TECHNISCHE UNIVERSITÄT
CHEMNITZ

Radar-based Application of Pedestrian and Cyclist Micro-Doppler Signatures for Automotive Safety Systems

Von der Fakultät für Naturwissenschaften der
Technischen Universität Chemnitz

genehmigte

Dissertation

zur Erlangung des akademischen Grades

doctor rerum naturalium
(Dr. rer. nat.)

vorgelegt von
M.Eng. Patrick Held
geboren am 10.07.1990 in Ulm

Tag der Einreichung: 18.05.2021
Tag der Verteidigung: 28.02.2022

Gutachter:
Prof. Dr. Ulrich T. Schwarz
Prof. Dr. Thomas Brandmeier

Danksagung

Kurzfassung

Die sensorbasierte Erfassung des Nahfeldes im Kontext des hochautomatisierten Fahrens erfährt einen spürbaren Trend bei der Integration von Radarsensorik. Fortschritte in der Mikroelektronik erlauben den Einsatz von hochauflösenden Radarsensoren, die durch effiziente Verfahren sowohl im Winkel als auch in der Entfernung und im Doppler die Messgenauigkeit kontinuierlich ansteigen lassen. Dadurch ergeben sich neuartige Möglichkeiten bei der Bestimmung der geometrischen und kinematischen Beschaffenheit ausgedehnter Ziele im Fahrzeugumfeld, die zur gezielten Entwicklung von automotiven Sicherheitssystemen herangezogen werden können.

Im Rahmen dieser Arbeit werden ungeschützte Verkehrsteilnehmer wie Fußgänger und Radfahrer mittels eines hochauflösenden Automotive-Radars analysiert. Dabei steht die Erscheinung des Mikro-Doppler-Effekts, hervorgerufen durch das hohe Maß an kinematischen Freiheitsgraden der Objekte, im Vordergrund der Betrachtung. Die durch den Mikro-Doppler-Effekt entstehenden charakteristischen Radar-Signaturen erlauben eine detailliertere Perzeption der Objekte und können in direkten Zusammenhang zu ihren aktuellen Bewegungszuständen gesetzt werden. Es werden neuartige Methoden vorgestellt, die die geometrischen und kinematischen Ausdehnungen der Objekte berücksichtigen und echtzeitfähige Ansätze zur Klassifikation und Verhaltensindikation realisieren.

Wird ein ausgedehntes Ziel (z.B. Radfahrer) von einem Radarsensor detektiert, können aus dessen Mikro-Doppler-Signatur wesentliche Eigenschaften bezüglich seines Bewegungszustandes innerhalb eines Messzyklus erfasst werden. Die Geschwindigkeitsverteilungen der sich drehenden Räder erlauben eine adaptive Eingrenzung der Tretbewegung, deren Verhalten essentielle Merkmale im Hinblick auf eine vorausschauende Unfallprädiktion aufweist. Ferner unterliegen ausgedehnte Radarziele einer Orientierungsabhängigkeit, die deren geometrischen und kinematischen Profile direkt beeinflusst. Dies kann sich sowohl negativ auf die Klassifikations-Performance als auch auf die Verwertbarkeit von Parametern auswirken, die eine Absichtsbekundung des Radarziels konstituieren. Am Beispiel des Radfahrers wird hierzu ein Verfahren vorgestellt, das die orientierungsabhängigen Parameter in Entfernung und Doppler normalisiert und die gemessenen Mehrdeutigkeiten kompensiert.

Ferner wird in dieser Arbeit eine Methodik vorgestellt, die auf Grundlage des Mikro-Doppler-Profils eines Fußgängers dessen Beinbewegungen über die Zeit schätzt (Tracking) und wertvolle Objektinformationen hinsichtlich seines Bewegungsverhaltens offenbart. Dazu wird ein Bewegungsmodell entwickelt, das die nichtlineare Fortbewegung des Beins approximiert und dessen hohes Maß an biomechanischer Variabilität abbildet. Durch die Einbeziehung einer wahrscheinlichkeitsbasierten Datenassoziation werden die Radar-Detektionen ihren jeweils hervorrufenden Quellen (linkes und rechtes Bein) zugeordnet und eine Trennung der Gliedmaßen realisiert. Im Gegensatz zu bisherigen Tracking-Verfahren weist die vorgestellte Methodik eine Steigerung in der Genauigkeit der Objektinformationen auf und stellt damit einen entscheidenden Vorteil für zukünftige Fahrerassistenzsysteme dar, um deutlich schneller auf kritische Verkehrssituationen reagieren zu können.

Abstract

Sensor-based detection of the near field in the context of highly automated driving is experiencing a noticeable trend in the integration of radar sensor technology. Advances in microelectronics allow the use of high-resolution radar sensors that continuously increase measurement accuracy through efficient processes in angle as well as distance and Doppler. This opens up novel possibilities in determining the geometric and kinematic nature of extended targets in the vehicle environment, which can be used for the specific development of automotive safety systems.

In this work, vulnerable road users such as pedestrians and cyclists are analyzed using a high-resolution automotive radar. The focus is on the appearance of the micro-Doppler effect, caused by the objects' high kinematic degree of freedom. The characteristic radar signatures produced by the micro-Doppler effect allow a clearer perception of the objects and can be directly related to their current state of motion. Novel methods are presented that consider the geometric and kinematic extents of the objects and realize real-time approaches to classification and behavioral indication.

When a radar sensor detects an extended target (e.g., bicyclist), its motion state's fundamental properties can be captured from its micro-Doppler signature within a measurement cycle. The spinning wheels' velocity distributions allow an adaptive containment of the pedaling motion, whose behavior exhibits essential characteristics concerning predictive accident prediction. Furthermore, extended radar targets are subject to orientation dependence, directly affecting their geometric and kinematic profiles. This can negatively affect both the classification performance and the usability of parameters constituting the radar target's intention statement. For this purpose, using the cyclist as an example, a method is presented that normalizes the orientation-dependent parameters in range and Doppler and compensates for the measured ambiguities.

Furthermore, this paper presents a methodology that estimates a pedestrian's leg motion over time (tracking) based on the pedestrian's micro-Doppler profile and reveals valuable object information regarding his motion behavior. To this end, a motion model is developed that approximates the leg's nonlinear locomotion and represents its high degree of biomechanical variability. By incorporating likelihood-based data association, radar detections are assigned to their respective evoking sources (left and right leg), and limb separation is realized. In contrast to previous tracking methods, the presented methodology shows an increase in the object information's accuracy. It thus represents a decisive advantage for future driver assistance systems in order to be able to react significantly faster to critical traffic situations.

Contents

Appendix 127

CHAPTER 1

Introduction

The ongoing progress in driver assistance systems results in a tremendous augmentation of automotive sensor technology. Forward-looking, contactless interpretation of the vehicle periphery has undergone a substantial innovation stimulation in the last few years. Sophisticated driver assistance systems such as "adaptive cruise control", "blind spot detection", or "automated emergency braking" significantly increase the driver's comfort and the safety of each occupant. These systems are summarized under the term advanced driver assistance systems (ADAS) [109].

Whereas ADAS were exclusive to luxury category vehicles in the past, there is now a strong trend towards equipping medium and compact vehicle types. The realization of autonomous driving and the associated complexity of the surrounding's gathered information requires comprehensive sensor technology that contributes to a granular perception of the vehicle environment. The aim is to merge exteroceptive sensory elements such as ultrasound, lidar, camera, and radar to generate a detailed environment model for highly automated driving (HAD) [109].

Besides the vehicle's possibility of self-localization in complex territories, a detailed environment model enables the acquisition of object parameters, e.g., position, velocity, and type, which can be employed to determine hazardous situations and potential collisions. There are concepts for incorporating predictive sensor technology into pre-crash detection to enable early intervention of the vehicle's guidance and activation of active and passive safety systems in order to prevent or considerably mitigate the accident for all involved [75]. The overall objective in current safety development is to detect a potential accident with foresight and gain valuable, often decisive milliseconds for the activation of safety systems. In developing efficient sensor technology for environment detection and suitable object recognition methods, specific questions regarding the vehicle's periphery's situational understanding must always be answered [68]. What traffic participants surround the vehicle? What might be their intentions? What conclusions can the sensors derive from them?

In contrast to clearly structured roads and motorways, urban traffic poses major challenges for the design of the mentioned sophisticated sensor technologies and reliable object detection algorithms. The ambitious task of automated driving in urban traffic encompasses the mastering of substantially more complex road layouts and scenarios. From a driver's point of view, urban roads denote high traffic density, large distraction potential, and a dynamic and lively environment. The variety of urban road users comprising the most vulnerable participants, such as pedestrians and cyclists (vulnerable road user (VRU)), and their interaction in unclear confined spaces constitute a host of complexities for HAD

functions.

Considering the fact that more than 50% of all road traffic fatalities occur among VRU, it is precisely the urban surrounding whose accident prevention formulates the highest priority concerning future autonomous vehicles [110]. Consequently, sophisticated sensor applications and object detection systems for VRU enter the spotlight of comprehensive next-generation ADAS.

1.1 Automotive environmental perception

Current automated vehicles are usually equipped with an advanced sensor suite consisting of various sensors for elaborate environmental perception tasks. Most of the contactless sensors used can be divided into two groups according to their resolvable range. For the vehicle's immediate vicinity up to approximately 5 m, ultrasonic and capacity sensors are employed for parking assistance tasks on which basis complete automated parking systems are already realized [8]. Besides ultra short-range applications, state-of-the-art exteroceptive sensor technologies for demanding object detection in road scenery consist of cameras, lidars, and radars. Each sensor technology tackles the extensive process from raw sensor data acquisition to object-level differently and contributes distinctive strengths to the overall perception system. The concept of complementary sensor networking, in which the data of all sensors are fused and applied, is ubiquitous in the focus of the current development.

Camera: Cameras belong to the most intuitive sensor types and can be divided according to their spectral configuration into visible spectrum (VS) and infrared (IR) cameras. VS cameras denote the majority of the currently used vision-based systems and are usually integrated behind the vehicle's windshield. For dedicated VRU detection purposes, VS cameras typically provide a large field of view and high resolution capabilities. VS cameras may be designed in stereo arrangement with lateral offsets enabling depth perception. However, VS cameras suffer from insufficient sensitivity in poor lighting conditions due to their limited dynamic range, i.e., the ability to reproduce both dark and bright areas in the images [109]. A possible solution is the use of IR cameras that operate in the infrared spectral range. The traffic scenery is illuminated with infrared light where the thermal radiation image of pedestrians and animals can be directly received and forwarded to specific safety functions [109].

In the automotive field, cameras are primarily used for the task of VRU detection and classification [22], [34], [80]. The tremendous progress in the domain of artificial intelligence establishes deep learning architectures as an integral part of the vehicle's perception system. Elaborate convolutional neural network techniques, such as YOLO or fast R-CNN enable real-time object classification and outperform traditional computer-vision object detection procedures [82], [84], [114].

Lidar: Lidar sensors illuminate their environment with optical pulses and measure the characteristics of the reflected return signals. The distances are computed based on the light's travel time. Due to the small wavelength of the collimated laser light in the range

of $800 - 1500$ nm, lidar sensors provide an extremely high spatial resolution and can detect even non-metallic objects, such as pedestrians, in considerable detail. Current solid-state lidars employ semiconductor materials, eliminating the need for spinning mechanical parts and hence cumbersome implementations. Though, lidars' range performance is significantly influenced by the intensity of the emitted light pulse, whereas eye safety requirements limit the pulse power [109]. Moreover, if there is an increased attenuation of the atmosphere due to rain or fog, individual pulses interfere with the water droplets in the air leading to severe damping of the luminous power. Besides, as with all active sensors, the transmitted power must travel the distance again in the opposite direction after reflection from a particle yielding further damping.

Theoretically, the Doppler effect can be utilized to determine the radial relative velocities of surrounding objects. However, the increased requirements and the costs involved in measuring the Doppler frequency in the light spectrum prevent its implementation [109].

Radar: Since its first market-ready use as a mere adaptive cruise control system for highway driving in 1998, radar sensors have undergone continuous development and are utilized for various ADAS applications [77]. State-of-the-art radars emit frequency-modulated millimeter waves in the chirp sequence mode, which achieves the best exploitation of the signal power, measuring time, and bandwidth among all waveform designs [109]. The steady advances in microelectronic architecture development and semiconductor technology provide powerful integrated circuits and voltage-controlled oscillators, enabling large signal bandwidths and hence considerable range resolution. However, in contrast to pronounced diffuse backscattering at small wavelengths, as in lidar sensors, automotive radars' millimeter waves cause enhanced specular scattering when the radiated power impinges upon a target. As a result, the reflected wave is subject to multipath propagation, leading to temporally shifted receiving components [31], [61]. Moreover, the received power amplitudes typically possess significant fluctuations in dynamic scenarios and may vary over time and aspect angle.

While radar-based vehicle applications were initially limited to the 24 GHz band, the global market now utilizes the 77 GHz band. In the next few years, a further transition to the already allocated 79 GHz wideband (77-81 GHz) will be realized, resulting in a drastic increase in range, velocity, and angular resolution [99], [109]. The trend towards fast chirp sequence modulations, sophisticated antenna structures, and higher processable data rates give rise to advanced applications and places high-performance radars in the spotlight regarding environmental scene understanding for HAD. The utilization of objects' micro-motions induced by the micro-Doppler (μD) effect is one of the most promising topics in the current radar signal processing domain [77]. In pedestrians, the μD effect is manifested by the periodic limb motions while walking, which generate unique pedestrian identifying signatures in the radar return signal [15], [50], [53]. Those additive frequency components in the low-level data enable VRU specialized detection algorithms that improve the derivation of potential collision-related object features for integral vehicle safety systems. With their unique characteristics such as high robustness to weather conditions and accurate measurements of range, radial relative velocity, and angle in multi-target scenarios, next-generation radar sensors take a key role in future HAD systems.

1.2 Contributions of this work

Automated driving requires the interpretation of situations and the consideration of each conceivable scenario when interacting with highly agile targets such as pedestrians and cyclists. There are radar systems that can detect and classify pedestrians in some situations. These systems partially incorporate Doppler and μD information to distinguish pedestrians from other surrounding objects. However, despite the compelling potential of the μD effect, its use in pedestrian detection systems of current automotive radars is not yet widespread, as it is still unclear how reliable and effective μD information are in field usage.

Next-generation radar sensors provide high-resolution capabilities, whereas additional features such as identifying extremities' motions are to be determined and used for improved classification of VRU.

This work's contribution is the detailed analysis of high-resolution μD signatures of pedestrians and cyclists in urban scenarios regarding their spectral composition by characteristic limb motions or wheel and pedal spinning, respectively. The potential of μD-based short-range pedestrian and cyclist perception is demonstrated, and novel approaches for algorithmic use are revealed. The results of this work enable a more detailed object detection than conventional radar-based signal processing techniques provide and hence help to increase the safety of pedestrians, cyclists, and occupants in automated functions. Existing radar systems are mostly limited to the detection and classification of the current situation. However, besides the mere perception of the vehicle environment, the sensory detectability of predictive features of VRU would be highly desirable concerning fatal collision avoidance and the overall goal of creating social acceptance for autonomous vehicles of the future. Pedestrians and cyclists exhibit high motion maneuverability with the result that any auxiliary information regarding their behavior intention is welcome to reduce the uncertainty. This raises the question on how to teach an autonomous vehicle to predict the intention of pedestrians who communicate solely with their head and the position of their feet to express intentions in road traffic. The specific detection of a change in movements, such as a pedestrian who may intend to leave the assumed track or a cyclist increasing his speed to make it through a green light still, denotes significant relevance for active safety systems that cover a range of possible actions from low-level driver warnings to autonomous emergency braking.

This work addresses the interaction between the involved protagonists in urban traffic and presents μD-based possibilities for a more precise determination of a VRU's intended behavior within a few sensor cycles. Particularly, the objects' kinematic properties, such as a human's spatial locomotion while walking or a cyclist's pedaling motion, are considered to develop sophisticated algorithms with low computational effort that allow the implementation in automotive series sensors and reveal valuable object information that significantly enhances VRU safety functions of the future. The presented algorithms are analytically developed and provide transparent structures. In contrast to artificial intelligence methods, they possess a chronological computation sequence and allow traceable decision making, which is crucial for irreversible automotive safety functions.

The following publications have been produced during the period of doctoral candidacy.

Internationally refereed journal articles

- P. Held, D. Steinhauser, A. Koch, T. Brandmeier and U. T. Schwarz, "A novel approach for model-based pedestrian tracking using automotive radar," *IEEE Transactions on Intelligent Transportation Systems*, in press.

- Y. Deep, P. Held, S. S. Ram, D. Steinhauser, A. Gupta, F. Gruson, A. Koch and A. Roy, "Radar cross-sections of pedestrians at automotive radar frequencies using ray tracing and point scatterer modelling," *IET Radar, Sonar Navigation*, vol. 14, no. 6, pp. 833-844, 2020.

Internationally refereed conference papers

- P. Held, D. Steinhauser, A. Kamann, A. Koch, T. Brandmeier and U. T. Schwarz, "Normalization of micro-Doppler spectra for cyclists using high-resolution projection technique," *IEEE International Conference on Vehicular Electronics and Safety (ICVES)*, Cairo, Egypt, September 2019 (won Runner-up for Best Paper Award).

- P. Held, D. Steinhauser, A. Kamann, A. Koch, T. Brandmeier and U. T. Schwarz, "Micro-Doppler extraction of bicycle pedaling movements using automotive radar," *IEEE Intelligent Vehicles Symposium (IV)*, Paris, France, June 2019.

- P. Held, D. Steinhauser, A. Kamann, T. Holdgrün, I. Doric, A. Koch and T. Brandmeier, "Radar-based analysis of pedestrian micro-Doppler signatures using motion capture sensors," *IEEE Intelligent Vehicles Symposium (IV)*, Changshu, China, June 2018.

- A. Kamann, P. Held, F. Perras, P. Zaumseil, T. Brandmeier and U. T. Schwarz "Automotive radar multipath propagation in uncertain environments," *IEEE International Conference on Intelligent Transportation Systems (ITSC)*, Maui, USA, November 2018.

- D. Steinhauser, P. Held, B. Thöresz, and T. Brandmeier, "Towards safe autonomous driving: challenges of pedestrian detection in rain with automotive radar," *IEEE European Radar Conference (EuRAD)*, Utrecht, Netherlands, January 2021.

- D. Steinhauser, P. Held, A. Kamann, A. Koch and T. Brandmeier, "Micro-Doppler extraction of pedestrian limbs for high resolution automotive radar," *IEEE Intelligent Vehicles Symposium (IV)*, Paris, France, June 2019.

- P. Zaumseil, D. Steinhauser, P. Held, A. Kamann and T. Brandmeier, "Radar-based near field environment perception using back projection algorithm," *IEEE European Radar Conference (EuRAD)*, Paris, France, October 2019.

Filed patent applications

- P. Held, D. Steinhauser, A. Koch, L. D'Errico, T. Brandmeier (Conti Temic microelectronic GmbH and Technische Hochschule Ingolstadt), "Verfahren zum Bestimmen einer Bewegung eines Objekts," Patent Application DE 10 2020 210 380.1

- P. Held, D. Steinhauser, T. Brandmeier, A. Koch, L. D'Errico (Conti Temic microelectronic GmbH and Technische Hochschule Ingolstadt), "Verfahren zum Erfassen einer Bewegung eines Objekts," Patent Application DE 10 2019 200 849.6

- P. Held, D. Steinhauser, A. Kamann, T. Brandmeier, A. Koch, A. Amar (Conti Temic microelectronic GmbH and Technische Hochschule Ingolstadt), "Verfahren zum Erfassen von Teilbereichen eines Objekts," Patent Application DE 10 2019 200 141.6

- A. Kamann, P. Held, T. Brandmeier, A. Koch, F. Gruson, A. Forster (Conti Temic microelectronic GmbH and Technische Hochschule Ingolstadt), "Verfahren zur Bestimmung von unfallrelevanten Parametern mittels eines Radarsystems eines Fahrzeugs," Patent Application DE 10 2018 214 338.2

1.3 Thesis overview

This section provides a structural overview of this thesis. The structure is pictorially illustrated in Figure 1.1. The gray boxes refer to chapters, which provide a review of state-of-the-art and cover the context of this thesis's addressed problems. The black boxes comprise original contributions.

