: Measured and best fit simulated Doppler frequency for (a) trajectory 1, (b) trajectory 3 and (c) trajectory 5. Measurements performed on person 1.

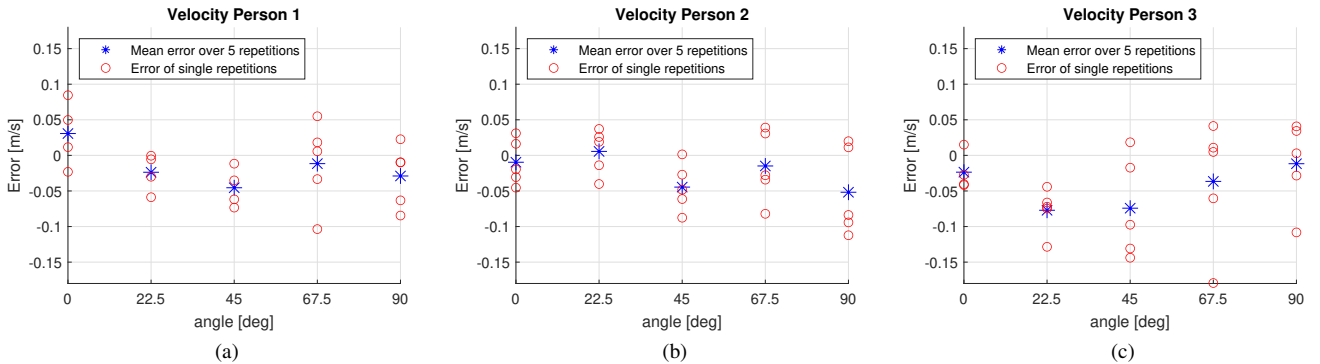


Fig. 23: Velocity estimation error values compared to ground truth for (a) person 1, (b) person 2 and (c) person 3.

range bins without any human Doppler signatures, and the estimated step frequency is inaccurate.

As described in the methodology in section III, the step length was determined from the radar data by dividing the gait velocity by the step frequency. Consequently, errors which appear in these parameters will propagate through to the estimated step length parameter as well. This effect is visible in the data shown in Fig. 25. When evaluating the step length error for the first three trajectories for person 1 and 2, respectively, the step length error is below 6.2 cm and 5.2 cm for all repetitions. For the last two walking angles, the errors in step frequency are relatively high and these errors reappear in the step length error as well. For these two walking angles, person 1 and 2 have a maximum error of 9.4 cm and 9.7 cm, respectively. Moreover, for person 3, the step length errors are higher compared to the other two individuals. The cause of this is the relatively high error in both velocity and step frequency, which again influences the step length estimation. Only for the angle of 0° the error is low because the velocity and step frequency errors are small here as well. The mean error of the step length for all three individuals averages out the outlying errors, making the mean error approach zero. The mean step length errors are maximally 5.0 cm, 4.3 cm and 9.3 cm, respectively.

Thus, to make the gait estimation more accurate, outliers with higher errors can be averaged out if the trajectory is long enough. In order to show this effect, the gait parameters are estimated over the five measurements instead of over one measurement. The results are listed in table V. The errors listed in this table were also plotted in Figs. 23, 24 and 25.

2) *Human height estimation:* The height of the participants was estimated using the described method. The results are shown in Fig. 26. For person 1, the height is estimated relatively accurately compared to the other persons, especially at low aspect angles. For this specific individual, most of the estimated heights have an absolute error below 8.2 cm. Nevertheless, there are a number of extreme cases (4 out of 25) in which the errors are higher than 14 cm. One possible explanation for occurring errors is the limited resolution in both range and angle. The range resolution is 5 cm, while the angular resolution is 14.3° . The resolution in the Cartesian plane depends on the resolution of the polar plane. This follows from equation 37. With increasing range, the area of range-angle cells in the Cartesian plane increase. Thus, resolution is reduced since more points in the plane are assigned to a single combination of r and ϕ . On top of that, the target is moving, so the scattering points are not stationary. So, the elevation angles of the scattering points on top of the head

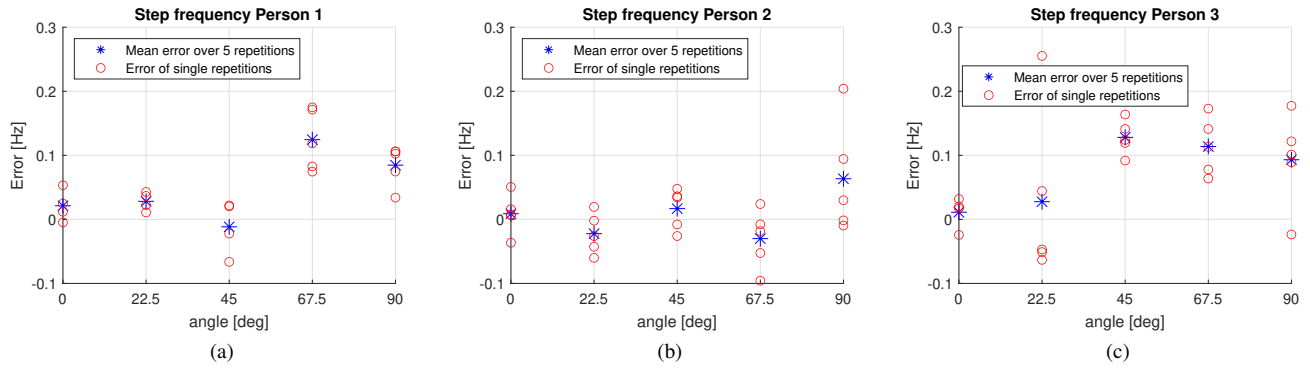


Fig. 24: Step frequency estimation error values compared to ground truth for (a) person 1, (b) person 2 and (c) person 3.

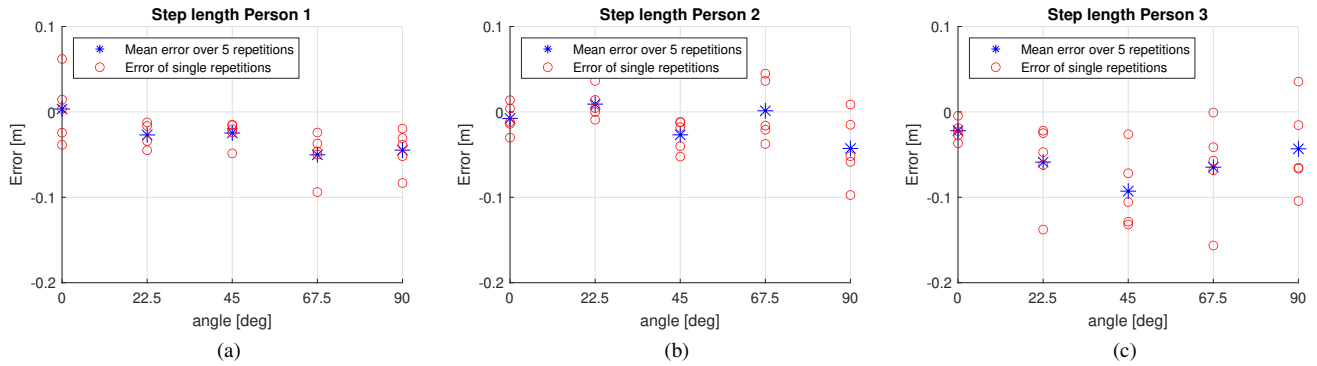


Fig. 25: Step length estimation error values compared to ground truth for (a) person 1, (b) person 2 and (c) person 3.

TABLE V: Mean gait parameter estimation for different angles and persons and errors with respect to ground truth.

Person	Trajectory	Gait speed estimation [m/s]	Gait speed error [m/s]	Step frequency estimation [Hz]	Step frequency error [Hz]	Step length estimation [m]	Step length error [m]
1	1	0.7990	0.0196	1.4921	0.0298	0.5364	0.0034
	2	0.8172	-0.0230	1.4936	0.0294	0.5469	-0.0269
	3	0.8427	-0.0419	1.5122	-0.0081	0.5572	-0.0247
	4	0.8200	-0.0116	1.6226	0.1245	0.5049	-0.0502
	5	0.7872	-0.0290	1.6387	0.0847	0.4804	-0.0448
2	1	0.9897	-0.0097	1.9171	0.0088	0.5087	-0.0076
	2	1.0287	0.0056	1.9462	-0.0223	0.5377	0.0091
	3	1.0739	-0.0446	1.9653	0.0168	0.5196	-0.0269
	4	1.0361	-0.0149	1.9218	-0.0301	0.5404	0.0013
	5	1.0179	-0.0518	1.9047	0.0634	0.4917	-0.0428
3	1	0.9619	-0.0222	1.4459	0.0137	0.6651	-0.0220
	2	0.9528	-0.0771	1.5383	0.0276	0.6230	-0.0588
	3	0.9761	-0.0742	1.7070	0.1280	0.5724	-0.0928
	4	1.0379	-0.0366	1.7210	0.1138	0.6013	-0.0647
	5	1.1170	-0.0117	1.7231	0.0931	0.6493	-0.0432