In Chapter 2, the physical fundamentals of automotive radar are given in order to empower the reader to understand the contributions in the following chapters. A detailed derivation of the underlying theories is provided. This includes an overview of the radiation and reception of electromagnetic waves in automotive radars as well as the phenomenon of the μD effect, which is derived in detail. Next, the radar measurement model is introduced, which comprises the theory of frequency-modulated waveforms for modern driver assistance systems. Besides, a general array signal model for direction-of-arrival estimation is presented. Subsequently, high-resolution radars' repercussions regarding what extent an object is detected are introduced, and corresponding pre-processing techniques for target extraction are presented.
In Chapter 3, we cover the occurrence of μD distributions at cyclists in field-relevant scenarios. The most prominent μD sources are physically derived and employed in novel
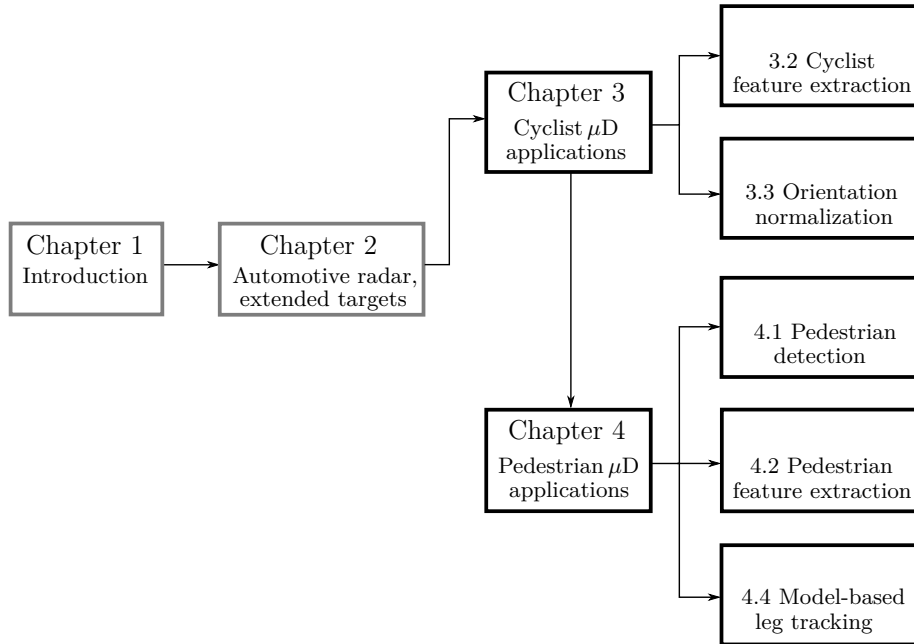
Figure 1.1: Structural overview of the thesis, gray boxes represent review and introductory parts, black boxes possess original contributions.

signal processing techniques for radar-based feature extraction. A real-time procedure for adaptive pedaling extraction is demonstrated in Section 3.2, and its relevance regarding the specific detection of behavior intents is presented using real radar data in controlled experiments. The orientation's effect on extended targets is described and discussed. In Section 3.3, a procedure for the normalization of orientation-dependent ambiguities at cyclists is introduced, enabling a spectral representation with static parameters for enhanced radar-based classification and feature extraction tasks, respectively.

Chapter 4 analyzes high-resolution μD signatures of human walking and presents limb extraction procedures for enhanced pedestrian tracking and intent estimation. For this case, in Section 4.4, we present a model-based tracking approach with low computational complexity that utilizes the μD signatures of a human's lower body locomotion to estimate the foot's spatial and kinematic propagation. The provided procedure's performance is analyzed with corresponding motion capture data in different movement trajectories, revealing real-time intent estimation features.

Finally, the conclusions of this thesis's contributions are given in Chapter 5, and an outlook to possible future work is drawn.

CHAPTER 2

Automotive radar

This chapter presents the fundamental principles of an automotive radar system for detailed environment perception. After deriving the physical foundations of electromagnetic wave propagation, a radar measurement model is introduced that comprises the emitting and processing of radar signals to infer essential information of surrounding targets. The principle of the frequency-modulated continuous wave (FMCW) design and the chirp sequence modulation as its efficient implementation is introduced in detail. A parameterizable FMCW radar system is used throughout this work, providing high-resolution measurement capability in several dimensions. As a consequence, targets may appear as extended, i.e., they give rise to a vast number of reflections making the point-target assumption obsolete. A signal processing routine including power detection and object clustering for target extraction is provided and exemplary applied to a real radar measurement.

2.1 Physical fundamentals

This section addresses the physical fundamentals of automotive radars and establishes basic definitions and nomenclature used throughout this work. The phsyical principles introduced in the following give a brief overview of the radiation and reception of electromagnetic energy in the context of automotive radars. The radar cross-section (RCS) as a measure for a target's reflectivitiy is introduced and utilized to formulate the radar equation. Moreover, the phenomenon of the micro-Doppler effect is defined in detail as it lays the fundament for this work.

2.1.1 Radar cross section

A radar-based vehicle environment detection system is based on the reflection of bundled electromagnetic waves by other road users and peripheral objects. The echoes reflected by these radar targets are received and evaluated. The reflectivity of a radar target, described by its RCS σ, has a significant influence on detectability. The RCS is defined by the Institute of Electrical and Electronics Engineers (IEEE) as [60]

$$\sigma = \lim_{r \to \infty} 4\pi r^2 \frac{|\boldsymbol{E}_r|^2}{|\boldsymbol{E}_i|^2}, \tag{2.1}$$

where r is the radial range between source and target and $\boldsymbol{E}_r$ and $\boldsymbol{E}_i$ are the far-field scattered and incident electric field intensities, respectively. The unity of σ, m^2, illustrates that the value depends primarily on the size of the illuminated area. Moreover, geometry,

the target's material, polarization, the wavelength λ, and orientation are further factors influencing the RCS. Table 2.1 shows examples of orders of magnitude of typical objects in the vehicle environment in both linear and logarithmic units. The following applies to the specification in dBsm

$$\frac{\sigma}{\text{dBsm}} = 10 \cdot \log_{10}\left(\frac{\sigma}{1\text{m}^2}\right). \tag{2.2}$$

Generally, the RCS is subject to an intense phase sensitivity to any radial movements. As a consequence, phase changes occurring within a frequency cell as a function of radial range can cause interferences and may lead to severely fluctuating power amplitudes [29], [70], [112]. This sensitivity can be illustrated with an example. Assuming a target somewhere in the radar's field-of-view is displaced by 1 mm in the radial direction to its initial distance, the resulting phase change at the receiver of the radar is

$$\Delta\varphi = \frac{2\pi}{\lambda} \cdot 1\,\text{mm} \cdot 2 \approx \pi, \tag{2.3}$$

where $2\pi/\lambda$ defines the wavenumber considering a wavelength of $\lambda = 3.9\,\text{mm}$ (77 GHz). The factor 2 takes into account the two-way propagation of the traveling wave. In other words, a small motion in the order of the half-wavelength provokes a phase shift of 2π.

Table 2.1: Order of RCS magnitudes for typical automotive objects [29].

object class	$\frac{\sigma}{\text{m}^2}$	$\frac{\sigma}{\text{dBsm}}$
truck	1000	30
car	100	20
motorcycle	10	10
pedestrian	1	0

2.1.2 Radar equation

Automotive radars typically operate in a frequency range of $76 - 81\,\text{GHz}$ ($\equiv 0.39 - 0.37\,\text{cm}$ wavelength), yielding to wavelengths that are much smaller than the actual distances to reflecting objects. As a result, the transmitting antenna can be regarded as a point source. The power density at a distance r_{TO} from an isotropic radiating point source emitting with power P_{T} is [38], [70]

$$S_{\text{T}} = \frac{P_{\text{T}}}{4\pi r_{\text{TO}}^2}. \tag{2.4}$$

Automotive radar antennas are designed to transmit stronger in one direction than another (*anisotropic* radiator). The directional effect of the transmitting path is achieved by the antenna gain G_{T}. Consequently, the power density in the direction of the main lobe becomes

$$S_{\text{T}} = \frac{P_{\text{T}}G_{\text{T}}}{4\pi r_{\text{TO}}^2}. \tag{2.5}$$

An object with RCS σ reflects the power $S_\mathrm{T}\sigma$ isotropically leading to the received power density S_R at the receiving antenna at distance r_OR

$$S_\mathrm{R} = \frac{P_\mathrm{T}G_\mathrm{T}\sigma}{(4\pi)^2 r_\mathrm{TO}^2 r_\mathrm{OR}^2}. \tag{2.6}$$

The consideration of the receiving antenna gain G_R with

$$G_\mathrm{R} = a_\mathrm{R,eff}\frac{4\pi}{\lambda^2}, \tag{2.7}$$

where $a_\mathrm{R,eff}$ is the effective area of the receiving antenna completes the radar equation expressed as

$$P_\mathrm{R} = P_\mathrm{T}\frac{G_\mathrm{T}G_\mathrm{R}\lambda^2\sigma}{(4\pi)^3 r_\mathrm{TO}^2 r_\mathrm{OR}^2}. \tag{2.8}$$

Since the distance between the transmitting and receiving antenna is much smaller than the distance to a point target, r_TO and r_OR appear to be almost equal. The resulting power of four shows the rapid decrease of the received power with increasing distance. However, with extended targets, the effective area increases with increasing distance r due to the widening of the radar's main lobe. In this case, the dependence according to the law of the fourth power is no longer valid [29].

2.1.3 Micro-Doppler effect

Radar sensors use the Doppler effect to determine the radial relative velocity of objects, as well as to differentiate between moving and stationary objects. The Doppler effect describes a change in frequency of an incident wave for a receiver moving relatively to its radiating source. The algebraic sign of this frequency shift can be positive or negative, depending on the target's motion direction. An approaching target compresses the impinging wave resulting in a smaller wavelength, whereas an object that moves away from the radar expands the wave and hence increases the wavelength of the reflected signal. To define objects' overall motion behavior, we first describe the macroscopic motion, i.e., the global translational movement behavior. Then, we cover the microscopic motion comprising the additional micro-motions an object can exhibit while moving.
In case of no relative movement, the reflection reaches the radar with a time delay τ due to the speed of light c and the distance of the radar target r according to

$$\tau = \frac{2r}{c}. \tag{2.9}$$

In case of existing relative movement between radar and target, the occurring time dependence of τ caused by the change in distance of the moving target must be considered. Assuming a target movement with constant velocity, the time of flight delay $\tau(t)$ can be expressed as [27], [109]

$$\tau(t) = \frac{2\left(r(0) + \boldsymbol{v}_\mathrm{r,rel}t\right)}{c + \boldsymbol{v}_\mathrm{r,rel}} = \frac{2r(t)/c}{1 + \frac{\boldsymbol{v}_\mathrm{r,rel}}{c}} \approx \frac{2r(t)}{c} \quad \text{for } \boldsymbol{v}_\mathrm{r,rel} \ll c, \tag{2.10}$$

where $r(0)$ refers to the initial distance of the target at $t = 0$, and $\boldsymbol{v}_{\text{r,rel}}$ is the radial relative velocity component that can be measured by the radar using the Doppler effect. This Doppler frequency shift can be derived using (2.10). Let $u_{\text{T}}(t)$ be a transmitted signal with constant frequency f_0 and initial phase φ_0 of the form [109]

$$u_{\text{T}}(t) = A_{\text{T}} \cos(2\pi f_0 t + \varphi_0), \tag{2.11}$$

the received signal including the Doppler frequency shift can be given by

$$u_{\text{R}}(t) = A_{\text{R}} \cos(2\pi f_0 \left(t - \tau(t) \right) + \varphi_0 + \pi) \tag{2.12}$$

$$= A_{\text{R}} \cos\left(2\pi f_0 \left(t - \frac{2 \left(r(0) + \boldsymbol{v}_{\text{r,rel}} t \right)}{c} \right) + \varphi_0 + \pi \right) \tag{2.13}$$

$$= A_{\text{R}} \cos\left(2\pi f_0 t \left(1 - \frac{2\boldsymbol{v}_{\text{r,rel}}}{c} \right) \underbrace{-2\pi f_0 \frac{2r(0)}{c} + \varphi_0 + \pi}_{\varphi} \right) \tag{2.14}$$

$$= A_{\text{R}} \cos\left(2\pi \left(f_0 - f_0 \frac{2\boldsymbol{v}_{\text{r,rel}}}{c} \right) t + \varphi \right). \tag{2.15}$$

Here, φ comprises the overall phase of the received signal, including a phase shift π resulting from an assumed total reflection on a metallic surface. Finally, the Doppler frequency f_{D} is the difference between the received and sent frequency

$$f_{\text{D}} = \left(f_0 - f_0 \frac{2\boldsymbol{v}_{\text{r,rel}}}{c} \right) - f_0 = -f_0 \frac{2\boldsymbol{v}_{\text{r,rel}}}{c} = -\frac{2\boldsymbol{v}_{\text{r,rel}}}{\lambda_0}. \tag{2.16}$$

Since most moving objects in the vicinity of a radar do not represent rigid bodies, their overall motion can consist of many individual movements. Besides linear movements, many objects show other dynamics, such as mechanical vibrations or rotations at different target parts. Every part of an object or any structure that is not rigidly connected to the object can have a deviating motion state and hence different relative velocities. Those so-called micro-motions induce side-band frequency modulations on the returned signal resulting in additional Doppler shifts known as the micro-Doppler (μD) effect. Initially applied in coherent laser technology, the μD effect occurs in various moving objects such as an aircraft propeller or rotors of helicopters and vibrating engines [24], [25], [113]. Especially in the areas of automotive environment perception, the μD effect is used in near field applications such as wheel and pedestrian detection, where the features of the μD effect allow entirely new methods for the algorithmic use of radar detections [23], [39], [62], [65].

In the following, the μD effect is explained using a point scattering model that is assumed to reflect the incident power perfectly [24], [25]. Figure 2.1 shows a stationary radar with coordinate system (U, V, W) and origin Q. The point scatterer P is located on a cylindrical rigid body with local coordinate system (x, y, z) showing translational and rotational motions with respect to the radar coordinates. The reference coordinate system with origin O has a distance $r(0)$ to the radar. Additionally, a reference coordinate

system (X, Y, Z) is added to capture the target's rotations, and hence is limited to move only in the target's translational direction. The target moves with translational velocity $\boldsymbol{v}_{\mathrm{r,rel}}$ that corresponds to the radial relative velocity, and rotation angular velocity $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)^T = (\omega_X, \omega_Y, \omega_Z)^T$ relatively to the radar sensor. Due to the motion of the rigid body, the particle P at time $t = 0$ is moving to P' at time $t = 1$. The movement comprises two parts: [24], [25]

1. A translation from P to P'' due to velocity $\boldsymbol{v}_{\mathrm{r,rel}}$, i.e., $\overrightarrow{OO'} = \boldsymbol{v}_{\mathrm{r,rel}}t$.

2. A rotation from P'' to P' due to $\boldsymbol{\omega}$.

The range vector at time $t = 1$ from the radar to the particle at P' can be expressed as

$$\overrightarrow{QP'} = \overrightarrow{QO} + \overrightarrow{OO'} + \overrightarrow{O'P'} = r(0) + \boldsymbol{v}_{\mathrm{r,rel}}t + \mathbf{R}_t r_{OP}, \tag{2.17}$$

with

$$r(t) = \|r(0) + \boldsymbol{v}_{\mathrm{r,rel}}t + \mathbf{R}_t r_{OP}\|, \tag{2.18}$$

and

$$r_{O'P'} = \overrightarrow{O'P'} = \mathbf{R}_t \overrightarrow{O'P''} = \mathbf{R}_t r_{OP}, \tag{2.19}$$

where $\|\cdot\|$ denotes Euclidean norm, and $r_0 = (X_0, Y_0, Z_0)^T$ is the location of particle P at $t = 0$ with respect to the reference coordinate system, and $\mathbf{R}_t$ is a rotation matrix performing the rotation from P'' to P'. Suppose the radar sensor transmits a wave with constant frequency f_0, the returned base-band signal from the point scatterer P is

$$u_{\mathrm{R}}(t) = A_{\mathrm{R}} e^{2\pi i f_0 \frac{2r(t)}{c}}, \tag{2.20}$$

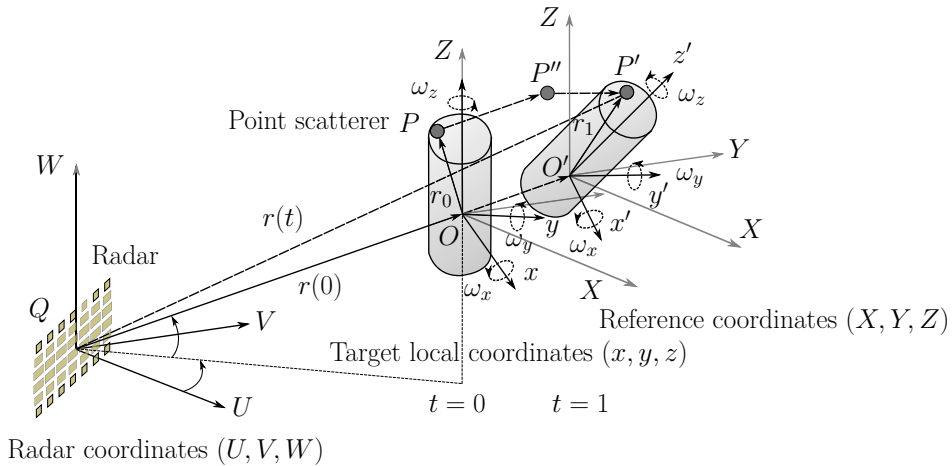**Figure 2.1:** Stationary radar sensor and target showing translational and rotational motion [24].

with a backscattering amplitude A_{R}, and phase

$$\varphi(t) = 2\pi f_0 \frac{2r(t)}{c}. \tag{2.21}$$

The time derivative of the phase yields the Doppler frequency shift of the target [24]

$$\boldsymbol{f}_{\mathrm{D}} = \frac{1}{2\pi} \frac{\mathrm{d}\varphi(t)}{\mathrm{d}t} = \frac{2f_0}{c} \frac{\mathrm{d}}{\mathrm{d}t} r(t), \tag{2.22}$$

$$= \frac{2f_0}{c} \frac{1}{2r(t)} \frac{\mathrm{d}}{\mathrm{d}t} \left((r(0) + \boldsymbol{v}_{\mathrm{r,rel}}t + \mathbf{R}_t r_{OP})^T (r(0) + \boldsymbol{v}_{\mathrm{r,rel}}t + \mathbf{R}_t r_{OP}) \right), \tag{2.23}$$

$$\approx \frac{2f_0}{c} \left(\boldsymbol{v}_{\mathrm{r,rel}} + \frac{\mathrm{d}}{\mathrm{d}t} (\mathbf{R}_t r_{OP}) \right)^T \boldsymbol{n}, \tag{2.24}$$

where

$$\boldsymbol{n} = \frac{(r(0) + \boldsymbol{v}_{\mathrm{r,rel}}t + \mathbf{R}_t r_{OP})}{(\|r(0) + \boldsymbol{v}_{\mathrm{r,rel}}t + \mathbf{R}_t r_{OP}\|)}, \tag{2.25}$$

is the unit vector of $\overrightarrow{QP'}$. The target rotates along the unit vector $\boldsymbol{\omega}' = \boldsymbol{\omega}/\|\boldsymbol{\omega}\|$ with angular rotation velocity $\|\boldsymbol{\omega}\|$. Assuming an infinitesimal rotational motion of the target during each observation interval, the rotation matrix $\mathbf{R}_t$ can be expressed by (see appendix A)

$$\mathbf{R}_t = e^{\hat{\omega}t}, \tag{2.26}$$

with

$$\hat{\boldsymbol{\omega}} = \begin{bmatrix} 0 & -\omega_Z & \omega_Y \\ \omega_Z & 0 & -\omega_X \\ -\omega_Y & \omega_X & 0 \end{bmatrix}, \tag{2.27}$$

being a skew-symmetric matrix that is associated with $\boldsymbol{\omega}$. Plugging (2.26) into (2.24) yields

$$\boldsymbol{f}_{\mathrm{D}} = \frac{2f_0}{c} \left(\boldsymbol{v}_{\mathrm{r,rel}} + \frac{\mathrm{d}}{\mathrm{d}t} \left(e^{\hat{\omega}t} r_{OP} \right) \right)^T \boldsymbol{n} \tag{2.28}$$

$$= \frac{2f_0}{c} \left(\boldsymbol{v}_{\mathrm{r,rel}} + \hat{\boldsymbol{\omega}} e^{\hat{\omega}t} r_{OP} \right)^T \boldsymbol{n} \tag{2.29}$$

$$= \frac{2f_0}{c} \left(\boldsymbol{v}_{\mathrm{r,rel}} + \hat{\boldsymbol{\omega}} r_{O'P'} \right)^T \boldsymbol{n} \tag{2.30}$$

$$= \frac{2f_0}{c} \left(\boldsymbol{v}_{\mathrm{r,rel}} + \boldsymbol{\omega} \times r_{O'P'} \right)^T \boldsymbol{n}. \tag{2.31}$$

Assuming $r(0) >> \|\boldsymbol{v}_{\mathrm{r,rel}}t + \mathbf{R}_t r_{OP}\|$, the unit vector $\boldsymbol{n}$ can be approximated by $\boldsymbol{n} \approx r(0)/\|r(0)\|$. Consequently, the total Doppler shift of the target is

$$\boldsymbol{f}_{\mathrm{D}} = \frac{2f_0}{c} \left(\boldsymbol{v}_{\mathrm{r,rel}} + \boldsymbol{\omega} \times r_{O'P'} \right). \tag{2.32}$$

Note that the first term in (2.32) is the Doppler shift induced by the translational movement, whereas the second term in (2.32) is the micro-Doppler shift induced by the rotation of the

target:

$$\boldsymbol{f}_{\mu\mathrm{D}} = \frac{2f_0}{c}\left(\boldsymbol{\omega} \times r_{O'P'}\right). \tag{2.33}$$

With the continuous improvement of microelectronics and the resulting increase in resolution, the μD effect is becoming more important for radar-based applications. Since most secondary movements in the automotive context inducing a μD effect are relatively weak targets in terms of their reflectivity compared to their bulk area (car wheels to the vehicle body, bicycle wheels to frame and cyclist, limbs to torso, etc.), μD detections typically show a low signal-to-noise ratio (SNR). Therefore, they are most pronounced in the immediate vehicle environment.