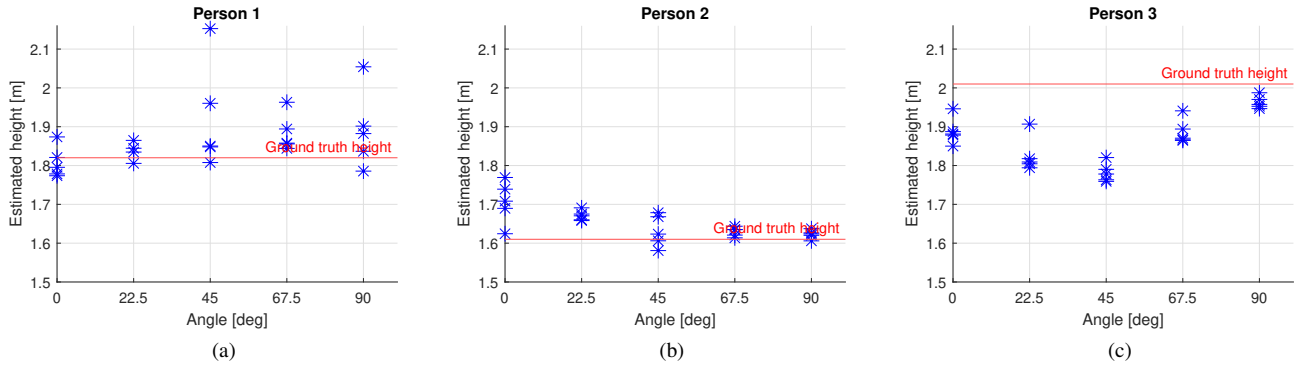


Fig. 26: Height estimations for (a) person 1, (b) person 2 and (c) person 3.

are not constant over time. Angular noise is introduced due to the changing phases of the echoes arriving at the receiver.

For person 2, the height is generally overestimated, especially at low aspect angles. The errors are maximally 15.9, 8.1, 6.9, 3.4 and 2.8 cm for the five angle values, respectively. So, there is an evident decline in height estimation error when the angle gets higher. The amount of points detected at and above head height is higher for low aspect angles. A possible explanation for this might be the increased backscattering power due to a higher RCS because of the shape of the head. Then, the CFAR detector detects more points around the actual scattering center. Compared to the other two persons, person 2 has a lower incident angle, which enhances this effect.

The height estimation of person 3 is always underestimated, with errors appearing to be even more than 20 cm for multiple measurement repetitions. There are a number of possible reasons for having such high error values. To start with, the FOV of the vertically-orientated radar is limited to a 3dB-beamwidth of ± 28 in the elevation plane [46]. So, the received power from high elevation angles corresponding to the location of the head is relatively low compared to the power received from boresight. Moreover, the RCS of the top of the head is low due to its shape. Incoming waves will be reflected to many different directions other than the direction where the radar is located due to diffuse reflection. Also, the backscattering from the head is low due to a high angle of incidence on the head. Contrary, the torso has a low angle of incidence and a smoother surface, in which the waves are reflected specularly to the radar.

The mean absolute error (MAE) of the height estimation is shown in Fig. 27. Evidently, the overall height estimation accuracy of person 3 is the worst. As explained before, this volunteer is too tall for the system to make an accurate height estimate for a number of reasons. Person 1 and person 2 show similar accuracy in MAE. However, for person 1 the error is higher for high angles, while for person 2 the error is higher for low aspect angles. Overall, the MAE for persons 1 and 2 stays below 10.9 cm.

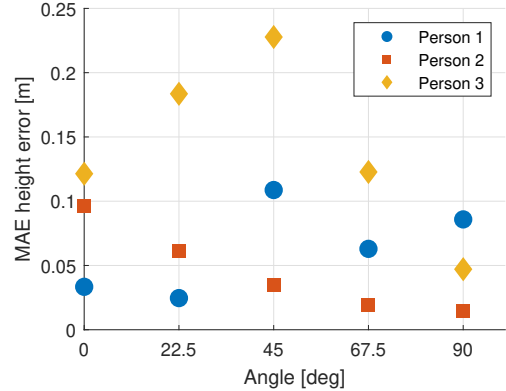


Fig. 27: Mean absolute error of height estimation.

C. Comparison with current state-of-the-art

To evaluate the outcome of this research, it is compared to previous (state-of-the-art) papers. The performance of the gait parameter estimation algorithm in this work is compared to a number of references is listed in Table VI. The listed parameters are the mean parameters taken over a certain measurement time (and covered distance) which were reported in the relevant work. The current work has errors in gait parameters which are in the same order of magnitude as previous work. It outperforms previous work on velocity parameter estimation when the target is walking in a straight trajectory towards the radar. The step length and step frequency errors in [8] are relatively low compared to other work, including this research. This is because the measurement scenario differs a lot from the scenario from the others; the target was walking on a treadmill for two consecutive minutes. Therefore, a lot of data is captured because the walk takes a long time. If the mean gait parameters are then computed, errors will be averaged out over an extensive time and the results approach the ground truth values. In other works, the walking distance varied between 5.2 and 10 m, which is much shorter, resulting in less accurate estimates. In the current research, the walking distance was short as well (approximately 3 m without acceleration and

deceleration phases). The error values in the table are based on 5 repetitions, so the total walking distance is about 15 m. No previous work exists in which the gait parameter estimation performance is analyzed as a function of the aspect angle. So, the error for high aspect angles in this research can not be compared to similar outcomes of previous work. Nevertheless, it can be noted the velocity and step frequency for trajectories with high aspect angles is deteriorated with respect to low aspect angles. However, the gait parameters could be estimated with similar accuracy compared to results in earlier work for an aspect angle of 0° .

Previous work on body dimension estimation or human keypoint estimation mainly rely on the use of ML techniques. The performance of the gait parameter estimation algorithm in this work is compared to a number of references listed in Table VII. A NN is a much more powerful tool compared to the technique used in this work. It has lower height estimation errors compared to the approach presented here. However, it takes a lot of training data comprising different persons and environments. The method presented here can be applied to arbitrary radar data, but requires refinement. Its performance can be improved if both the FOV and the resolution of the radar are increased. Moreover, the method based on the empirical formula stated in [22] is far from accurate and should not be used.

D. Limitations

This study has several limitations. The first limitation is the number of participants, repetitions and trajectories. If measurements had been performed on a larger amount of people with more repetitions, more data would be available. This would have a positive effect on the reliability of the research. Now, conclusions have to be drawn based on only three participants performing three repetitions for five different trajectories. Additional participants and trajectories would be useful to verify the pipeline with more supportive data. Greater participant variability would lead to a wider range of step lengths, gait speeds, and body dimension parameters being introduced. In addition, the current work was limited to measuring one participant in the room. However, in realistic scenario's, multiple people could be present in the room. Furthermore, if large amounts of data had been captured, this could have been used to develop and train a NN which could capture the dimensions of multiple human body part. It should be stated that an attempt had been made to infer human key points from radar data using the open-source model proposed in [20]. However, the model was only trained on two persons in a fixed environment and the input requires being a point cloud. Hence, with the data from the current work, no accurate key point estimations could be made.

Secondly, the room in which the measurements were performed had limited space. Therefore, only a short trajectory could be measured. On one side, this is realistic since rooms where the system can potentially be deployed do not always have to be big. On the other side, the gait estimation accuracy is limited because the gait is measured for only a short

period of time. A more reliable estimate could be made if the individual is measured for a long period. This has been confirmed by earlier work.

Thirdly, the hardware used is limiting the ability to estimate the human body dimensions. The radars used were two separate FMCW radars, both having a virtual ULA of eight elements. The radars work independently and are only able to estimate the angle in one plane (i.e. either azimuth plane or elevation plane). If a more sophisticated virtual array with more elements had been used, a two-dimensional virtual array could be realized. For example, if the resources would allow, this virtual array could have been 8×8 elements (comparable to the one used in [19]) instead of two arrays of 8×1 . Then, the location of the every detected point could be estimated in both azimuth and elevation because beamforming can be realized for θ and ϕ instead of only for either one of the two. Then, a point cloud could have been realized, from which the dimensions of more human body parts can be potentially estimated instead of just the height.

Another limitation is the method of computing the step length from the radar data and the velocity from the groundtruth data. This was both done by using the relation in Eq. 43. Consequently, the computed gait parameter is dependent on the other two gait parameters. Therefore, inaccuracies inherent in the previously obtained parameters propagate into the computed parameter itself. This decreases the precision of the estimated step length and reduces the reliability of the established ground truth velocity.