The Doppler frequency in automotive radar follows $\boldsymbol{f}_{\mathrm{D}} = 510\,\mathrm{Hz}\cdot\boldsymbol{v}_{\mathrm{r,rel}}$ assuming a carrier frequency of $f_c = 76.5\,\mathrm{GHz}$ and are consequently in the range of several kHz. The μD frequencies depend on the object and its degrees of motion. For rotating objects such as tires, the frequencies range from 0 to twice the Doppler frequency $\boldsymbol{f}_{\mathrm{D}}$. For pedestrians, the μD frequencies correspond to the limb velocity components and occupy frequencies from 0 to three times the Doppler frequency $\boldsymbol{f}_{\mathrm{D}}$.

2.2 Radar measurement model

State-of-the-art automotive radars use the FMCW waveform design [77]. Contrary to pulse-Doppler radars, FMCW radars provide cost-effective analog radio frequency (RF) hardware, possess modest digital processing requirements, and achieve the best exploitation of the signal power, measuring time, and modulation bandwidth [108], [109]. The following subsections introduce the radar's physical measurement model used in this work in order to estimate attributes of surrounding targets, i.e., radial distance, radial relative velocity, and angular information.

2.2.1 FMCW radar

In linear frequency-modulated radars, the idea is to continuously increase or decrease the transmit frequency with a defined slope. In doing so, an offset to the frequency of the reflected signal is created. This offset allows both radial distance and Doppler velocity to be determined. The transmit frequency for one chirp is of the form [108]

$$f_{\mathrm{T}}(t) = f_c + \dot{f}t \quad \text{for} \quad t \in [-T/2;\, T/2], \tag{2.34}$$

where T is the modulation time, Δf is the frequency bandwidth, $\dot{f} = \Delta f/T$ is the slope, and f_c is the carrier frequency of the chirp. The transmitted signal is then

$$S_{\mathrm{T}}(t) = \cos\left(\varphi_{\mathrm{T}}(t)\right), \tag{2.35}$$

where

$$\varphi_{\text{T}}(t) = 2\pi \int_{-T/2}^{t} f_{\text{T}}(t)\mathrm{d}t \tag{2.36}$$

$$= 2\pi \left(f_c t + \frac{1}{2}\dot{f}t^2 \right) \Big|_{-T/2}^{t} \tag{2.37}$$

$$= 2\pi \left(f_c t + \frac{1}{2}\dot{f}t^2 \right) - \varphi_{T0}, \tag{2.38}$$

is the overall phase of the integrated transmit frequency where φ_{T0} denotes an initial arbitrary phase. The received signal, which can be regarded as a damped and delayed version of the transmitted signal, reaches the sensor with delay τ:

$$S_{\text{R}}(t) \propto \delta \cdot S_{\text{T}}(t - \tau), \tag{2.39}$$

where δ denotes the attenuation. Signal attenuation is mainly due to path loss $1/(r_{\text{TO}}^2 r_{\text{OR}}^2)$ in (2.8), target parameter like RCS, and weather conditions such as heavy rain. In case of no relative movements between radar and target, the frequency modulation and the time delay, τ, describe a frequency difference, or intermediate frequency (IF), δf, between the transmitted and received signal whose relation reads [27]

$$\frac{\tau}{T} = \frac{\delta f}{\Delta f}. \tag{2.40}$$

A typical way of determining δf is to mix the received signal, $S_{\text{R}}(t)$, with a replica of the transmitted signal, $S_{\text{T}}(t)$, to the base-band. Since this corresponds mathematically to a multiplication of two oscillations, the following trigonometric relationship applies to the phases $\varphi_{\text{R}}(t) = \varphi_{\text{T}}(t - \tau)$ and $\varphi_{\text{T}}(t)$ of the received and transmitted signal, respectively [14]

$$\cos\left(\varphi_{\text{R}}(t)\right) \cdot \cos\left(\varphi_{\text{T}}(t)\right) = \frac{1}{2}\left(\cos(\varphi_{\text{R}}(t) - \varphi_{\text{T}}(t)) + \cos(\varphi_{\text{R}}(t) + \varphi_{\text{T}}(t))\right). \tag{2.41}$$

A subsequent low-pass filter removes the upper sideband of the real-valued[1] correlation mixer output in (2.41), which is almost twice the carrier frequency $2f_c$. Considering a static τ, as defined in (2.9), the down-converted phase can be given by

$$\Delta\varphi(t) = 2\pi \left(\gamma + \dot{f}\frac{2r}{c}t \right), \tag{2.42}$$

---

1 For hardware cost and power consumption reasons, current off-the-shelf monolithic microwave integrated circuits use real receivers only instead of in-phase and quadrature receivers [32].

where γ denotes a constant range-dependent phase term. The generated IF of the received signal is directly proportional to r and reads

$$\delta f = \dot{f}\frac{2r}{c}.\tag{2.43}$$

Assuming several targets, δf consists of several range-dependent frequencies. Range resolution refers to the ability to resolve two targets in the distance dimension. Two frequencies can be distinguished if they have a difference of at least $1/T$ in the frequency domain

$$|\delta f_1 - \delta f_2| \geq 1/T.\tag{2.44}$$

Using (2.40) and the condition of (2.44), range resolution Δr can be defined as

$$\Delta r = \frac{c}{2}\left(\tau_1 - \tau_2\right) = \frac{c}{2}\frac{(\delta f_1 - \delta f_2)}{\dot{f}} \geq \frac{c}{2}\frac{1}{\Delta f},\tag{2.45}$$

and is consequently determined by the bandwidth Δf of the chirp.

Suppose radar and target are moving relative to each other, a Doppler shift is induced to the received signal. Hence, both the range-dependent difference frequency, δf, and the Doppler frequency, f_D, contribute additively to the resulting total intermediate frequency f_IF. Using the time dependence of τ induced by the relative movement given in (2.10), and dropping all terms containing $1/c^2$ and neglecting the range-Doppler-coupling, the mixed phase in (2.42) extends to [108]

$$\Delta\varphi(t) = 2\pi\left(\gamma + \left(f_c\frac{2\boldsymbol{v}_\mathrm{r,rel}}{c} + \dot{f}\frac{2r(0)}{c}\right)t\right),\tag{2.46}$$

with total intermediate frequency

$$f_\mathrm{IF} = \underbrace{f_c\frac{2\boldsymbol{v}_\mathrm{r,rel}}{c}}_{f_\mathrm{D}} + \underbrace{\dot{f}\frac{2r(0)}{c}}_{\delta f}.\tag{2.47}$$

This leads to an ambiguity of the quantities r and $\boldsymbol{v}_\mathrm{r,rel}$ to be determined and additional measures have to be taken to compute the respective values. One way to resolve the ambiguity is to adjust the modulation pattern to chirps of different slopes. Figure 2.2 illustrates the principle of transmitting a signal with different chirp slopes. The intermediate frequencies are equal in magnitude with a stationary target as the chirp rises or falls, but reversed in sign. With relative movement, the intermediate frequencies change both by the magnitude of the Doppler shift f_D. Thus, the difference of the intermediate frequencies, $f_\mathrm{IF,1}$ and $f_\mathrm{IF,2}$, yields the distance r, whereas the sum yields the radial relative velocity $\boldsymbol{v}_\mathrm{r,rel}$ [29].

2.2.2 Chirp sequence modulation

Despite the possibility to measure r and $\boldsymbol{v}_\mathrm{r,rel}$ simultaneously, the classical FMCW approach has weaknesses in multiple target scenarios. Each target creates another pair of intersection
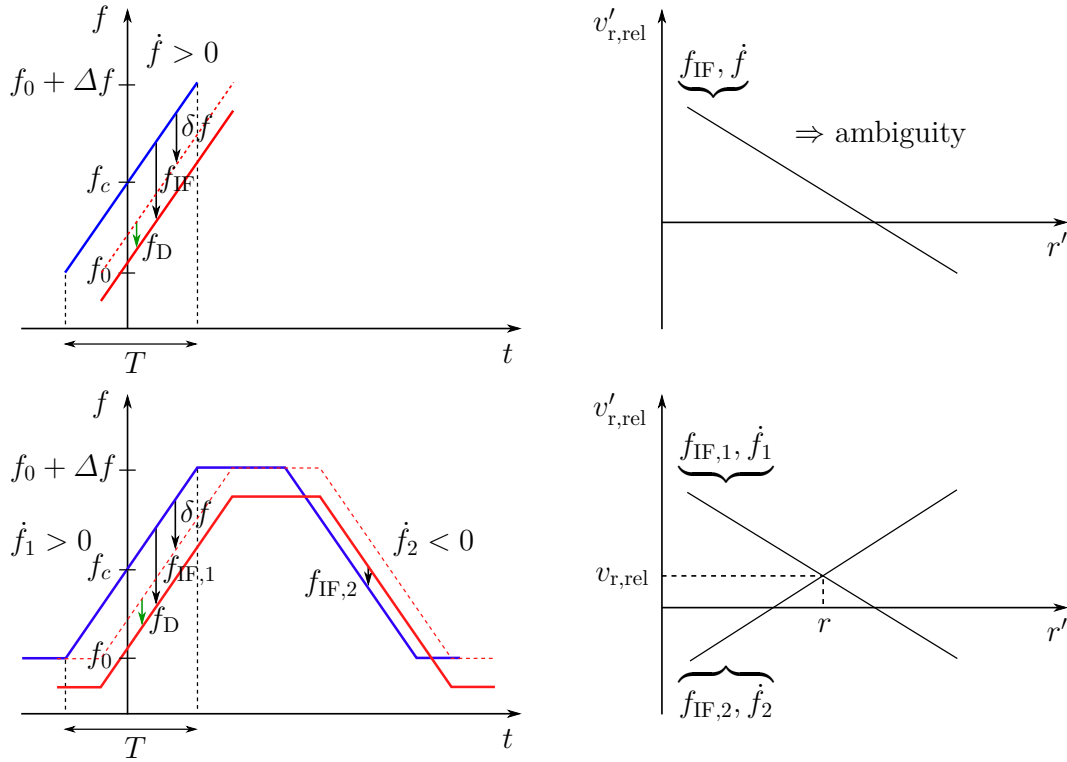
Figure 2.2: Measuring principle of an FMCW radar with transmitting signal (——), and receiving signal (——). The travel time and the frequency modulation of the transmitted signal lead to a difference frequency, f_{IF}, whose magnitude is proportional to the distance of the target in case of a static scenario. If $v_{r,rel} \neq 0$, the Doppler frequency, f_D, contributes to f_{IF} resulting in an ambiguous difference frequency. Using a transmitting signal with different chirp slopes, $\dot{f}$, yields several difference frequencies and hence resolve the ambiguity in the r-$v_{r,rel}$-space [29].

lines in the r-$\boldsymbol{v}_{r,rel}$-space, which results in twice the number of intersection points than actual existent targets. Consequently, each target requires a further transmitted chirp slope to resolve the ambiguity. Considering a typical automotive scenery containing a variety of surrounding targets, this may result in a computational demanding linear equation system, which has to be solved.

Alternatively, a modified form of the FMCW method - the chirp sequence modulation - can be applied [27]. From (2.47), it gets apparent that the range-dependent frequency contribution is influenced by the chirp slope, while the Doppler-dependent contribution is not. The chirp sequence modulation is characterized by the successive use of very short chirps of high frequency change per measurement cycle. By doing so, the range-dependent term of one chirp dominates the Doppler contribution making it negligible. Instead, $\boldsymbol{v}_{r,rel}$ is obtained by determining the phase shift from chirp to chirp occurring only due to the velocity. This allows several targets with different Doppler velocities within a range cell to be resolved without ambiguity.

To emphasize the essential characteristic of the chirp sequence modulation, (2.35) is

repeated under a redefinition of the independent time variable t given by [14], [27]

$$S_\text{T}(t_m, m_s) = \cos\left(\varphi_\text{T}(t_m, m_s)\right), \tag{2.48}$$

with

$$\varphi_\text{T}(t_m, m_s) = 2\pi\left(f_c t_m + \frac{1}{2}\dot{f}t_m^2\right) - \varphi_{T0},$$

and

$$t = t_m + m_s G, \;\; \forall\; \{t_m \in \mathbb{R}|-T/2 \le t_m < T/2\}\;\{m_s \in \mathbb{Z}|-M_s/2 \le m_s < M_s/2\}, \tag{2.49}$$

where the integer m_s numbers the consecutive chirps, and t_m is the time elapsed since the m_sth of M_s chirps has started, as illustrated in Figure 2.3. $S_\text{T}(t_m, m_s)$ is repeated every G seconds. The received signal includes the time of flight delay $\tau(t)$ and can be given by

$$S_\text{R}(t_m, m_s) = \delta \cdot \cos\left(\varphi_\text{R}(t_m, m_s)\right), \tag{2.50}$$

with

$$\varphi_\text{R}(t_m, m_s) = 2\pi\left(f_c\left(t_m - \tau(t)\right) + \frac{1}{2}\dot{f}\left(t_m - \tau(t)\right)^2\right) - \varphi_{T0}.$$

Analogous to (2.10), $\tau(t)$ is a function of the initial range $r(0)$ and the radial velocity $\boldsymbol{v}_\text{r,rel}$ of the target and can be expressed using (2.49) as

$$\tau(t) = \frac{2r(0)}{c} + \frac{2\boldsymbol{v}_\text{r,rel}}{c}m_s G + \frac{2\boldsymbol{v}_\text{r,rel}}{c}t_m. \tag{2.51}$$

Note, that $\tau(t)$ changes steadily during a chirp because of $(2\boldsymbol{v}_\text{r,rel}/c)t_m$, and changes progressively from chirp to chirp because of $(2\boldsymbol{v}_\text{r,rel}/c)m_s G$. This approach makes it clear that a Doppler shift occurs both within a chirp and from chirp to chirp. However, the
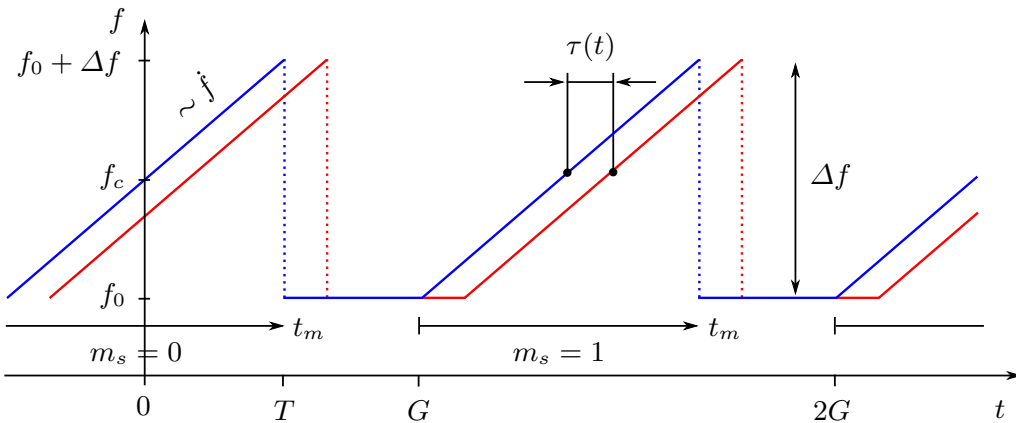

Figure 2.3: Representation of the chirp sequence modulation frequency characteristics of transmitted (——) and received (——) chirps over time [45].

influence of the Doppler within a chirp is negligible[1] due to the high frequency modulation gradient $\dot{f}$ of the range-dependent frequency component δf. Instead, it is exactly the rate of change of $\tau(t)$ from chirp to chirp that is used to determine the radial velocity of a target [27]. Due to the negligible Doppler shift within a chirp, it can be assumed that (2.40) is still valid and $\tau(t)$ is therefore directly proportional to the difference in frequency δf and hence to the distance r at any instant. Following the down-conversion to the base-band in (2.41), the two-dimensional IF-signal is (see appendix B)

$$S_{\mathrm{IF}}(t_m, m_s) = \delta \cdot \cos\left(\Delta\varphi(t_m, m_s)\right). \tag{2.52}$$

with

$$\Delta\varphi(t_m, m_s) = 2\pi\left(\left(f_c\frac{2\boldsymbol{v}_{\mathrm{r,rel}}}{c} + \dot{f}\frac{2r(0)}{c}\right)t_m + f_c\frac{2\boldsymbol{v}_{\mathrm{r,rel}}}{c}m_sG + f_c\frac{2r(0)}{c}\right). \tag{2.53}$$

Note, that the current frequency of a chirp is

$$f_c\frac{2\boldsymbol{v}_{\mathrm{r,rel}}}{c} + \dot{f}\frac{2r(0)}{c}, \tag{2.54}$$

whereas

$$f_c\frac{2\boldsymbol{v}_{\mathrm{r,rel}}}{c}m_sG, \tag{2.55}$$

changes the phase from chirp to chirp incrementally only due to the radial relative velocity. Finally,

$$f_c\frac{2r(0)}{c}, \tag{2.56}$$

is a constant phase term due to the initial distance $r(0)$. The occurring phase difference between consecutive chirps can be illustrated for a moving target with constant radial velocity. Suppose we hold t_m constant, and hence view $S_{\mathrm{IF}}(t_m, m_s)$ at the same value of t_m for all chirps. We get

$$S_{\mathrm{IF}}(t_m = k, m_s) = \delta \cdot \cos\left(f_c\frac{2\boldsymbol{v}_{\mathrm{r,rel}}}{c}m_sG + \Theta\right), \tag{2.57}$$

where

$$\Theta = f_c\frac{2r(0)}{c} + \left(f_c\frac{2\boldsymbol{v}_{\mathrm{r,rel}}}{c} + \dot{f}\frac{2r(0)}{c}\right)k, \tag{2.58}$$

---

1   The frequency present during a chirp is

$$\underbrace{f_c\frac{2\boldsymbol{v}_{\mathrm{r,rel}}}{c}}_{f_{\mathrm{D}}} + \underbrace{\dot{f}\frac{2r(0)}{c}}_{\delta f}$$

$\Rightarrow \quad f_{\mathrm{D}}/\delta f \approx 10^{-3}$ with $v_{\mathrm{r,rel}} = 20\,\mathrm{m/s}$, $f_c = 76.5\,\mathrm{GHz}$, $T = 60\,\mu\mathrm{s}$, $\Delta f = 1\,\mathrm{GHz}$, $r = 50\,\mathrm{m}$.