Moreover, the method of measuring the ground truth step length does not enable to link the marked dots to a specific repetition. The step lengths were measured only after the five consecutive repetitions were performed. Therefore, the dots were not marked with the repetition number and could not be retraced. The step lengths could only be expressed using a mean value with a certain variation, instead of the absolute values between the marked steps.

VI. CONCLUSION

This research proposes a novel pipeline for developing an individualized human gait model utilizing a monostatic FMCW radar. The methodology was validated and essential components of the pipeline were verified. Furthermore, the study explored the gait parameter estimation error under varying target locations and orientations concerning the radar. Leveraging experimental data from three individuals walking in five different trajectories, the study successfully estimated three gait parameters. However, it was observed that the accuracy of gait parameter estimation diminishes with an increase of the aspect angle of the target. This is because the Doppler signature is less expressed for higher aspect angles due to the radar beam and the vector of the movement direction approaching orthogonality. Furthermore, single repetitions exhibited higher parameter errors compared to the average over five repetitions. Consequently, enhancing the trajectory length or increasing the number of repetitions would be advantageous to improve parameter estimation

TABLE VI: Comparison of gait parameter estimation in this work to previous work.

Reference	Scenario	Hardware	Extracted gait parameters	Mean reported velocity error	Mean reported step length error	Mean reported step frequency error
[8]	Walking on treadmill for two consecutive minutes with radar at knee height	two CW radars (24 GHz) with one radar at the front and one in the back	11 parameters among which step time, cadence and step length	-	up to 0.80 cm	up to 6.33e-4
[9]	Walking back and forth twice in a 10m-long hallway with radar at chest height	single FMCW radar (77 GHz)	Gait speed, step length, step count, step time	up to 0.0435 m/s	up to 4.82 cm	-
[10]	Walking straightly for 5.2 m towards radar with one radar on ground and one at torso height	two pulse-Doppler radars (5.8 GHz)	Foot velocity, torso velocity, step time	from foot radar: 0.144 m/s from torso radar: 0.092 m/s	-	from foot radar: 0.000 Hz from torso radar: 0.124 Hz
[11]	Walking back and forth 7 m in a straight line in front of two types of radars	one pulse-Doppler radar (10 GHz) and one FMCW radar (24 GHz)	Gait speed	Pulse-Doppler: 0.14 m/s FMCW: 0.06 m/s	-	-
[12]	Walking back and forth 6 m directly to radar at torso height	one FMCW radar (77 GHz)	Step time	-	-	Up to 0.140 Hz
This work	Walking for 3.5 m for five times in different trajectories with respect to radar at torso height	two co-located FMCW radars (77 GHz)	Step frequency, gait velocity, step length	Straight trajectory towards radar: up to 0.022 m/s Maximum for all trajectories: 0.077 m/s	Straight trajectory towards radar: up to 2.2 cm Maximum for all trajectories: up to 9.3 cm	Straight trajectory towards radar: up to 0.030 Hz Maximum for all trajectories: up to 0.128 Hz

TABLE VII: Comparison of human height parameter estimation in this work to previous work.

Reference	Hardware	Goal	Technique utilized	Mean absolute error
[18]	One FMCW radar, 12 virtual elements	Classify gait patterns and estimate 3D keypoint coordinates	convolutional neural network (CNN), trained with walking data of 74 persons, 3 minutes per person	Head keypoint: 7.10 cm
[19]	One FMCW radar, 64 virtual elements	Infer human skeletons from RF signals	CNN, trained with 12 hours of activity data in 22 environments	Head keypoint: 4.8 cm
[20]	Two FMCW radars, both 12 virtual elements	Skeletal pose estimation	NN, trained with 2 persons performing different actions in the same environment	Overall: 1.8 cm
[22]	Not mentioned	Motion classification and height estimation of pedestrians	Either empirical formula or random forest	Height: Empirical formula 19.4 cm Random forest 6.3 cm
This work	Two co-located FMCW radars, both 8 virtual elements	Human height estimation	Maximum height values over frames	Height: up to 10.9 cm for persons 1 and 2

performance. Compared to previous work employing similar scenarios, this work has similar accuracy in velocity, step length and step frequency. Additionally, the research examined the feasibility of estimating body part dimensions from processed radar data. Using the hardware employed in this study, the only dimensional parameter that could be estimated was the total human height. The sparse nature and the noise of the detected points with noise make it hard to determine the height accurately. The performance of body part dimension estimation could be potentially enhanced by increasing the number of antenna elements, the FOV, and the resolution in range and angle. Nevertheless, this improvement is constrained by the limited available resources (time, frequency, and space) that must be carefully considered. Compared to earlier work on human keypoint estimation, ML techniques have shown to be more accurate than the technique used in this work.

Several questions arise from this research, warranting future

investigations. First of all, the body part sizes in the model were scaled based on a single parameter, namely the human height. For future research, a method should be developed to extract dimensions for each individual body part. Then, the physical dimensions of the body parts in the gait model are more representative because they all will be independently determined instead of having a certain ratio between them. Utilizing a NN may be effective, given the promising results showcased in Table VII. Furthermore, future efforts could be directed towards creating a more sophisticated walking model. The current model used in this study, presented in [23], was derived from experimental data, thereby limiting its level of individualization in terms of flexing of joints. Nevertheless, gait is a complex and distinctive human characteristic, with joint rotational properties that may vary significantly from person to person. Some individuals might not exhibit the typical gait due to physical disabilities, resulting in deviating motions. It would be valuable to have a gait model which is individualized according to the exact movement of the target,

instead of having some generalized properties. Another aspect of the research which demands future investigation is the hardware and its deployment. The radar setup used in this study limits the performance of the system in terms of height estimation. Future research could analyze the influence of the size of the virtual array and investigate the effects of different radar deployment heights and tilt angles. This analysis may lead to improvements in the overall system performance.

ACKNOWLEDGMENT

Foremost, many thanks to my daily supervisor, dr. Yang Miao, for her excellent guidance and support throughout the entire process, and for participating in the experiments. Also, I want to thank the other committee members, prof. dr. ir. André Kokkeler and dr. ir. Alex Chiumento, for their willingness to be part of the committee and for evaluating my work. Furthermore, I would like to thank Hadi Alidoustaghdam for providing additional insights in the topic which helped me to advance in my research. Finally, I would like to thank the volunteers who showed interest in my research and participated in the experiments. Without them, this research would not have been possible.