does not change from chirp to chirp. Figure 2.4 emphasizes the proportionality between the change of phase over M_s consecutive chirps and the radial relative velocity $\boldsymbol{v}_{\text{r,rel}}$. Assuming that the distance r of the target has not changed during one entire coherent integration period over all M_s chirps in such a way that the target migrates[1] into another range-dependent frequency cell, the complex amplitudes of (2.57) (depicted as blue arrows in Figure 2.4) of $S_{\text{IF}}(t_m = k, m_s)$ remain constant but rotate with frequency $f_c(2\boldsymbol{v}_{\text{r,rel}}/c)$ in the complex plane. Consequently, the measured phase difference between two consecutive chirps reads

$$\delta\varphi = 2\pi f_{\text{D}} G = 2\pi \left(f_c \frac{2\boldsymbol{v}_{\text{r,rel}}}{c} \right) G. \tag{2.59}$$

In order to separate two velocity contributions, $\delta\varphi$ has to satisfy

$$\delta\varphi > \frac{2\pi}{M_s}. \tag{2.60}$$

Thus, using (2.59) and (2.60), velocity resolution can be expressed as

$$\Delta\boldsymbol{v}_{\text{r,rel}} = \frac{c}{2 f_c M_s G}. \tag{2.61}$$

As can be seen from (2.61), the velocity resolution is inversely proportional to the total coherent integration time $M_s \cdot G$. Following, the time signal $S_{\text{IF}}(t_m, m_s)$ is digitized along t_m yielding discrete values of t_m. The sampled signal can be expressed by

$$S_{\text{IF}}(l_s, m_s) = \delta \cdot \cos\left(\Delta\varphi(l_s, m_s) \right), \tag{2.62}$$

with

$$\Delta\varphi(l_s, m_s) = 2\pi \left(\left(f_c \frac{2\boldsymbol{v}_{\text{r,rel}}}{c} + \dot{f} \frac{2r(0)}{c} \right) \frac{l_s}{f_{\text{s}}} + f_c \frac{2\boldsymbol{v}_{\text{r,rel}}}{c} m_s G + f_c \frac{2r(0)}{c} \right), \tag{2.63}$$

where f_{s} is the sampling frequency $S_{\text{IF}}(t_m, m_s)$ is sampled with, l_s is the sample index of t_m with $l_s = 0, \ldots, L_s-1$, and m_s is the sample index of the M_s chirps with $m_s = 0, \ldots, M_s-1$. According to the Nyquist-Shannon sampling theorem, $1/f_{\text{s}}$ and G determine the maximum resolvable distance and radial velocity of a target. The constraints are [27]

$$f_{\text{s}} > 2 \left(f_c \frac{2\boldsymbol{v}_{\text{r,rel}}}{c} + \dot{f} \frac{2r(0)}{c} \right), \tag{2.64}$$

---

1 This effect is called *range migration* and describes the target's change (migration) of range-dependent frequency cells (range gates) due to its velocity while the measurement takes place. Using the provided signal processing technique, where the phase behavior within a range gate over the entire measurement interval is used to determine the target's velocities, range migration can lead to wider peaks in the spectrum and hence to an erroneous velocity determinations of targets. The occurrence of range migration is proportional to the range and velocity resolution (decreasing the length of range gates and longer observation time), resulting in a continuous parameterization conflict [89].
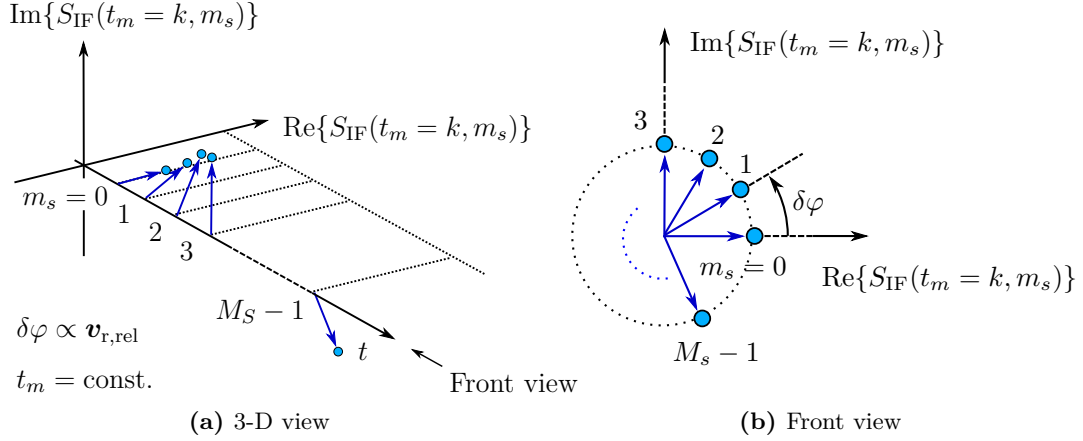
(a) 3-D view   **(b)** Front view

Figure 2.4: Scheme representation of the phase progression of a chirp sequence modulation at a constant r in the complex plane. The blue arrows symbolize the amplitudes of the complex numbers within one frequency cell rotating proportionally to a constant radial velocity of a target [4].

for the distance, and

$$\frac{1}{G} > 2f_c \frac{2\boldsymbol{v}_{\mathrm{r,rel}}}{c} \;\; \Rightarrow \;\; \boldsymbol{v}_{\mathrm{r,rel}} < \frac{1}{G} \frac{c}{4f_c}, \tag{2.65}$$

for the velocity, respectively.

With regard to automotive scenarios, and the trend towards growing resolution capabilities for the detailed environment perception, the chirp sequence modulation outperforms the classical FMCW approach and manifests the most efficient waveform design in current radar sensors [109].

2.2.3 Direction-of-arrival estimation

Besides radial distance and radial relative velocity, angular information of an incident wave reflected from a target is highly relevant for the entire perception of the vehicle's surroundings. In the previous subsection, we covered the radar measurement model for range and velocity for only one sensor element. Now, we extend the observation perspective to several sensor elements. Automotive radar antenna structures consist of at least one transmitting antenna and several receiving antennas. The objective is to determine the arriving angle of a spatially propagating signal concerning the radar antenna array. Direction-of-arrival (DOA) estimation formulates a significant aspect in array signal processing where the antennas measure a spatial field at various array locations to reveal quantitative and angular information about the sources. Here, we will focus on the concept of digital beamforming for DOA estimation. At first, a general model for the output signal of a receiving sensor array will be introduced. Then, a particular case of the model, the one-dimensional uniform linear array (ULA) is derived and its performance to beamforming is explained.

Consider an array of N_s receiving antennas at positions $\boldsymbol{p}_{n_s}$, $n_s = 1, \ldots, N_s$, and a radiating point source, located in the far-field of the sensor, at azimuth and elevation angle ϕ and ϵ, respectively. Suppose that the antenna array receives the signal $s(t)e^{j2\pi f_c t}$ emitted by the

source, where $s(t)$ denotes the complex modulating function. The corresponding geometry is visualized in Figure 2.5(a). The impinging wavefront is considered to be planar and can be described with the wavenumber vector [49], [91], [104]

$$\boldsymbol{k} = -\underbrace{\frac{2\pi}{\lambda}}_{\kappa} \begin{bmatrix} \sin(\epsilon)\cos(\phi) \\ \sin(\epsilon)\sin(\phi) \\ \cos(\epsilon) \end{bmatrix}. \tag{2.66}$$

The received signals at the individual antenna elements with inter-element distance d appear as delayed versions of the incoming signal with respect to the first reference element. The snapshot vector $\boldsymbol{x}_c(t) \in \mathbb{C}^{N_s \times 1}$ collects the receiving signals at the array elements according to

$$[\boldsymbol{x}_c]_{n_s} = s(t - \tau_{n_s})e^{j2\pi f_c(t - \tau_{n_s})}, \quad n_s = 1, \ldots, N_s, \tag{2.67}$$

where

$$\tau_{n_s} = \frac{1}{2\pi f_c}\boldsymbol{k}^T\boldsymbol{p}_{n_s}, \tag{2.68}$$

denotes the delay at sensor element n_s. Since we follow the narrow-band assumption, i.e., the bandwidth of $s(t)$ is narrow enough that the modulations stay almost constant during the propagation between the antenna elements, we can approximate $s(t - \tau_{n_s}) \approx s(t)$ [97]. The received signals are demodulated with a known carrier frequency resulting in the base-band array output vector

$$\boldsymbol{x}(t) = e^{-2\pi f_c t}\boldsymbol{x}_c(t) = s(t)\boldsymbol{a}(\phi, \epsilon), \tag{2.69}$$

where $\boldsymbol{a}(\phi, \epsilon)$ denotes the complex steering vector possessing the set of phase delays with elements

$$[\boldsymbol{a}(\phi, \epsilon)]_{n_s} = e^{-j\boldsymbol{k}^T\boldsymbol{p}_{n_s}}, \quad n_s = 1, \ldots, N_s. \tag{2.70}$$

Suppose that the antenna array receives K ($K < N_s$) impinging wavefronts from K different point sources with explicit DOA angles $\boldsymbol{\phi} = [\phi_1, \ldots, \phi_K]^T$, and $\boldsymbol{\epsilon} = [\epsilon_1, \ldots, \epsilon_K]^T$ with respect to the radar coordinate system, the general signal model for the sampled base-band snapshot vector at the l_sth of L_s available sample instants in noise reads [49]

$$\boldsymbol{x}(l_s) = \sum_{k=1}^{K} s_k(l_s)\boldsymbol{a}(\phi_k, \epsilon_k) + \boldsymbol{e}(l_s) = \boldsymbol{A}(\boldsymbol{\phi}, \boldsymbol{\epsilon})\boldsymbol{s}(l_s) + \boldsymbol{e}(l_s), \quad l_s = 1, \ldots, L_s, \tag{2.71}$$

where

$$\boldsymbol{A}(\boldsymbol{\phi}, \boldsymbol{\epsilon}) = [\boldsymbol{a}(\phi_1, \epsilon_1), \ldots, \boldsymbol{a}(\phi_K, \epsilon_K)] \in \mathbb{C}^{N_s \times K}, \tag{2.72}$$

is the array steering matrix whose columns contain the K steering vectors of unknown DOAs, the vector $\boldsymbol{s}(l_s) \in \mathbb{C}^{K \times 1}$ contains the source waveforms, and $\boldsymbol{e}(l_s) \in \mathbb{C}^{N_s \times 1}$ is an additive vector of white randomly modeled process noise covering external diffuse noise sources in the hardware.

Note, that (2.70) is valid for arbitrary array geometries. However, most automotive radar sensors possess a ULA for one-dimensional DOA estimation. A ULA in y-direction is

depicted in Figure 2.5(b) with $\phi \in [-\pi/2, \pi/2)$ and $\epsilon = \pi/2$. Using spatial frequency $\psi_y = \kappa d \sin(\phi)$, the steering vector then becomes [49]

$$\boldsymbol{a}(\phi) = \boldsymbol{a}(\psi_y) = \boldsymbol{a}_y(\psi_y), \tag{2.73}$$

with

$$\boldsymbol{a}_y(\psi_y) = \left[1, e^{j\psi_y}, \ldots, e^{j(N_s-1)\psi_y}\right]^T \in \mathbb{C}^{N_s \times 1}. \tag{2.74}$$

The phase shift's spatial sampling across all antenna elements results in a sinusoidal signal whose frequency is proportional to ϕ. Beamforming describes the concept of conventional non-parametric DOA estimation, where the array output vector is used to calculate the signal components of the particular directions. A spatial matched filter $\boldsymbol{a}(\phi)^H \boldsymbol{x}(l_s)$, the beamformer, coherently combines the antenna outputs to the hypothetical angle ϕ to create a spatial power spectrum [49]

$$P_{\mathrm{BF}}(\phi) = \frac{1}{L_s N_s} \sum_{l_s=1}^{L_s} |\boldsymbol{a}(\phi)^H \boldsymbol{x}(l_s)|^2, \tag{2.75}$$

whose power distribution reveals the DOA estimates. In the case of ULA, the inner product in (2.75) is equivalent to a spatial Fourier transform. Thus, a discrete Fourier transform (DFT) can be used to calculate the frequency of the spatially sampled phase propagation ψ. Among the established DOA estimation techniques, beamforming represents a computationally simple approach but is limited in its resolution. Considering automotive frequencies, the antenna aperture N_A, i.e., the overall space occupied by the antenna array
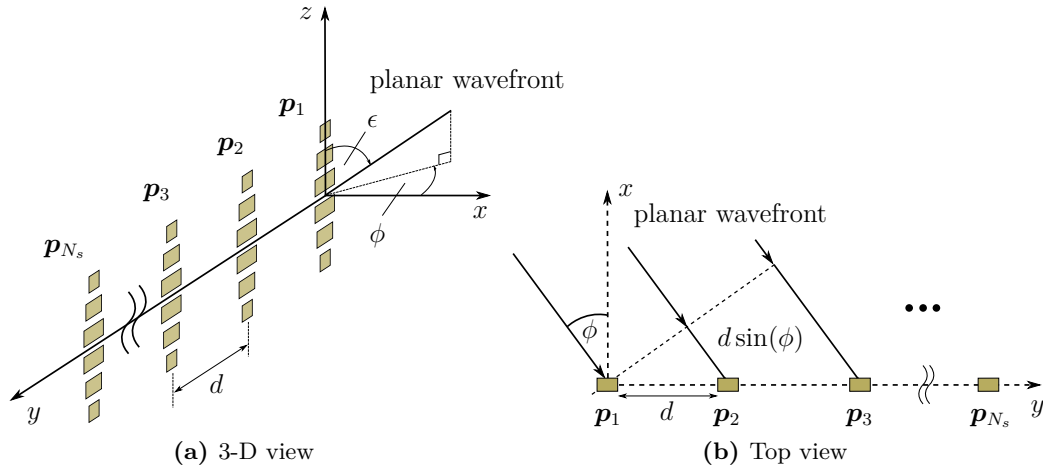

(a) 3-D view   **(b)** Top view

Figure 2.5: Uniform linear array geometry for azimuthal DOA estimation. Planar wavefront is impinging on antenna elements at positions $\boldsymbol{p}_{n_s}$ ($n_s = 1, \ldots, N_s$) with azimuth angle ϕ and elevation angle ϵ, respectively.

in the considered direction, determines the angular resolution $\Delta\phi_{\mathrm{BF}}$ according to

$$\Delta\phi_{\mathrm{BF}} \approx \frac{\lambda}{N_{\mathrm{A}}}. \tag{2.76}$$

As can be seen from (2.76), a larger aperture is desirable. This can be achieved by either increasing the number of antennas or the element spacing. However, we derive from (2.74) that $\boldsymbol{a}(\phi)$ is uniquely defined, i.e., without spatial aliasing, only if ψ holds:

$$|\psi| \leq \pi. \tag{2.77}$$

As a consequence, (2.77) is tantamount to

$$d|\sin(\phi)| \leq \frac{\lambda}{2} \ \Rightarrow\ d \leq \frac{\lambda}{2} \quad \text{for } \phi \in [-\pi/2, \pi/2), \tag{2.78}$$

From (2.78), it gets apparent that the spatial sampling period d ought to be smaller than half the wavelength in order to avoid spatial aliasing. These aliasing effects are expressed by grating lobes, which appear periodically to the inverse of the element spacing with equal gain to the main lobe, causing directional ambiguity. On the other side, an increase in the number of elements comes with higher cost and computational demand, resulting in a tradeoff regarding element spacing [97].

2.3 Signal processing

In the previous subsections, the radar measurement model based on chirp sequence modulation and DOA is derived and introduced. The combined overall signal model for a coherent processing interval (CPI), i.e., the total measurement cycle time to be sampled, can be expressed as a superposition of K discretized 3-D complex sinusoids in noise [31]

$$\boldsymbol{x}(l_s, m_s, n_s) = \sum_{k=1}^{K} a_k e^{j(\lambda_k l_s + \mu_k m_s + \nu_k n_s)} + \xi(l_s, m_s\, n_s), \tag{2.79}$$

for $l_s = 0, \ldots, L_s - 1$, $m_s = 0, \ldots, M_s - 1$, and $n_s = 0, \ldots, N_s - 1$, where L_s, M_s, and N_s indicate the number of samples per chirp, the number of chirps, and the number of receiving antennas, respectively. a_k is the complex amplitude of the kth target and ξ denotes Gaussian distributed circular complex noise. λ_k, μ_k, and ν_k denote the radian frequency parameters with

$$\lambda_k \simeq 2\pi \left(f_c \frac{2\boldsymbol{v}_{\mathrm{r,rel},\,k}}{c} + \dot{f} \frac{2r(0)_k}{c} \right) \tag{2.80}$$

$$\mu_k \simeq 2\pi \left(f_c \frac{2\boldsymbol{v}_{\mathrm{r,rel},\,k}}{c} \right) G \tag{2.81}$$

$$\nu_k \simeq \kappa d \sin\left(\phi_k\right), \tag{2.82}$$

that correspond to range, radial relative velocity, and azimuth angle, respectively. The constant phase term in (2.56) is neglected. A practical approach for estimating the spectral components of the digitized time signals in (2.79) is the successive use of three DFTs[1] resulting in the 3-D spectrum [31]

$$\boldsymbol{X}\left(\lambda,\mu,\nu\right) = \sum_{l_s=0}^{L_s-1}\sum_{m_s=0}^{M_s-1}\sum_{n_s=0}^{N_s-1} \mathbf{w}_\lambda(l_s)\mathbf{w}_\mu(m_s)\mathbf{w}_\nu(n_s) \tag{2.83}$$

$$\times\,\boldsymbol{x}\left(l_s,m_s,n_s\right)e^{-j(\lambda l_s+\mu m_s+\nu n_s)} \tag{2.84}$$

for $\lambda = (2\pi/L)l$, $l = 0,\dots,L-1$, $\mu = (2\pi/M)m$, $m = 0,\dots,M-1$, and $\nu = (2\pi/N)n$, $n = 0,\dots,N-1$, where L, M, and N denote the number of frequency samples in range dimension, Doppler dimension, and angular dimension, respectively. $\mathbf{w}_\lambda(l_s)$, $\mathbf{w}_\lambda(l_s)$, and $\mathbf{w}_\lambda(l_s)$ are normalized window functions[2] in range dimension, Doppler dimension, and azimuth dimension, respectively.

Note that the summations in (2.83) imply that the order of the DFT's is irrelevant and hence exchangeable. Typically, the range-DFT along l_s is performed first to determine the objects' distance-dependent frequencies in the radar's field of view. Due to the hermitian symmetry of the real-valued intermediate frequency signal $S_{\text{IF}}(l,m)$, the DFT is completely specified by $L/2-1$ complex coefficients, and hence the redundant half of the range-dependent spectrum can be removed for further processing. The frequency vector obtained can be assigned to distances using (2.45) and thus provide range gates. A second DFT applied orthogonally to the first across all chirps, yields the Doppler frequencies, and completes the 2-D range-Doppler spectrum. This procedure is done for each of the N_s receiving antennas individually. Consequently, the third DFT combines the individual receiving antennas' independent measurement spaces by mapping the range-Doppler spectra to several directions through digital beamforming, as introduced in Subsection 2.2.3. Since (2.83) is a 3-D frequency spectrum, multiple targets within a range gate can be separated at different radial velocities and azimuth angles.

2.3.1 Target properties

The continuous development of microelectronics and antenna structures used in automotive radar sensors influences their resolution capacity. The level of resolution determines to what extent an object is detected concerning its level of detail. Depending on the measured geometric and kinematic properties, objects can be assigned to different expansion

---

1   The DFT is typically calculated using a fast Fourier transform (FFT). The FFT represents a computationally efficient algorithm that reduces the N^2 complex multiplications and additions of a DFT to $\frac{1}{2}N\log_2 N$ by using a radix-2-FFT implementation, given a sequence of length N [97].

2   The use of window functions is necessary to reduce the occurrence of side lobes emerging from the spectral leakage effect due to possible discontinuities at the signals' ends [46]. However, the minimization of side lobes is always accompanied by a broadening of the main lobe in the spectrum. This can lead to a spreading of the power to neighboring range gates and hence can affect the effective range resolution and the detectability of weaker spectral components [27], [46]. The intensity of the main lobe widening depends on the selected window function.

types. In the following, four different extension types are distinguished and defined [36], [63].

- **From point target to extended target**:
  A point target is an object, where the physical dimensions in all considered dimensions (radial distance, Doppler velocity, azimuth angle, elevation angle) are smaller than the corresponding resolution cells. Consequently, the object is perceived as a point target, and the radar cannot determine the actual physical dimension. High-resolution radar sensors are characterized by the fact that objects are no longer detected as point targets but appear as extended targets. This is particularly true for near-field radars where the reduced maximum range and Doppler velocity is compensated with higher resolution. As described in Subsection 2.2.1, the bandwidth of the transmitted chirp defines the range resolution in FMCW radars. However, the IF signal and hence the range estimation accuracy are further influenced by various aspects, such as phase noise, SNR, and especially non-linearities in the frequency modulation as they lead to a broadend peak in the base-band signal [6]. Typically, extended targets have several scattering points, meaning that an object's physical extent is larger than the cell resolutions in at least one dimension. Thereby, an object can occupy several resolution cells in several dimensions, as illustrated in Figure 2.6 using the example of a pedestrian. However, due to the measurement principle of range, Doppler, and angle, and due to the range-dependent SNR, the initially detected extension of an object may change during a movement and may also change from an extended target to a point target and vice versa [36], [63]. Consequently, whether an object appears to be extended depends not only on the physical size of the object, but also on the ratio of the physical size to the sensor resolutions.

- **Geometric extent**:
  If the radar's resolution in angle or range dimension is larger than the physical extent of the object at the object position, a geometric extent exists. The geometric extent is depicted in Figure 2.6(b).

- **Kinematic extent**:
  If an object exhibits several measurable velocities in one range or angle cell, this object is considered kinematically extended. The kinematic extent can be caused by the μD effect, which is defined in Subsection 2.1.3, or by differently measured radial velocities within an angle cell. The kinematic extent is depicted in Figure 2.6(c).

- **Double extent**:
  If an object shows both a kinematic and a geometric extent, a double extent exists leading to a detection area in the r-$\boldsymbol{v}_{\mathrm{r,rel}}$-plane or the ϕ-$\boldsymbol{v}_{\mathrm{r,rel}}$-plane or ϵ-$\boldsymbol{v}_{\mathrm{r,rel}}$-plane, respectively. This means that the object's reflection is detected in at least two different velocity cells and angle or distance cells, respectively. The double extent is depicted in Figure 2.6(d).

2.3.2 Target extraction

With the signal processing provided in the last section, the discretized time signals have been transformed into 3-D frequency space. The data structure arrangement can be regarded as a 3-D data cube of size $L/2 - 1 \times M \times N$. The data cube comprises the raw data signals in three dimensions, including random noise and clutter effects. Clutter refers to any object that may induce undesired reflections, such as environmental influences, vegetation, and ground terrain and hence interfere with regular radar returns in the IF signal [70]. However, concerning practical aspects, it is not necessary to generate and store the entire data cube for a CPI to extract potential targets. High-resolution radar parametrizations quickly exceed hardware storage capacities so that alternative methods are needed. Instead, a more efficient calculation sequence is used, where power detection, i.e., the separation of power amplitudes from the noise level, is carried out in the range-Doppler spectrum. For further processing, only the power-detected range-Doppler cells are considered for angular frequency estimation. The objective is to create a detection list per measurement cycle, in which the properties of detected cells, e.g., range r, Doppler velocity[1] v^D, azimuth angle ϕ, and target parameters like RCS and SNR, are stored. The detection list lays the foundation for various further processing steps such as object tracking or fusion methods with comparable variables of other environment detection sensors.

Power detection

For power detection, the magnitudes of the individually calculated complex range-Doppler spectra are averaged over all N_s antennas by means of non-coherent integration according to [85]

$$|\bar{\boldsymbol{X}}(\lambda, \mu)|_{N_s} = \frac{1}{N_s} \sum_{n_s=1}^{N_s} |\boldsymbol{X}(\lambda, \mu)|_{n_s}. \tag{2.85}$$

By doing so, the SNR is significantly increased, which helps to distinguish between noise and actual target detections and hence improves the detection performance. The accumulated matrix provides the foundation for the power detection procedure. Due to non-homogeneous noise and interference statistics, a global threshold may cause various wrong target detections and is not suitable for power detection. Accordingly, an adaptive threshold value must be determined depending on the local environment of a cell of interest. Constant false alarm rate (CFAR) filters are supposed to keep the false positive rate in the measurement space constant and still detect all targets. This is especially true for situations where severe clutter occurs, e.g., clutter due to rain or multipath reflections. The ordered-statistics constant false alarm rate (OS-CFAR) represents an efficient procedure to remove noise from detections in dense multi-target and extended target scenarios where classical cell averaging CFAR filters exhibit negative performance [88].

---

1 Due to the classification of every range gate's velocity spectrum into cells of different power amplitudes, the variable $\boldsymbol{v}_{\mathrm{r,rel}}$ will be substituted with v^D for detected Doppler cells from now on to underline the identification.

(a) Point target

(b) Geometric extent

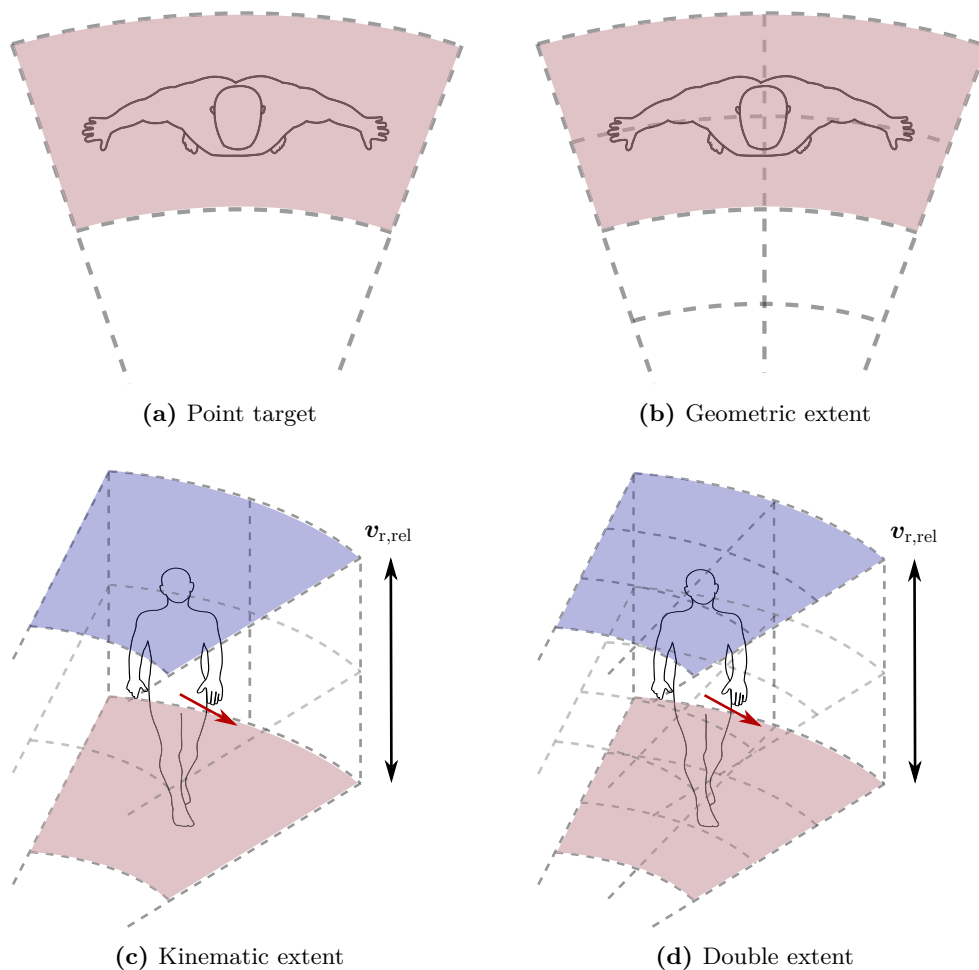**(c)** Kinematic extent

(d) Double extent

Figure 2.6: Overview of the different types of extents depending on the resolution [63].

The principle of the one-dimensional OS-CFAR filter in Doppler direction is introduced in the following. Suppose $\boldsymbol{X}(l,m) \in |\bar{\boldsymbol{X}}(\lambda,\mu)|_{N_s}$ for $m = 1, \ldots, M$ spans the M reference cells of the lth range gate. Note that $\boldsymbol{X}(l,m)$ can comprise random noise as well as target emerging Doppler amplitudes. Then, in OS-CFAR, the cells of $\boldsymbol{X}(l,m)$ are sorted in ascending order according to their amplitude yielding a new sequence

$$X_{(1)} \leq X_{(2)} \leq \cdots \leq X_{(k_{\mathrm{CFAR}})} \leq \cdots \leq X_{(M)}. \tag{2.86}$$

From this ordered statistic, the amplitude of a single cell $X_{(k_{\mathrm{CFAR}})}$ with rank k_{CFAR} is selected and serves as an average estimate of the noise level in the reference window. The detection threshold is then given by

$$\varrho = X_{(k_{\mathrm{CFAR}})} \cdot \alpha. \tag{2.87}$$

The scaling factor α can be derived according to

$$\alpha = (P_{fa})^{-1/k_{\mathrm{CFAR}}} - 1, \tag{2.88}$$

where P_{fa} is an assumed false alarm rate distribution [88]. For the optimal determination of k_{CFAR}, it must be taken into account that sudden clutter does not lead to a temporarily increased P_{fa}. Moreover, low target emerging amplitudes should be detected despite the presence of strong targets in the reference cells. A typical value for k_{CFAR} is given in literature with $0.6M \leq k_{\mathrm{CFAR}} \leq 0.85M$ [69]. Finally, each cell whose amplitude is below ϱ is dropped, whereas cells with amplitudes above ϱ are considered to be detections and are kept for further processing. This procedure is now repeated for every range gate which completes the OS-CFAR filtering in the Doppler dimension.

Note that the provided OS-CFAR approach, which is used in this work for power detection, considers all Doppler cells of a range gate as reference cells to determine the threshold. By doing so, only one threshold is specified for the data sequence of one range gate, which reduces the computational effort compared to the sliding window technique in classical CFAR filtering [88].

Clustering

After OS-CFAR filtering, the detected power amplitudes are available without any specific target assignment. Hence, procedures are required that associate single detections to an unknown number of different targets and identify leftover noise amplitudes simultaneously. Clustering describes the management of spatially extended data by assigning single detections to different groups based on their similarity. By grouping single detections to a cluster, an object is created. Especially with a view to the optimization of subsequent tracking algorithms and object parameter estimation, closed formed objects per measurement cycle are indispensable.

Conventionally, clustering algorithms can be divided into partitional and hierarchical approaches. The partitional clustering approach iteratively partitions a database into a
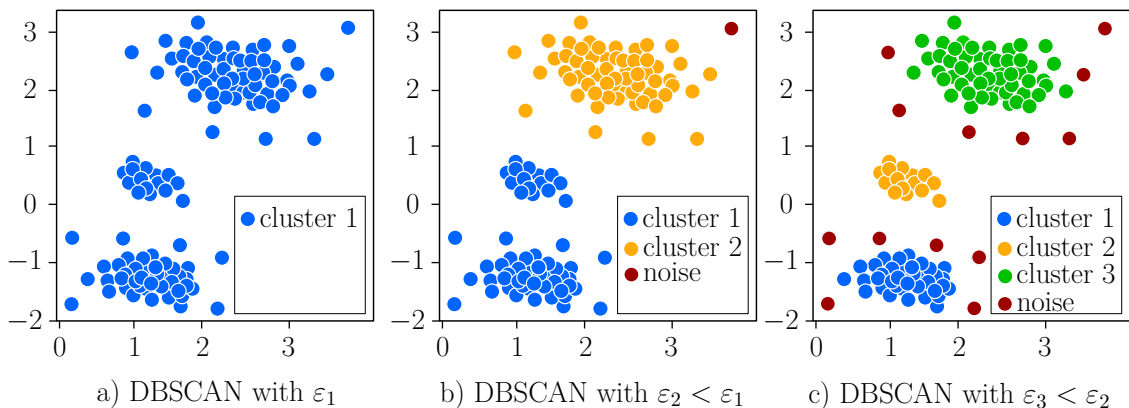

Figure 2.7: Principle of the DBSCAN clustering algorithm with different radii ε and constant density threshold.

set of clusters. Here, an input variable indicating the number of clusters to be expected has to be determined a priori. Contrary, hierarchical approaches do not require upfront domain knowledge. However, a termination condition for the decomposition process must be derived in advance [33]. For automotive clustering applications, where both the number of objects and their shape are usually unknown, classical clustering algorithms do not provide adequate performance. A more intuitive way of grouping data is the density-based method. A well-known representative is the density-based spatial clustering of applications with noise (DBSCAN) used in this work. The idea in DBSCAN is to define a cluster based on the density of a point's neighborhood. The density criterion can be formulated as follows: "for each point of a cluster the neighborhood of a given radius has to contain at least a minimum number of points, i.e., the density in the neighborhood has to exceed some threshold" [33]. Every point that fulfills this criterion or is at least inside the search area is considered a core point. All other points denote outliers and are considered to be noise. The choice of the radius ε and the density threshold significantly influence the clustering process. Figure 2.7 exemplary illustrates the DBSCAN clustering procedure for different radii while keeping the minimum number of data points constant. If ε is chosen too large, different targets might be combined into a single cluster, including potential noise leftovers, whereas a small ε tends to label target emerging detections as noise erroneously. Concerning targets that exhibit μD components, the correct identification can be particularly challenging.

2.3.3 Real radar data example

Figure 2.8 visualizes the entire signal processing, including OS-CFAR filtering and DBSCAN clustering of real high-resolution radar data, and presents the object's detection list, $D = \{r, v^D, \phi, \sigma\}$, carrying information about range, Doppler, azimuth angle and RCS. Here, a snapshot of a moving bicycle is used exemplary. The bicycle's front wheel is located at a radial distance of about $12.5\,\mathrm{m}$, and the bicycle exhibits a geometric extent of about $1.5\,\mathrm{m}$ in this particular snapshot. The bicycle's metallic frame can be identified due to the highest power reflections in the antenna averaged range-Doppler periodogram (see Figure 2.8(a)), and shows a radial translational velocity, which is slightly below $3\,\mathrm{m/s}$. Moreover, the μD amplitudes emerging from the rotating wheels are clearly visible and range from $0\,\mathrm{m/s}$ to about $6\,\mathrm{m/s}$ for the front and back wheel, respectively. OS-CFAR filtering removes noise from locally high power amplitudes, i.e., detections (see Figure 2.8(b)). Subsequent DBSCAN clustering groups the detections to objects and noise, respectively. Figure 2.8(c) shows the largest remaining object corresponding to the bicycle. Based on the power-detected processing cells of this object, angle estimation is calculated using a third DFT over all antennas according to (2.83). Finally, Figure 2.8(d)) illustrates the angular resolved peak frequencies in three dimensions of the moving bicycle.
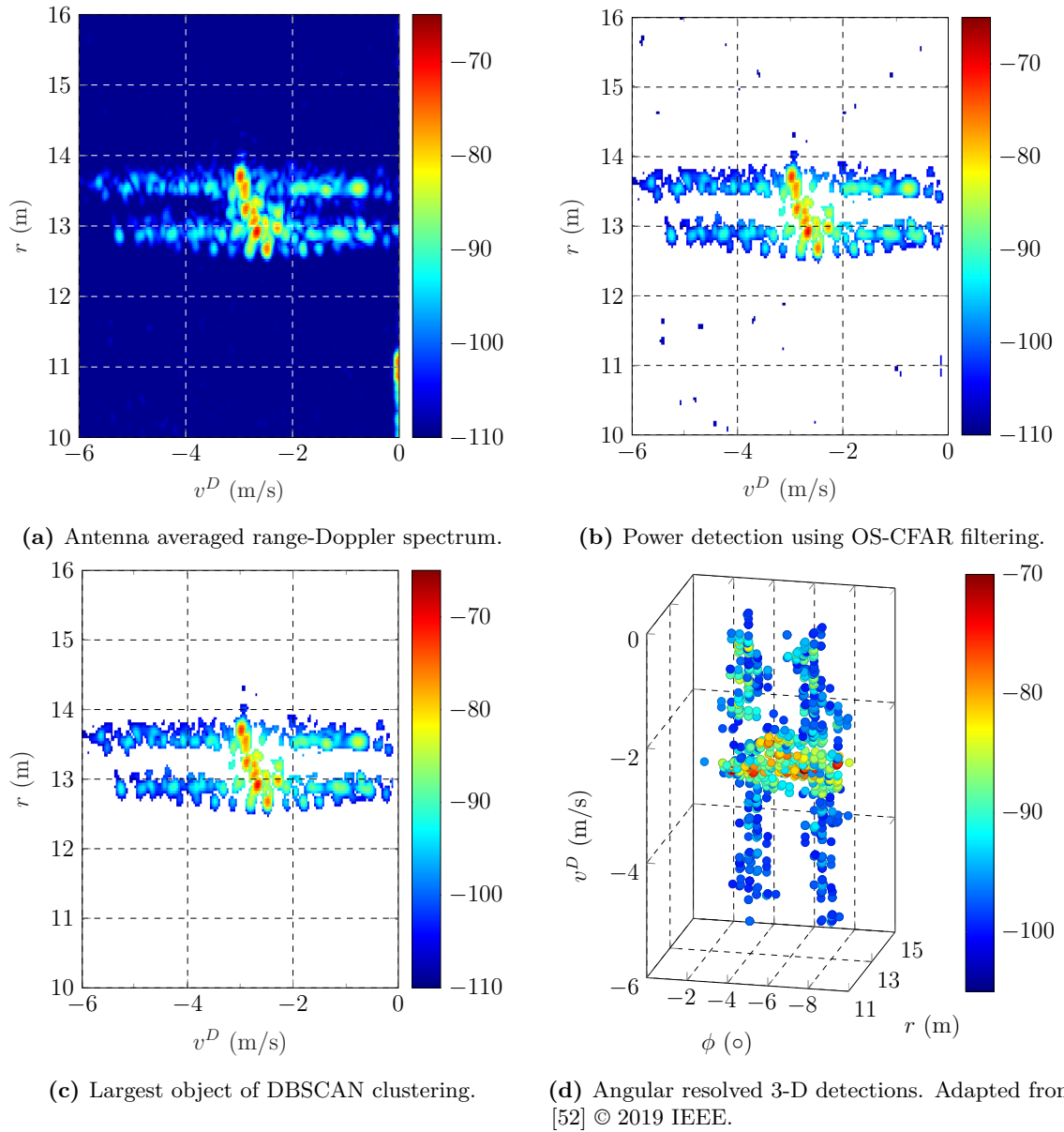
(a) Antenna averaged range-Doppler spectrum.

(b) Power detection using OS-CFAR filtering.

(c) Largest object of DBSCAN clustering.

(d) Angular resolved 3-D detections. Adapted from [52] © 2019 IEEE.

Figure 2.8: Visualization of the provided signal processing steps including OS-CFAR filtering and DBSCAN clustering using a real radar data snapshot of a moving bicycle yielding a range-Doppler-azimuth detection list. The logarithmized backscatter intensities are in dBV.

2.4 Conclusion

This chapter presents automotive radar's essential operation, including the underlying physical effects such as Doppler and μD effect. Starting with the propagation behavior of directed electromagnetic waves and the RCS introduction, the frequency modulation principle for radial distance and radial relative velocity determination is explained. An optimized procedure, the chirp sequence modulation, which compensates the classic FMCW procedure's weaknesses in multi-target scenarios, is derived and considered in detail. DOA estimation is introduced to determine a reflection's orientation by evaluating the occurring phase shift across all antenna elements, which completes the radar measurement model. Potential targets are extracted from the 3-D measurement space utilizing OS-CFAR thresholding and DBSCAN clustering procedure showing measured detections for the distance r, the Doppler velocity v^D, and the azimuth angle ϕ, respectively. The concept of extended targets is introduced, and its spatial and kinematic extensions are defined. In the end, the entire signal processing chain is applied to real high-resolution radar data for an approaching bicycle that reveals the μD phenomenon of the rotating wheels.

CHAPTER 3

Micro-Doppler applications of a cyclist

3.1 Physical fundamentals

This section addresses the physical principles of the μD appearance of a cyclist and the associated challenges regarding safety-relevant signal processing techniques. After analyzing and illustrating the primary sources of μD phenomena using real radar raw data, the influence of the relative orientation dependence on the radar signals is considered in detail.

3.1.1 Micro-Doppler signatures of a cyclist

The ability of radar sensors to utilize the μD effect is crucial for a detailed interpretation of the vehicle's surroundings. Besides linear motions, it is predominantly the additive frequency modulations that give rise to a thoroughly new perspective of radar-based VRU detection. In cyclists, μD velocity distributions are particularly pronounced due to wheel spinning and pedal rotations becoming a permanent feature in sophisticated safety functions. However, the research field of bicycle detection and μD utilizing automotive radar approaches is still relatively new and hence insufficiently investigated.