REFERENCES

- [1] A. P. J. Zanardi, E. S. da Silva, R. R. Costa, E. Passos-Monteiro, I. O. Dos Santos, L. F. M. Krueel, and L. A. Peyré-Tartaruga, "Gait parameters of Parkinson's disease compared with healthy controls: A systematic review and meta-analysis," *Scientific reports*, vol. 11, no. 1, pp. 1–13, 2021.
- [2] L. di Biase, A. Di Santo, M. L. Caminiti, A. De Liso, S. A. Shah, L. Ricci, and V. Di Lazzaro, "Gait Analysis in Parkinson's Disease: An Overview of the Most Accurate Markers for Diagnosis and Symptoms Monitoring," *Sensors*, vol. 20, no. 12, 2020. [Online]. Available: <https://www.mdpi.com/1424-8220/20/12/3529>
- [3] J. Welzel, D. Wendtland, E. Warmerdam, R. Romijnders, M. Elshehabi, J. Geritz, D. Berg, C. Hansen, and W. Maetzler, "Step Length Is a Promising Progression Marker in Parkinson's Disease," *Sensors*, vol. 21, no. 7, 2021. [Online]. Available: <https://www.mdpi.com/1424-8220/21/7/2292>
- [4] T. B. Moeslund, A. Hilton, and V. Krüger, "A survey of advances in vision-based human motion capture and analysis," *Computer vision and image understanding*, vol. 104, no. 2-3, pp. 90–126, 2006.
- [5] J. P. Singh, S. Jain, S. Arora, and U. P. Singh, "Vision-Based Gait Recognition: A Survey," *IEEE Access*, vol. 6, pp. 70 497–70 527, 2018.
- [6] C. Benedek, B. Gálai, B. Nagy, and Z. Jankó, "Lidar-Based Gait Analysis and Activity Recognition in a 4D Surveillance System," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 1, pp. 101–113, 2018.
- [7] S. C. Mukhopadhyay, "Wearable Sensors for Human Activity Monitoring: A Review," *IEEE Sensors Journal*, vol. 15, no. 3, pp. 1321–1330, 2015.
- [8] A.-K. Seifert, M. Grimmer, and A. M. Zoubir, "Doppler radar for the extraction of biomechanical parameters in gait analysis," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 2, pp. 547–558, 2020.
- [9] H. Abedi, J. Boger, P. P. Morita, A. Wong, and G. Shaker, "Hallway gait monitoring using novel radar signal processing and unsupervised learning," *IEEE Sensors Journal*, vol. 22, no. 15, pp. 15 133–15 145, 2022.
- [10] F. Wang, M. Skubic, M. Rantz, and P. E. Cuddihy, "Quantitative gait measurement with pulse-Doppler radar for passive in-home gait assessment," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 9, pp. 2434–2443, 2014.
- [11] P. P. Morita, A. S. Rocha, G. Shaker, D. Lee, J. Wei, B. Fong, A. Thatte, A.-H. Karimi, L. Lin Xu, A. Ma *et al.*, "Comparison of gait speed estimation of multiple sensor-based technologies," in *Proceedings of the International Symposium on Human Factors and Ergonomics in Health Care*, vol. 8, no. 1. SAGE Publications Sage CA: Los Angeles, CA, 2019, pp. 135–139.
- [12] M. M. Rahman, D. Martelli, and S. Z. Gurbuz, "Gait Variability Analysis with Multi-Channel FMCW Radar for Fall Risk Assessment," in *2022 IEEE 12th Sensor Array and Multichannel Signal Processing Workshop (SAM)*. IEEE, 2022, pp. 345–349.
- [13] K. Saho, K. Uemura, K. Sugano, and M. Matsumoto, "Using micro-doppler radar to measure gait features associated with cognitive functions in elderly adults," *IEEE Access*, vol. 7, pp. 24 122–24 131, 2019.
- [14] T. Xu, D. An, Y. Jia, and Y. Yue, "A Review: Point Cloud-Based 3D Human Joints Estimation," *Sensors*, vol. 21, no. 5, 2021. [Online]. Available: <https://www.mdpi.com/1424-8220/21/5/1684>
- [15] N. Grammalidis and M. Strintzis, "Head detection and tracking by 2-D and 3-D ellipsoid fitting," in *Proceedings Computer Graphics International 2000*, 2000, pp. 221–226.
- [16] C. Panagiotakis and A. Argyros, "Parameter-free modelling of 2D shapes with ellipses," *Pattern Recognition*, vol. 53, pp. 259–275, 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0031320315004240>
- [17] M. Ye and R. Yang, "Real-time Simultaneous Pose and Shape Estimation for Articulated Objects Using a Single Depth Camera," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2014.
- [18] M. A. Alanazi, A. K. Alhazmi, O. Alsattam, K. Gnau, M. Brown, S. Thiel, K. Jackson, and V. P. Chodavarapu, "Towards a low-cost solution for gait analysis using millimeter wave sensor and machine learning," *Sensors*, vol. 22, no. 15, p. 5470, 2022.
- [19] M. Zhao, Y. Tian, H. Zhao, M. A. Alsheikh, T. Li, R. Hristov, Z. Kabelac, D. Katabi, and A. Torralba, "RF-Based 3D Skeletons," ser. SIGCOMM '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 267–281. [Online]. Available: <https://doi.org/10.1145/3230543.3230579>
- [20] A. Sengupta and S. Cao, "mmPose-NLP: A Natural Language Processing Approach to Precise Skeletal Pose Estimation Using mmWave Radars," *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–12, 2022.
- [21] Z. Sun, J. Wang, J. Sun, and P. Lei, "Parameter estimation method of walking human based on radar micro-doppler," in *2017 IEEE Radar Conference (RadarConf)*. IEEE, 2017, pp. 0567–0570.
- [22] M. Horn, O. Schumann, M. Hahn, J. Dickmann, and K. Dietmayer, "Motion classification and height estimation of pedestrians using sparse radar data," in *2018 Sensor Data Fusion: Trends, Solutions, Applications (SDF)*. IEEE, 2018, pp. 1–6.
- [23] R. Boulic, N. M. Thalmann, and D. Thalmann, "A global human walking model with real-time kinematic personification," *The visual computer*, vol. 6, pp. 344–358, 1990.
- [24] C. Iovescu and S. Rao, "The fundamentals of millimeter wave sensors," *Texas Instruments*, pp. 1–8, 2017.
- [25] X. Li, X. Wang, Q. Yang, and S. Fu, "Signal processing for TDM MIMO FMCW millimeter-wave radar sensors," *IEEE Access*, vol. 9, pp. 167 959–167 971, 2021.
- [26] A. Venon, Y. Dupuis, P. Vasseur, and P. Merriaux, "Millimeter wave FMCW radars for perception, recognition and localization in automotive applications: A survey," *IEEE Transactions on Intelligent Vehicles*, vol. 7, no. 3, pp. 533–555, 2022.
- [27] M. I. Skolnik, *Radar handbook*. McGraw-Hill Education, 2008.
- [28] P.-J. Chung, M. Viberg, and J. Yu, "DOA estimation methods and algorithms," in *Academic Press Library in Signal Processing*. Elsevier, 2014, vol. 3, pp. 599–650.
- [29] J. Capon, "High-resolution frequency-wavenumber spectrum analysis," *Proceedings of the IEEE*, vol. 57, no. 8, pp. 1408–1418, 1969.
- [30] J. B. Keller, "Geometrical theory of diffraction," *Josa*, vol. 52, no. 2, pp. 116–130, 1962.
- [31] V. C. Chen, D. Tahmoush, and W. J. Miceli, *Radar micro-Doppler signatures*. Institution of Engineering and Technology London, UK, 2014.
- [32] B. Erol and S. Z. Gurbuz, "A Kinect-based human micro-doppler simulator," *IEEE Aerospace and Electronic Systems Magazine*, vol. 30, no. 5, pp. 6–17, 2015.