From a backscattering point of view, it can be shown that the typically metallic frame of a bicycle generates the most extensive reflections. While this is especially true for longitudinal directions, the wheels' reflectivity, particularly emerging from the rims and spokes, increases significantly for other orientations and contributes to the overall perceptible RCS [18].

77 GHz analyses of a jacked-up bicycle's μD signatures highlight the contributing velocity components, i.e., spokes rotation and leg pedaling captured in different orientations [16]. In [94], the first range-Doppler spectra of radially and laterally moving cyclists are presented, displaying the wheels' characteristic velocity distributions with the rotating pedals in between. Considering a spectral view, the rotating wheels and pedals possess the most relevant and distinct μD components of a cyclist and can be split into three spectral groups: the front wheel denoted $\boldsymbol{X}(\lambda_{\mathrm{fw}}, \mu_{\mathrm{fw}})$, the rear wheel denoted $\boldsymbol{X}(\lambda_{\mathrm{rw}}, \mu_{\mathrm{rw}})$, and the pedaling motion denoted $\boldsymbol{X}(\lambda_{\mathrm{p}}, \mu_{\mathrm{p}})$. The converted pedal movement causes the rolling of a wheel that can be viewed as pure rotation. In doing so, a wheel's motion exhibits both a rotational velocity component $\boldsymbol{v}_{\omega} = \omega_{\mathrm{w}} \times r_{\mathrm{w}}$, which is defined as the cross-product of the angular velocity, ω_{w}, and the wheel's radius, r_{w}, and a constant translational velocity component, $\boldsymbol{v}_{\mathrm{ego}}$, as depicted in Figure 3.1(a). If we assume a non-slipping wheel having contact with the substrate, the total measurable velocity $\boldsymbol{v}_{\mathrm{total}}$ is the vectorial sum of the linear velocities [44]

$$\boldsymbol{v}_{\mathrm{total}} = \boldsymbol{v}_{\omega} + \boldsymbol{v}_{\mathrm{ego}}. \tag{3.1}$$

As can be seen from Figure 3.1(a), $\boldsymbol{v}_{\text{total}}$ comprises instantaneous velocities from zero at the ground up to $2 \cdot \boldsymbol{v}_{\text{ego}}$ at the wheel's highest point. Under the assumption of a static radar and a radially approaching bicycle, the corresponding range-Doppler spectrum would possess a line-shaped broadening in the Doppler dimension. In different orientations, the scatterers are distributed along with the wheel and arise from varying distances to the radar. Thereby, the wheels' μD components, $\boldsymbol{X}(\lambda_{\text{fw}}, \mu_{\text{fw}})$, and $\boldsymbol{X}(\lambda_{\text{rw}}, \mu_{\text{rw}})$, assume elliptical shapes in the spectrum, as shown in Figure 3.1(b). The pedaling denotes a cyclist's third significant μD component besides the wheels' velocity distributions, which occurrs as a periodically oscillating signal between the wheels [16], [94]. The pedaling motion is considerably affected by the rotating pedals and the knees of the cyclist. While the pedals produce the largest amplitudes in Doppler and RCS dimensions, the knees occur as superimposed motions with attenuated amplitudes. The knees' signals also exhibit a temporal phase shift depending on the cyclists' biomechanical properties [44].

3.1.2 Orientation dependence

The chance of employing μD components of a cyclist for the specific extraction of safety-relevant parameters, such as the wheels' velocity distributions or the pedaling characteristic, highly depends on the relative orientation between radar sensor and object due to the radar's radial measuring principle. As a result, signal ambiguity arise that may greatly problematize both the classification and behavioral indications' perception [52].

Certainly, lateral motions, which are of significant relevance concerning accident prevention with VRU, complicate the μD information's qualitative usability. Figure 3.2(a) illustrates
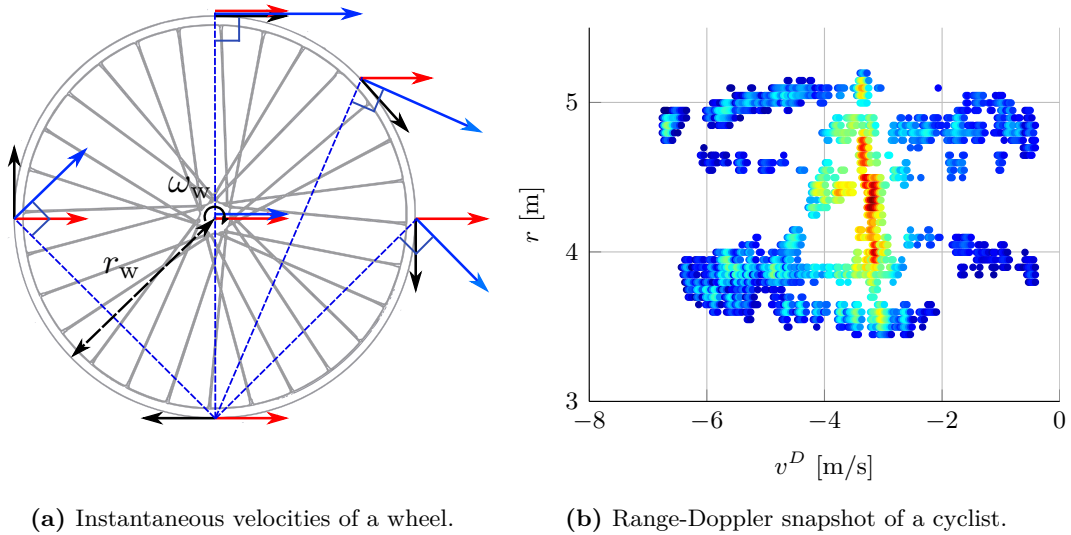

(a) Instantaneous velocities of a wheel.   **(b)** Range-Doppler snapshot of a cyclist.

Figure 3.1: Composition of a bicycle's μD components caused by the velocity distribution of the wheels and pedalling. (a) Instantaneous linear velocities of single locations on a clockwise turning wheel having contact to the ground with $\boldsymbol{v}_\omega$ (black arrows), $\boldsymbol{v}_{\text{ego}}$ (red arrows), and $\boldsymbol{v}_{\text{total}}$ (blue arrows). (b) High-resolution range-Doppler snapshot of a radially towards the sensor moving bicycle. The color-coding depicts the backscattering intensities. © 2019 IEEE [51].

the problem of ambiguous reflections. As can be seen, parts of the front wheel and the left pedal and parts of the rear wheel and the right pedal possess equal radial distance and are subject to an orientation-related overlap. Thus, various reflections are allocated to the same range gate yielding overlapping Doppler components. Besides, the radar perceives generally lower Doppler amplitudes due to the orientation angle compared to a fully radial motion. Consequently, both the geometric and the kinematic extent of an object may decrease during the approaching process. Figure 3.2(b) presents the simulated results of a laterally approaching cyclist highlighting the orientation dependence. Here, a bicycle with a length of 2 m moves in a lateral direction to the sensor. The starting position of the front wheel possesses the coordinates $y = 20$ m and $x = 10$ m according to the radar's coordinate system given in Figure 3.2(a). The blue curves in the diagram describe the radial distances of the foremost (solid curve) and the rearmost (dashed curve) location of the bicycle to the sensor while moving a route of 20 m. The red curve depicts the discrepancy in magnitude between the blue curves. After starting with an initial discrepancy of approximately 1.8 m, it decreases steadily under a nonlinear behavior until reaching a value of around 0.2 m at the end of the traveled route. Furthermore, the same reference points' radial velocities are displayed based on a constant translational velocity of 4 m/s. The severe reduction of their magnitudes can be distinctly caught. It is apparent that the orientation dependence influences the shape of the object in both dimensions and, particularly in lateral movements, leads to a sharp deterioration that may affect subsequent signal processing steps.
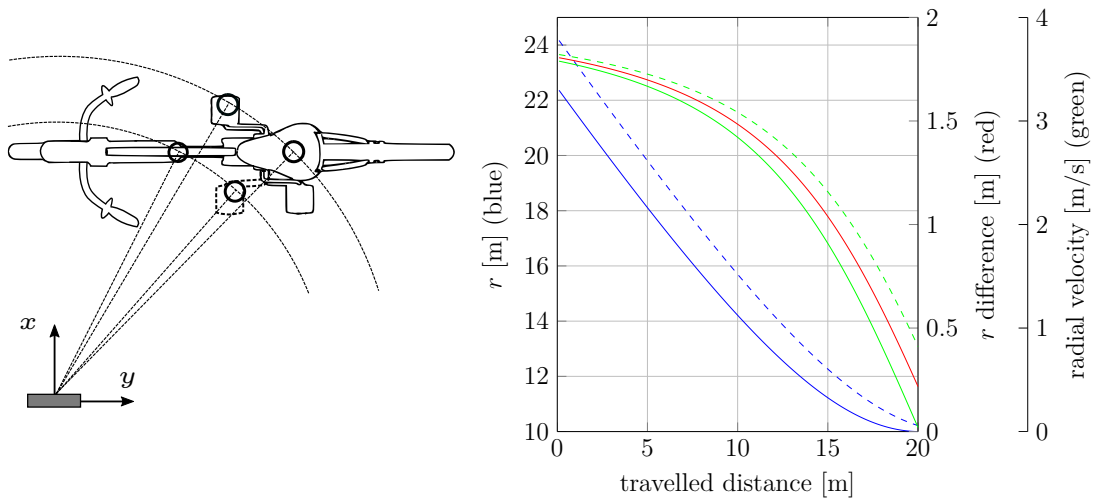
(a) Orientation dependent ambiguities.

(b) Geometric and kinematic behavior.

Figure 3.2: Illustration of the realtive orientation's influence on radar signals using a laterally moving bicycle. (a) Occurring range ambiguities at different locations on the bicycle. (b) Simulated results of the laterally moving cylist's orientation dependency. (──) and (- - -) display the radial distances of the foremost and rearmost locations of a bicycle. (──) is the difference of the blue lines. (──) and (- - -) represent the radial velocities of the foremost and rearmost locations of the bicycle, respectively. © 2019 IEEE [52].

3.2 Cyclist feature extraction

The preceding subsections introduce the prominent μD distributions of a moving cyclist caused by the wheel and pedal rotations and analyze their relation to orientation. It is notably the pedaling movement that may be of exceptional significance for establishing cyclists' radar-based behavior prediction. It is legitimate to assume that changes of the pedaling frequency reveal imminent changes of the cyclist's state and can be detected faster than ordinary acceleration or deceleration processes, which are only perceivable over a sequence of sensor cycles.

Since the shaft that carries the pedals is near the rear wheel, the μD components, $\boldsymbol{X}(\lambda_{\mathrm{p}}, \mu_{\mathrm{p}})$, and $\boldsymbol{X}(\lambda_{\mathrm{rw}}, \mu_{\mathrm{rw}})$, may interfere. Additionally, the span between the wheels' ellipsoidal velocity distributions, $\boldsymbol{X}(\lambda_{\mathrm{rw}}, \mu_{\mathrm{rw}})$, and $\boldsymbol{X}(\lambda_{\mathrm{fw}}, \mu_{\mathrm{fw}})$, relies on the bicycle and the radar orientation and can change while driving. Accordingly, an extraction procedure based on a designated part of $\boldsymbol{X}(\lambda, \mu)$, where the pedaling is presumed to arise, is not expedient. What would likely happen is that either not all detections belonging to the pedal motion are extracted, or accessory detections that do not belong to the pedal motion, such as $\boldsymbol{X}(\lambda_{\mathrm{fw}}, \mu_{\mathrm{fw}})$, and $\boldsymbol{X}(\lambda_{\mathrm{rw}}, \mu_{\mathrm{rw}})$, would be extracted.

For those reasons, we are striving for an adaptive pedaling extraction procedure capable of separating the wheels' μD components in each cycle to reveal the pedaling contribution. The use of a robust ellipse approximation appears to be a practical method for localizing the wheels and hence for the pedaling extraction objective and is demonstrated in the following.

3.2.1 Adaptive pedaling extraction

The following application's purpose is to identify and localize the wheels' μD distributions using an adaptive ellipse fitting approach in order to reveal and extract the μD components that comprise the motion of the pedals. The wheels' extended velocity distributions that may partially interfere with the pedaling components require a stringent and adaptive constraint. The following approach, which is published in [51], aims to utilize a nonlinear least-squares technique that fits ellipses to the wheels' μD extensions in range and Doppler to achieve a segmentation from the pedaling area. A robust approximation is achieved by performing the least-squares minimization iteratively within a random sample consensus (RANSAC) structure.

Ellipticity constraints

A conic section in a plane can be represented as the locus of coordinates $\mathbf{z} = (z_1, z_2)^T$ satisfying [79], [100], [101]

$$a_1 z_1^2 + a_2 z_1 z_2 + a_3 z_2^2 + a_4 z_1 + a_5 z_2 + a_6 = 0, \tag{3.2}$$

with real coefficients $a_1, a_2, a_3, a_4, a_5, a_6$ such that $a_1^2 + a_2^2 + a_3^2 > 0$. Using a vector of parameters $\mathbf{a} = (a_1, a_2, a_3, a_4, a_5, a_6)^T$ and a vector of variables $u(\mathbf{z}) = (z_1^2, z_1 z_2, z_2^2, z_1, z_2, 1)^T$, the conic can be alternatively expressed as

$$\mathbf{a}^T u(\mathbf{z}) = 0. \tag{3.3}$$

The discriminant $\Delta = a_2^2 - 4a_1a_3$ categorizes any non-degenerate conic section into classes of parabolas, hyperbolas or ellipses. The constraint $\Delta < 0$ implies that the conic possesses ellipse-specifity leading to [100]

$$\mathbf{a}^T \boldsymbol{F} \mathbf{a} > 0, \tag{3.4}$$

where

$$\boldsymbol{F} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & 0 & 2 \\ 0 & -1 & 0 \\ 2 & 0 & 0 \end{pmatrix}, \tag{3.5}$$

where $\otimes$ denotes the Konecker product.

Ellipse fitting algorithm

Least-squares problems represent the process of fitting an over-determined parameterized model, with parameters $\mathbf{a}$, to a set of data pairs, $\mathbf{z}_j$ for $j = 1, \ldots, m$, where m denotes the number of available observations, by minimizing the squared residual error between the function, $f(\mathbf{t}_j|\mathbf{a})$, and the observations according to [98]

$$\chi(\mathbf{a}) = \sum_j w_j \cdot [\mathbf{z}_j - f(\mathbf{t}_j|\mathbf{a})]^2 \Rightarrow \min, \tag{3.6}$$

where the vector $\mathbf{t}_j$ comprises the condition for observation j and can be considered as the set of independent variables, and w_j denotes a weighting variable expressing the reliability of the assigned observation.

The process of fitting an ellipse to a set of observations demands an expressive cost function that describes the degree to which any distinct $\mathbf{a}$ fails to fulfill the observed system according to (3.3). The proposed cost function uses the approximated maximum likelihood (AML) distance originally introduced by *Sampson* in [92] and further developed by *Szpak et al.* in [100], [101], which combines the Gaussian model of errors with the principle of maximum likelihood, and takes the form

$$\chi_{\mathrm{AML}}(\mathbf{a}) = \sum_{j=1}^{m} \frac{\mathbf{a}^T u(\mathbf{z}_j) u(\mathbf{z}_j)^T \mathbf{a}}{\mathbf{a}^T \nabla u(\mathbf{z}_j) \boldsymbol{\Lambda}_{\mathbf{z}_j} \nabla u(\mathbf{z}_j)^T \mathbf{a}}, \tag{3.7}$$

where $\nabla u(\mathbf{z}_j)$ denotes the partial derivatives of $u(\mathbf{z}_j)$, and $\boldsymbol{\Lambda}_{\mathbf{z}_j}$ is a symmetric covariance matrix expressing the data uncertainty [26]. If the data points are assumed to be corrupted by Gaussian distributed noise, the function χ_{AML} rests upon the sum of orthogonal distances between the ellipse and the data points.

Nonlinear least-squares problems, such as the straightforward task of fitting an ellipse to noisy data, require iterative solutions to minimize χ_{AML}. In order to be able to guarantee that the minimization takes place in compliance with the ellipse constraint, the search space is limited to the subset $E = \{\mathbf{a}|\Delta < 0\}$ comprising all feasible non-degenerate conic sections. However, the presence of a minimizer of the function χ_{AML} in E can not be guaranteed due to possible ill-conditioning. As a solution, the optimization algorithm is regularized. By adaptively restricting the search domain in E, a stopping criterion is

employed, preventing the algorithm from generating estimates too close to the parabolic class [100].

The nonlinear least-squares minimization of $\chi_{\text{AML}}(\mathbf{a})$ is carried out using the Levenberg-Marquardt (LM) algorithm. The LM algorithm combines the well known Gauss-Newton method with the gradient-descent approach in order to enhance the convergence performance and overall robustness to initial ill-parametrization [98]. Based on a direct linear ellipse estimate as a seed, the LM method iteratively walks towards the global minimum of $\chi_{\text{AML}}(\mathbf{a})$ according to [98]

$$\chi_{\text{AML},k+1}(\mathbf{a} + \Delta\mathbf{a}) = \chi_{\text{AML},k}(\mathbf{a}) + \Delta\mathbf{a}_k, \tag{3.8}$$

with

$$\Delta\mathbf{a}_k = \left(\mathbf{J}_k^T \mathbf{W}_k \mathbf{J}_k + \lambda_k^{\text{LM}} \mathbf{I}\right)^{-1} \mathbf{J}_k^T \mathbf{W}_k \mathbf{r}_k, \tag{3.9}$$

where $\mathbf{J}_k$ denotes the Jacobian matrix at the kth iteration containing the partial derivatives of the ellipse function with respect to the observation condition $\mathbf{t}$. $\mathbf{W}_k$ is a diagonal matrix containing the weights of each observation, $\mathbf{r}_k$ contains the residuals between the function and the observations, and $\mathbf{I}$ is the identity matrix. The damping factor λ_k^{LM} influences the minimization behavior towards the gradient-descent method or towards the Gauss-Newton method, respectively. For large λ_k^{LM}, $\lambda_k^{\text{LM}} \cdot \mathbf{I}$ becomes dominant which results in gradient-decent steps, whereas small values of λ_k^{LM} attaches weight towards the Gauss-Newton method [98].

The proposed ellipse fitting approach incorporates the RANSAC algorithm to realize robust parameter estimation. The RANSAC algorithm was initially introduced by *Fischler and Bolles* in 1981 for solving the location determination problem by estimating a model's parameters based on a set of data comprising a substantial amount of outliers [1]. In the RANSAC's context, an outlier denotes a contaminant observation, i.e., a data point that significantly differs from the majority of the available data set and hence appears to be generated by a different mechanism [98]. The RANSAC algorithm aims to implicitly detect those outliers by essentially following two iteratively repeated steps [98]:

- **Hypothesize**. At first, minimal sample sets (MSSs) are randomly chosen from all available data points to compute the model's parameters using only the elements of the MSS. This is done with the expectation that this MSS is free of outliers. The number of elements is chosen to be the smallest sufficient to calculate the model parameters which maximizes the probability that a certain MSS does not contain outliers (as opposed to the least-squares approach, where all available data are employed to estimate the parameters).

- **Test**. Next, RANSAC determines the subset of elements of the entire data set whose distance to the model is smaller than a certain threshold and hence supports the model. This subset of elements is called consensus set (CS).

The CS that possesses the maximum cardinality is subsequently used to estimate the final ellipse. Figure 3.3(a) pictorially illustrates the RANSAC parameter estimation based on two different CS iterations. For the process of fitting ellipses to the μD distributions of the
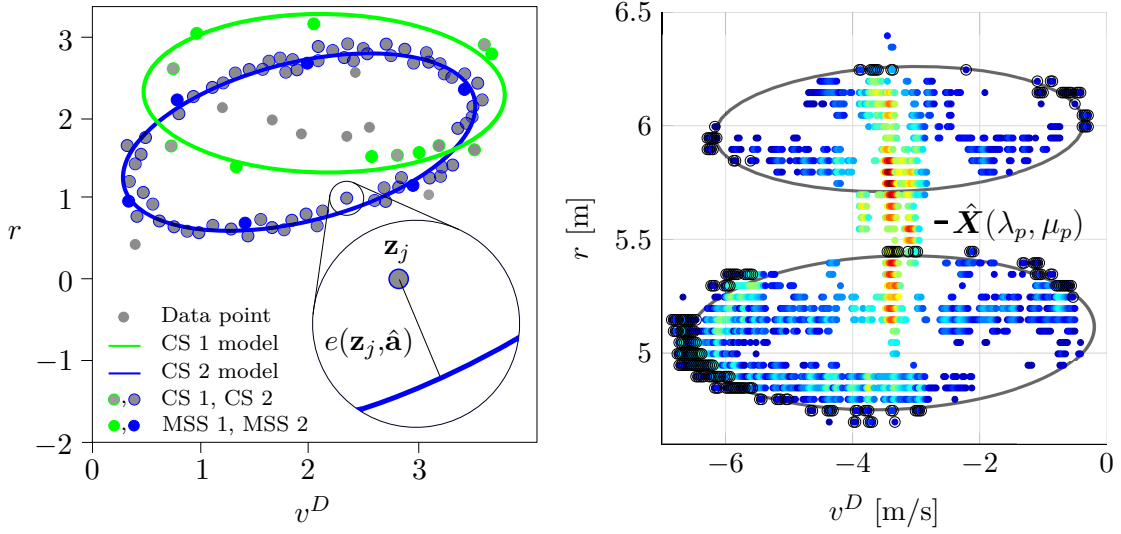
(a) RANSAC parameter estimation based on two different CS iterations.