- [33] V. C. Chen, F. Li, S.-S. Ho, and H. Wechsler, "Analysis of micro-doppler signatures," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 150, no. 4, pp. 271–276, 2003.
- [34] B. Godana, A. Barroso, and G. Leus, "Estimating human movement parameters using a software radio-based radar," *International Journal on Advances in Systems and Measurements*, vol. 4, no. 1–2, pp. 20–31, 2011.
- [35] H. M. Finn, "Adaptive detection mode with threshold control as a function of spatially sampled clutter-level estimates," *Rca Rev.*, vol. 29, pp. 414–465, 1968.
- [36] H. Rohling, "Radar CFAR thresholding in clutter and multiple target situations," *IEEE transactions on aerospace and electronic systems*, no. 4, pp. 608–621, 1983.
- [37] A. Jalil, H. Yousaf, and M. I. Baig, "Analysis of CFAR techniques," in *2016 13th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*. IEEE, 2016, pp. 654–659.
- [38] M. Ester, H.-P. Kriegel, J. Sander, X. Xu *et al.*, "A density-based algorithm for discovering clusters in large spatial databases with noise." in *kdd*, vol. 96, no. 34, 1996, pp. 226–231.
- [39] R. E. Kalman, "A new approach to linear filtering and prediction problems," 1960.
- [40] G. Welch, G. Bishop *et al.*, "An introduction to the Kalman filter," 1995.
- [41] A. Becker. (2023) Kalman filter tutorial. [Online]. Available: <https://www.kalmanfilter.net/default.aspx>
- [42] S. K. Mitra, *Digital signal processing: a computer-based approach*. McGraw-Hill Higher Education, 2001.
- [43] C. C. Gordon, T. Churchill, C. E. Clauser, B. Bradtmiller, J. T. McConville, I. Tebbetts, and R. A. Walker, "Anthropometric survey of us army personnel: Summary statistics, interim report for 1988," *Anthropology research project inc yellow springs OH*, 1989.
- [44] D. M. Patterson and R. A. Ashley, *A Nonlinear Time Series Workshop: A Toolkit for Detecting and Identifying Nonlinear Serial Dependence*. Boston, MA: Springer US, 2000, ch. Glint Tracking Errors in Radar", pp. 121–136.
- [45] P. van Dorp and F. Groen, "Human walking estimation with radar," *IEE Proceedings-Radar, Sonar and Navigation*, vol. 150, no. 5, pp. 356–365, 2003.
- [46] Texas Instruments, *IWR1443BOOST Evaluation Module mmWave Sensing Solution User's Guide*, December 2018–Revised May 2020. [Online]. Available: <https://www.ti.com/tool/IWR1443BOOST>
- [47] Texas Instruments., *Programming Chirp Parameters in TI Radar Devices*, May 2017–Revised February 2020. [Online]. Available: <https://www.ti.com/lit/an/swra553a/swra553a.pdf>
- [48] Diligent, *Cmod S7 Reference Manual*. [Online]. Available: <https://diligent.com/reference/programmable-logic/cmod-s7/reference-manual>
- [49] Bosch Sensortec, *BMI260: Accurate, low-power Inertial Measurement Unit (IMU)*, October 2018. [Online]. Available: <https://www.bosch-sensortec.com/products/motion-sensors/imus/bmi260/technical>