(b) Results of the ellipse fitting algorithm using real radar data. Black marks denote the CS. © 2019 IEEE [51]

Figure 3.3: Exemplary illustration of the RANSAC ellipse fitting principle (a), and the results of the ellipse fitting algorithm applied to the front and rear wheels' μD distributions using a real radar data snapshot of a moving cyclist (b). The color-coding corresponds to the backscattering intensities. The generated ellipses reveal the spectral components, $\hat{\mathbf{X}}(\lambda_p, \mu_p)$, emerging from the pedaling.

front and rear wheel, the range-Doppler spectrum is vertically segmented in two parts based on the largest backscattering intensities along the translational Doppler components. An appropriate search range for the ellipse fitting procedure for both segments is determined considering the wheels' physical principles described in Subsection 3.1.1. The semi-major axis of the initial ellipse is $|\boldsymbol{v}_{\text{ego}}|$, whereas $r_{\text{w}}/\Delta r$ denotes the length of the semi-minor axis with r_{w} referring to the radius of a standard 28 inch wheel. The center point of the ellipsoidal search area is then the mean value in range direction, and the Doppler value of the averaged largest backscattering intensities on which the spectrum was initially segmented.

Let $\mathbf{Z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_m\}$ be the set of m remaining radar detections, and $\mathbf{a}_0(\{\mathbf{z}_1, \ldots, \mathbf{z}_h\})$ denotes the MSS with cardinality h on which the initial ellipse estimation is based, the LM algorithm minimizes the function $\chi_{\text{AML}}(\mathbf{a}_0)$ according to (3.8) yielding the optimal geometric fit $\hat{\chi}_{\text{AML}}(\hat{\mathbf{a}})$. The residual error between the detection $\mathbf{z}_j \in \mathbf{Z}$ and $\hat{\chi}_{\text{AML}}(\hat{\mathbf{a}})$ is given by [96]

$$e_{\hat{\chi}_{\text{AML}}}(\mathbf{z}_j, \hat{\mathbf{a}}) = \min_{\mathbf{z}_j' \in \hat{\chi}_{\text{AML}}(\hat{\mathbf{a}})} \text{dist}(\mathbf{z}_j, \mathbf{z}_j'), \tag{3.10}$$

where $\text{dist}(\mathbf{z}_j, \mathbf{z}_j')$ is the Euclidean distance between $\mathbf{z}_j$ and its orthogonal projection of $\mathbf{z}_j'$ onto the function $\hat{\chi}_{\text{AML}}(\hat{\mathbf{a}})$. The CS is consequently determined using the constraint

$$S(\hat{\mathbf{a}}) = \{\mathbf{z}_j \in \mathbf{Z} : e_{\hat{\chi}_{\text{AML}}}(\mathbf{z}_j, \hat{\mathbf{a}}) \le \delta\}, \tag{3.11}$$

where δ is a constant threshold. The RANSAC process is repeated for a suitable number of different MSS whereby in each iteration the CS is updated if

$$|S(\hat{\mathbf{a}}_k)| > \max_{l \in 1, \dots, k-1} |S(\hat{\mathbf{a}}_l)|. \tag{3.12}$$

Finally, the model is fitted to the largest CS which completes the RANSAC approach for the current radar frame. Figure 3.3(b) depicts the ellipse fitting results for both wheels in the range-Doppler spectrum revealing the spectral area that most likely comprises the estimated pedaling motion $\hat{X}(\lambda_p, \mu_p)$. Consequently, every detection laying in between the ellipses is considered to be part of $\hat{X}(\lambda_p, \mu_p)$.

3.2.2 Experimental results

Figure 3.4(a) illustrates the extracted spectral regions, $\hat{\boldsymbol{X}}(\max(\lambda_p), \mu_p)$, assigned to a cyclist's pedaling motion moving parallelly to the static radar sensor with a lateral displacement of $2\,\text{m}$. The entire signal processing provided in Subsection 3.2.1 is applied for successive radar frames. The extracted signal is primarily constituted of the parts facing the sensor, i.e., the cyclist's right leg, due to the given orientation. Thereby, the signal covers both the translational velocity at around -3.5 m/s, and the superimposed pedaling motion. Consequently, the signal reveals two essential sinusoidal contributions with diverse amplitudes and phases. The associated pedaling rotation possesses high amplitudes in velocity and backscattering due to the more reflective material. At the same time, the second contribution originates most likely from the cyclist's knee showing generally weaker backscattering and a phase shift due to biomechanical circumstances [44].

As can be seen, the extracted signal includes single outliers that do not belong to the pedaling motion. In some situations the occurrence of the wheels' μD is not sufficiently pronounced due to occlusion. As a result, the ellipse fit does not reproduce an optimal approximation of the respective velocity distribution, but rather erroneously assigns other scatterers as pedaling contributions.

Figure 3.4(b) shows the analogously obtained spectral excerpt of a non-pedaling cyclist for the identical motion trajectory. The spectrum reveals no perceptible pedaling components. Instead, it is indicated solely by the bicycle's translational velocity and occasional scatterers. To emphasize the spectral discrepancy between existing pedaling and non-pedaling motions, the standard deviations specified by the respective normalized weighted velocity variance

$$\sigma^2 = \Big(\sum_{i=1}^{n} \gamma_i \Big)^{-1} \sum_{i=1}^{n} \gamma_i \cdot (\mu_{p_i} - \mu_{\text{trans}})^2, \tag{3.13}$$

are evaluated and depicted in Figure 3.4(c) and Figure 3.4(d) for the corresponding measurement, respectively. The variance is computed in each frame based on a weighted

squared difference between the μD value μ_{p_i} and the μD value of the detection possessing the largest backscattering amplitude μ_{trans} for each $i = 1, \ldots, m$, where m is the number of detections in $\hat{\boldsymbol{X}}(\lambda_p, \mu_p)$. The variances are weighted according to the detections' backscattering intensities. Here, γ refers to the power amplitudes of $\hat{\boldsymbol{X}}(\lambda_p, \mu_p)$ and incorporates them through multiplication. The methodology was also carried out for measurements with various distances between the sensor and the cyclist. Apparent variance distinctions between pedaling and non-pedaling can be conducted up to a range of 4 meters. The computation of the variances is of little computational effort and is updated for each
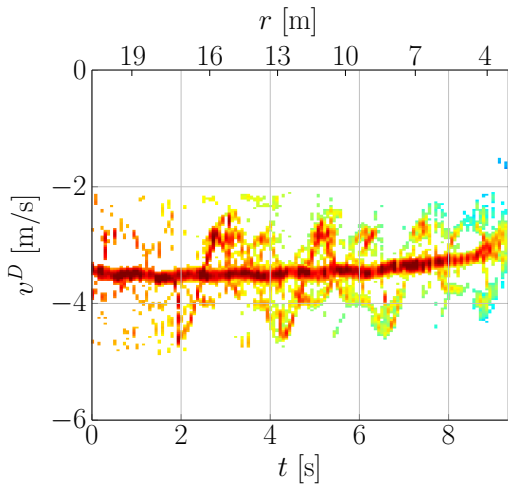
(a) Extracted $\hat{\boldsymbol{X}}(\max(\lambda_p), \mu_p)$ for a pedaling cyclist.
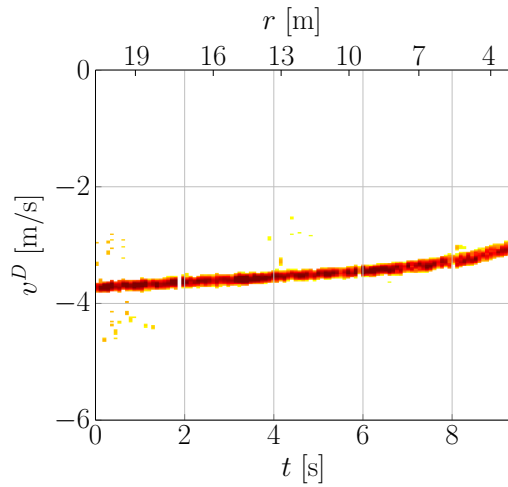
(b) Extracted $\hat{\boldsymbol{X}}(\max(\lambda_p), \mu_p)$ for a non-pedaling cyclist.
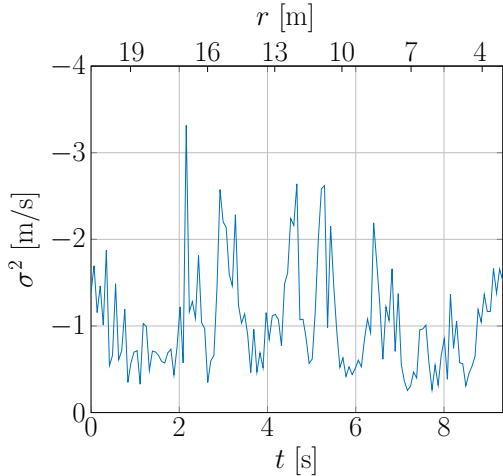
(c) Velocity variances of $\hat{\boldsymbol{X}}(\max(\lambda_p), \mu_p)$ for a pedaling cyclist.
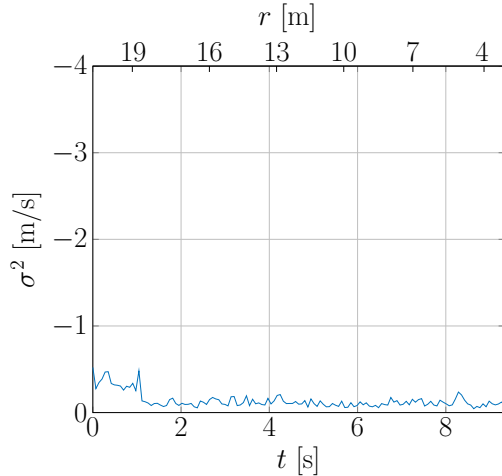
(d) Velocity variances of $\hat{\boldsymbol{X}}(\max(\lambda_p), \mu_p)$ for a non-pedaling cyclist.

Figure 3.4: Extracted μD pedaling spectra over time and range with corresponding velocity variances for a moving cyclist. The color-coding in (a) and (b) corresponds to the backscattering intensities. © 2019 IEEE [51]

cycle. Therefore, its temporal transition can be used explicitly as a decision parameter for safety functions affecting cyclists' behavioral indication. The results demonstrate the ellipse fitting procedure's functionality and the clear distinctness between a pedaling and non-pedaling cyclist.

3.3 Normalization of the orientation dependence

Radar-based classification algorithms for VRU, such as deep learning approaches, make effective use of the distinctive μD distributions during the training step. Therefore, signal corruption due to orientation dependence may affect accurate classifications and data associations. Especially in the case of intersecting VRU in urban areas, the orientation dependence drastically problematizes reliable detection and feature extraction procedures due to ambiguities in range and Doppler, as exemplarily shown for a bicycle in Figure 3.2(a). As a result, the object's overall extent reduces during the process of approaching, as illustrated in Figure 3.2(b).

In the following, we present a procedure that normalizes a cyclist's geometric and kinematic extension using a back-projection technique that allows a static representation of the range-Doppler parameters, which increases correct classifications and the extractability of object features. The proposed method was published in [52].

3.3.1 Geometric correction

The geometric correction strives to normalize the orientation-dependent range ambiguities. The idea is to create an axis that runs along with the bicycle's longitudinal orientation onto which the measured detections are projected. The projection is performed on the
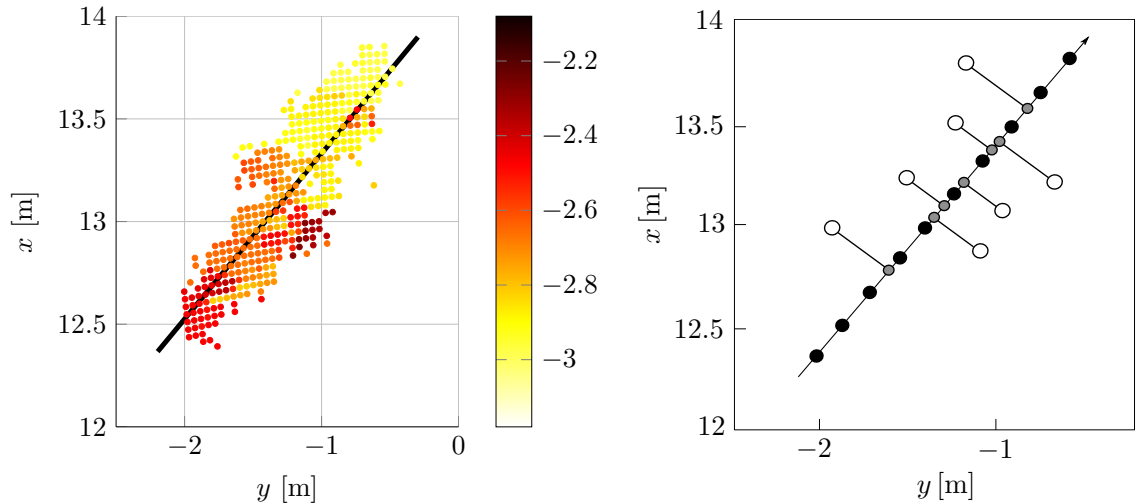
(a) Projection axis $\chi(\alpha, \beta)$ resulting from a regression fit based on data points $\tilde{D}$ with color-coded Doppler velocities in m/s.

(b) Illustration of data projection onto artificial range vector. ● : range gates. ○ : detections $\in D$. ● : orthogonal projections $\mathbf{z}'$.

Figure 3.5: Creation of projection axis $\chi(\alpha, \beta)$ based on detections $\tilde{D}$. © 2019 IEEE [52]

angular resolved Cartesian detections that reveal the overall orientation of the object. By projecting every Cartesian detection onto the bicycle's orientation axis, a superficial range vector is created that depicts the normalized radial distances of each detection.

Let $D = \{\mathbf{z}_1, \ldots, \mathbf{z}_m\}$ be the detection list of the current frame that forms the input dataset of the following procedure. In order to achieve satisfying results, the projection axis needs to comply precisely with the bicycle's actual orientation. Hence, solely detections possessing large backscattering intensities are considered for the axis' creation. It is legitimate to presume that these detections emanate from the actual bicycle and are no false positives. Let $\tilde{D} \in D$ be the $\tilde{m}$ remaining detections after applying a hard threshold. The projection axis is created using least-squares minimization as already defined in (3.6) and given again for convenience

$$\chi(\alpha, \beta) = \sum_{j=1}^{\tilde{m}} w_j \cdot \left[\mathbf{z}_{x_j} - f(\mathbf{z}_{y_j}|\alpha, \beta) \right]^2 \Rightarrow \min_{\alpha, \beta}, \tag{3.14}$$

with x-intercept α and axis slope β. Figure 3.5(a) shows the bicycle's angular resolved Cartesian detections, $\tilde{D}$, of a single radar snapshot and the fitted line, $\chi(\alpha, \beta)$, that approximates the bicycle's longitudinal axis. The depicted color-coding of the detections represents their measured Doppler amplitudes v^D in m/s highlighting the occurring kinematic spread due to the orientation.

Before performing the projection, the axis $\chi(\alpha, \beta)$ is divided into equidistant length units and thus can be interpreted as an artificially generated range vector. Those range gates' length can be arbitrarily chosen, allowing a substantially enhanced representation of the projected spectrum. Note that all detections in D are projected on the generated axis. Consequently, $Z = \{\mathbf{z}'_1, \ldots, \mathbf{z}'_m\}$ comprises all projected detections, where $\mathbf{z}'$ denotes the orthogonal projections of $\mathbf{z}$ onto the axis $\chi(\alpha, \beta)$. Next, each distance of the projected detections to the zero-point of $\chi(\alpha, \beta)$ is computed and allocated to the connected range gate, as pictorially shown in Figure 3.5(b).

3.3.2 Kinematic correction

The kinematic correction assumes a linearly moving bicycle that is free of any yawing dynamics. Given that, each of the bicycle's reflection points possesses the same translational velocity vector. The moving bicycle's actual orientation, which corresponds to the translational velocity vector's orientation, can be approximated by the projection axis's orientation $\hat{\Phi}$. Consequently, each point's actual velocity can be estimated by [64]

$$\hat{v}_j = \frac{v_j^D}{\cos(\hat{\Phi} - \phi_j)}, \tag{3.15}$$

where v_j^D for $j = 1, \ldots, m$ is the measured Doppler velocity and ϕ_j denotes the measured azimuth angle of the jth detection, respectively.

3.3.3 Experimental results

Figure 3.6 illustrates the normalization procedure results by comparing the measured range-Doppler spectrum of a moving bicycle with the normalized correspondent. Here, the

normalized Doppler velocity is plotted against the projection axis $\chi(\alpha, \beta)$ in m. Note that the zero-point of $\chi(\alpha, \beta)$ is relative and thus does not possess any significance. Though, regarding the purpose of comparability, both excerpts cover a distance of 4 m. The enhanced geometric extent in Figure 3.6(b) reveals the normalization of the original spectrum's radial range components due to the projection procedure. While the original object exhibits a geometric extent of around 1.35 m, the projected extent spans a length of around 2.05 m which corresponds to an increase of approximately 34 %. Consequently, the current extent approximately equals the actual bicycle's length.

The kinematic correction yields the normalized Doppler amplitudes and hence the actual velocity distributions of the bicycle. It becomes apparent that both the wheels' μD amplitudes, as well as the translational velocity amplitudes, are now significantly increased. Furthermore, the bicycle's translational velocity is now consistent along the entire spatial extent having an amplitude of approximately 3.7 m/s.

Besides, the presented approach enables enhanced detection abilities of μD features such as pedaling. By allocating the Cartesian detections to the bicycle's longitudinal axis, motion components, such as the pedaling, can be depicted with improved accuracy. Contrary to the conventionally generated spectrum, the pedaling emanating reflections are distinctly perceptible in the normalized image. This may be crucial for behavior intention recognition of cyclists that may be manifesting on pedaling changes. Moreover, the proposed method promotes standardization of range-Doppler spectra, which are of great significance in the field of supervised and unsupervised machine learning approaches. By normalizing orientation-dependent radar data, it can be assumed that classification algorithms' identification rates may rise considerably. Especially in scenarios with broad lateral orientations, the normalization approach provides an essential advancement for correct detection and classification. Besides the benefits of the proposed technique, it is striking to note that the
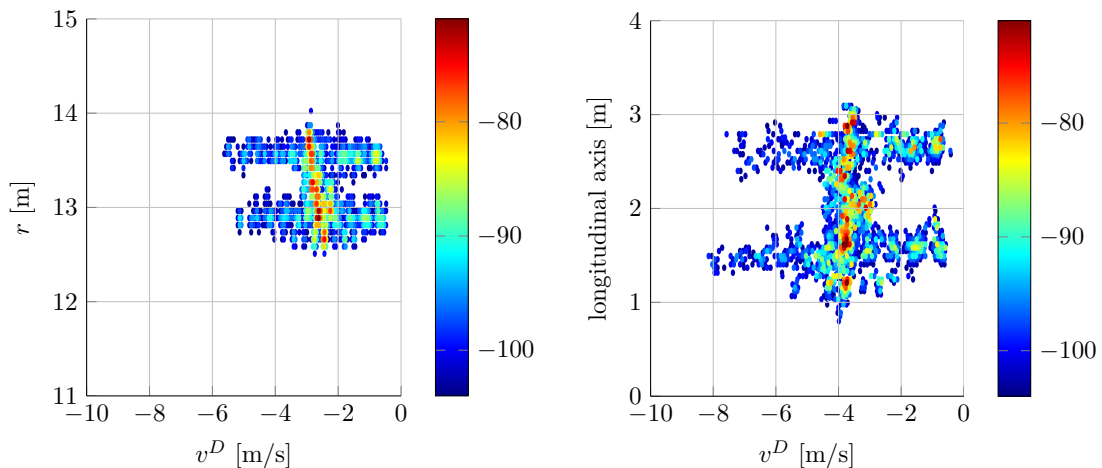
(a) Original range-Doppler spectrum.   **(b)** Normalized range-Doppler spectrum.

Figure 3.6: Original and normalized range-Doppler spectra of a cyclist. The logarithmized backscatter intensities are in dBV. © 2019 IEEE [52]

performance relies practically on the projection axis's orientation. Since the projection axis serves as an approximation of the bicycle's longitudinal axis, deviations from the actual bicycle's longitudinal axis result in distortions due to the orthogonal data projection.

3.4 Conclusion

The presented work introduced high-resolution radar signatures of cyclists and pointed out challenges of extended targets, such as the orientation dependence between the cyclist and the radar sensor. After describing the fundamental physical principles that lead to extended μD distributions, signal processing approaches are presented in order to enable essential feature extraction procedures, such as the extraction of pedaling motions, which are crucial for designated automotive safety functions. In particular, an adaptive ellipse fitting algorithm was applied to the range-Doppler spectrum to separate the wheels' μD components from those of the pedaling motion. The approach allows the extraction of the pedaling movements for a series of consecutive radar frames. The extracted spectral segment consists of sinusoidal signals picturing the rotating pedals and the lower leg motions and was compared with non-pedaling measurement. The related velocity variances have been computed and display apparent differences in pedaling or non-pedaling scenarios, respectively. The presented method's results demonstrate the potential importance of cyclists' behavioral indications for short-range automotive safety applications of the future. Furthermore, a procedure has been proposed for normalizing the orientation-dependent ambiguities in range and Doppler of a cyclist. A projection technique using angular resolved data combined with a correction of the radial velocity components was applied to radar measurements in the 77-GHz band using chirp sequence modulation. The measured and normalized range-Doppler spectra have been compared and analyzed. The results emphasize the significance of the presented approach regarding reliable identification of cyclists in the short-range and enhanced feature extraction of radar-based behavioral indications such as pedaling for automotive use.

3.5 Discussion and outlook

The proposed signal processing techniques aim to tackle the orientation's influence on the received radar signals in order to extract crucial features, such as the pedaling, for sophisticated safety functions. The idea to approximate the wheels' characteristic μD distributions by an ellipsoidal nonlinear least-squares procedure allows the adaptive segmentation of the bicycle's structure in range and Doppler and hence uncovers the pedaling contribution. The results prove the presented implementation's functionality up to distances of 20 m and demonstrate its general suitability as a near-field application for automotive systems.

Elliptical approximation as a technique for the adaptive disclosure of pedaling motions can also be utilized to detect steering movements. The change in orientation of the front wheel as a result of a steering movement directly influences the μD's geometric shape and thus of the approximated ellipse. An abrupt steering movement results in a shortening of the ellipse's semi-axes due to the reduction of the reflecting wheel surface and the radial velocity components, respectively. These features could be used as behavioral indications in addition to the pedaling motion, for example, to identify conspicuous driving patterns

such as zigzag movements and to declare them according to their safety relevance.

The presented method for the normalization of the orientation dependence is based on the detections' projection onto the bicycle's approximated longitudinal axis. For this purpose, the largest scatterers are employed on the assumption that they result from the bicycle's frame and thus represent the actual longitudinal axis. Consequently, the performance of the normalization algorithm depends on the detectability of the bicycle frame. Influencing factors such as large relative angles between the bicycle and the radar, occlusion, or the bicycle frame's reflectivity can affect the overall performance.

CHAPTER 4

Micro-Doppler applications of a pedestrian

4.1 Pedestrian detection

From a radar point of view, a pedestrian can be described by the measured information containing RCS, geometric extent, and velocity behavior. With high-resolution radars, which are steadily becoming more standard in short-range applications, a pedestrian is considered an extended object. Hence, it gives rise to a varying number of spatially distributed reflection points in several dimensions fulfilling the double extent definition provided in Subsection 2.3.1. Figure 4.1 illustrates a high-resolution radar snapshot of a walking pedestrian in Cartesian transformed x-y-coordinates with respect to the sensor coordinate system and range-Doppler perspective. As can be seen, the detections of the object form a reflection surface in both representations. The color-coded amplitudes allow basic conclusions about the reflection behavior of the body. However, the backscattering is typically subject to a substantial phase sensitivity, i.e., phase shifts within a range resolution cell leading to interference and severe fluctuations of the RCS amplitudes [112]. Although the interference effects tend to decrease with increasing range resolution due to the Fourier transform's improved frequency separation capability, they still occur. As already stated in Subsection 2.1.1, range changes of a few millimeters can lead to drastic interferences. Consequently, the distinction between pedestrians and arbitrary targets with comparable size and backscattering solely based on RCS information is impossible. Instead, velocity distributions, i.e., μD information, generated by limb motions are statistically uncorrelated to the RCS and are predominantly employed for pedestrian detection and classification objectives [15], [53], [77], [87]. Nevertheless, Doppler and μD amplitudes rely on the objects' orientation due to the radial measuring principle and typically assume small values or even converge towards zero for entirely lateral directions. In other words, entirely lateral moving pedestrians show similar velocity characteristics as static objects. With occlusion effects via stationary objects and antenna effects like parasitic side-lobing, pedestrian detection techniques can be vigorously affected [15]. These aspects are also assessed concerning the development of naturalistic pedestrian dummies and simulation frameworks and thus establishing test criteria for integral safety systems in the domain of pedestrian detection [28], [30]. Nevertheless, the measured radial velocities show unique perceptual characteristics that make robust pedestrian detection feasible and enable further behavior indication procedures.

4.1.1 Human kinematics

The concept of structure motions not rigidly connected to an object, which was already presented in Subsection 2.1.3, can be expanded and appointed to a human being's loco-
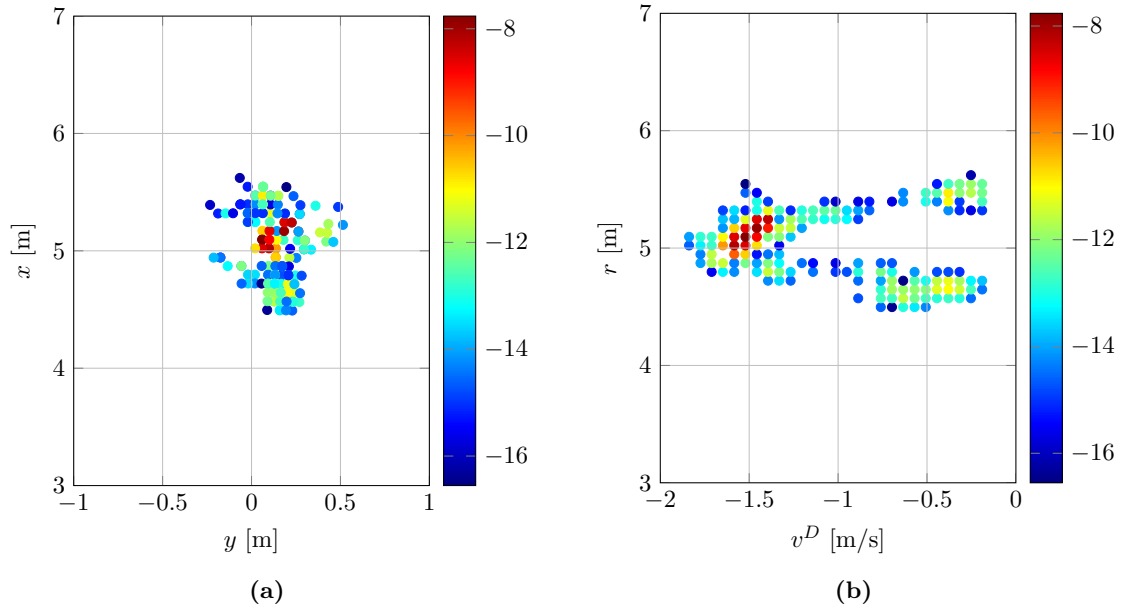
Figure 4.1: High-resolution radar snapshot of a walking pedestrian displayed in Cartesian coordinates and range-Doppler, respectively. The logarithmized backscatter intensities are in dBsm.

motion. For this purpose, two perspectives are introduced for the pedestrians' level of motion ability. The *macroscopic* motion defines global positioning while walking, i.e., the approximately linear translational movement combined with many degrees of freedom, such as the capacity to change direction and walking speed quickly. *Microscopic* motion incorporates the pedestrian's highly versatile motion apparatus revealing various limb motion courses and frequency variations, laying the foundation for utilizing the μD effect. A human body's motion expresses articulated locomotion where the extremities expose a periodically repeating motion routine in the gait cycle. A gait cycle incorporates a stance and a swing phase. The stance phase takes 60 % of the cycle duration, and is characterized by the fact that the feet are permanently in contact with the ground showing a heel strike and a toe-off, respectively. When the body shifts its weight to the supporting leg, the swing phase, which occupies the remaining 40 % of the cycle, is initiated by lifting one foot. The forward-moving process is divided into three parts: acceleration of the foot, mid-swing when the foot passes the supporting leg, and deceleration of the foot preparing the next cycle [59]. The permanent change of a pedestrian's center of gravity while walking induces microscopic motion by the extremities' inclination. Those patterns are predestined to be used as indicators for pedestrian detection and, based on this, for intentional recognition. The human motion recording is carried out with the Xsens MTw Awinda wireless full-body motion capture (MoCap) system. The system comprises 17 inertial measurement units attached to designated locations on the test person's body, as shown in Figure 4.2, to record the individual body segments' motion. Data transmission frequency between the accelerometers and the receiving station is 60 Hz yielding to incoming data packets every

16.7 ms [50]. Figure 4.3 shows representative MoCap data for a single gait cycle of a continuous walking test person equipped with inertial measurement units. The motion data of the feet, hands, and head are shown in relation to an initially defined body-related longitudinal axis in x direction.

4.1.2 Micro-Doppler signatures of a pedestrian

The utilization of the μD effect enables entirely new approaches in radar-based detection and classification of VRUs like pedestrians. Especially the limb motion can provide fundamental attributes that can influence the design of automotive safety systems. Consequently, detailed investigations of human μD articulation due to limb movement plays an essential role in current ADAS developments [4], [54], [76]. By utilizing high resolutions in range and Doppler, a moving pedestrian appears as an multi-reflective object whose μD contributions are superimposed in the frequency domain. The spectral composition of those characteristic pedestrian identifying μD patterns is of particular interest concerning safety applications. Several investigations use radar simulations instead of measured data to create μD signatures for automotive use [2], [93]. Generally, simulations of radar returns used to analyze specific motion features come along with disadvantages and need to be considered deliberately. Although μD simulations are reasonably suitable to display specific characteristics of real radar data, they are still simplified. Typically, they are limited to a few individual scatter points and thus do not represent the complexity of multi-reflective surfaces, such as a human leg or the torso. Also, radar simulations usually do not sufficiently cover environmental influences such as noise or multipath reflections e.g., on the ground or between limbs. Other researches use real radar data to compare measured μD signals with randomly recorded MoCap data [17], [48]. Due to the independence of MoCap data and radar data, the assignment of individual body parts by μD signals is based on unreferenced interpretations and hence only on similarities. However, the direct comparison of measured
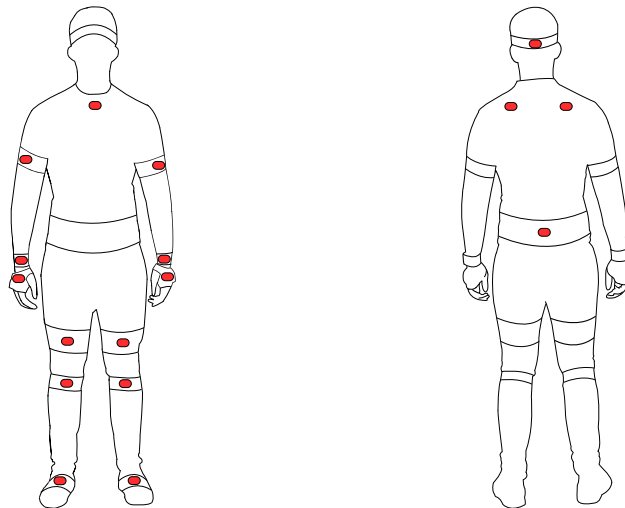

Figure 4.2: Test person wearing 17 acceleration sensors (red markings) attached to defined locations on the body in front and back view, respectively. © 2018 IEEE [50]
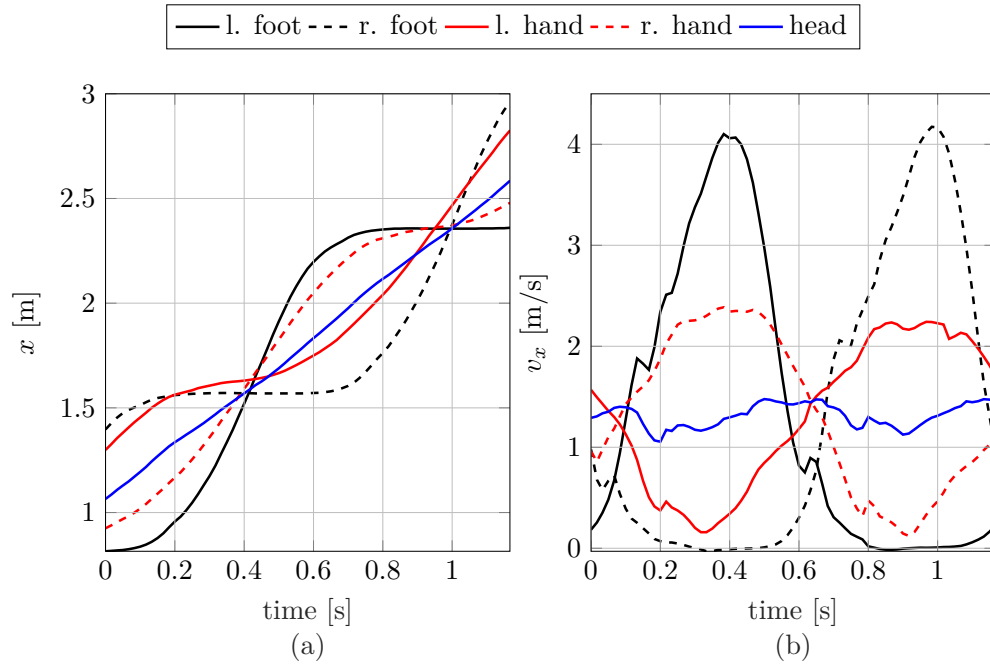
Figure 4.3: MoCap data for one gait cycle of a continuous walking test person equipped with inertial measurement units at defined limb locations.

radar data and synchronized MoCap reference data is crucial to understanding human motions' appearances and identifying individual body parts in the radar image. Various motion patterns are performed and analyzed. The first motion pattern is continuous walking. The second covers velocity transitions, e.g., the transition between walking and running. Finally, the third motion pattern includes crossing maneuvers such as abrupt turning movements, e.g., 90° transitions.

4.1.3 Experimental results

Radially moving pedestrian

Based on [50], in the following, the OS-CFAR-filtered and clustered signatures are visualized as time-frequency distributions and analyzed under consideration of the reference MoCap data depicted on top of the spectra. For the sake of clarity, only the most prominent body components of the reference data are pictured. Those include the feet, lower legs, hands, and pelvis. The lower legs comply with the knee locations since the corresponding sensors have been installed directly below the knee joints, as shown in Figure 4.2. Figure 4.4(a) shows the measured radar data of a pedestrian moving towards the static radar as a time-Doppler function. Contrary to the provided signal processing from Subsection 2.3, where the window length of the Doppler-DFT contains all chirps of a CPI, a short-time-Fourier-transform (STFT) is used to generate the time-Doppler spectra in Figure 4.4 and Figure 4.5, respectively. The STFT represents the Fourier transform of a signal, which is first divided into equal-length segments. The Fourier transform is then computed

individually on each segment. Hence, the STFT only determines the frequencies in a sub-segment of the whole CPI. The temporal sequence of the spectral Doppler frequencies obtained in this way represents a satisfactory tradeoff between Doppler and time resolution and highlights the human walking μD effect [50]. Note that the radar data are compensated by the distance-dependent signal attenuation to achieve a consistent power amplitude throughout the traveled distance. Besides, the synchronized reference curves corresponding to the measurement are depicted on top of the spectrum, as shown in Figure 4.4(b). This form of imaging allows a detailed understanding of the composition of the spectrum. The periodically changing Doppler amplitudes induced by the limb movements can be clearly recognized. The reference data elucidate that the leg swinging causes the largest Doppler amplitudes with values up to $-6\,\mathrm{m/s}$. The discrepancy between the feet's reference amplitude and the Doppler amplitude can be explained by the fact that the motion capture system's accelerometers of the feet are placed on top of the feet. Assuming that the leg's motion in the swing phase can be approximated to that of a pendulum, the maximum velocity occurs at the sole. Hence, in addition to occurring signal processing effects for extended target reflections, the difference in length between the accelerometer's position and the actual leg length is the main reason for the deviation. The magnitudes of the lower legs' velocity amplitudes and the hands' amplitudes are almost identical with maximum and minimum values of about $-3\,\mathrm{m/s}$, and $-0.5\,\mathrm{m/s}$, respectively. Besides, their μD components are gently perceptible in the spectrum due to their relatively high backscattering. As the largest reflecting body surface, the torso exhibits the highest RCS up to $-5\,\mathrm{dBsm}$. This can be confirmed by the pelvis' velocity behavior, which coincides with that of the torso. The torso backscattering decreases with decreasing distance to the radar due to the radar's narrow field of view in elevation.

The results show that the knee movement represents the initial movement of a beginning gait cycle. Figure 4.6 illustrates a range-Doppler sequence of a starting quarter gait cycle from which the lower leg's initial acceleration emerges. The corresponding MoCap data are directly superimposed in range and velocity over consecutive frames. Beginning with the stance phase where both feet are in contact with the ground, the left (rear) leg's incipient acceleration, more precisely the left knee joint, can be captured. Figure 4.4(a) reveals the lower leg's beginning acceleration in the form of a gentle protrusion of the Doppler envelope. By moving the angled left knee forward, the left foot also experiences an acceleration. The right hand, however, is the last to be accelerated. As the acceleration progresses, the angle between the upper and lower leg is steadily reduced, thus increasing the foot's velocity until it reaches its maximum, as shown in Figure 4.6(f). A gait cycle analysis indicates the tremendous potential of the range information regarding the proclaimed objective of radar-based pedestrian feature extraction.

The scenario presented in Figure 4.4(c) visualizes the transition from walking to running. At first, the test person walks radially towards the sensor. After approximately $4\,\mathrm{s}$ the subject starts to run out of the walking motion abruptly and decelerates vigorously at the end. From the reference data, it becomes apparent that the running initiating motion is the increase of momentum of one hand, which is detectable in the μD signal. Moreover, the timing of this action needs to be emphasized. It is striking to note that the hand's acceleration happens between the gait cycles and thus contrary to the previous rhythm.

This might indicate a pedestrian's abrupt transition movement, which can be used for behavior prediction. For the duration of running, in addition to a pronounced increase in all velocity amplitudes up to 14 m/s, a significant velocity increase of the hands compared to walking is noticeable. The hands show a quadrupled velocity value of approximately 10 m/s and reach or even exceed the feet' velocity values. Moreover, the stance phases' duration is shortened by about three quarters compared to the walking phase.

Crossing pedestrian

Besides radial movements, lateral movements are of particular relevance regarding radar-based pedestrian detection in the automotive domain. Due to the radar's measurement principle, only radial velocities can be evaluated, making lateral movement patterns challenging to detect. On this account, the results of a pedestrian who suddenly leaves the sidewalk to cross the road are presented in the following. At first, the test person walks parallel to the radar sensor's longitudinal axis and then transitions into a laterally crossing movement. Figure 4.5 depicts the μD spectrum and the affiliated MoCap data of the mentioned trajectory. Note that the reference data are adapted to reflect the limbs' radial velocities and are therefore suitable for comparison. The test person approximately strides at the same velocity for the longitudinal and sidewards path, respectively. In the beginning, the pedestrian gradually enters the radar's azimuthal field of view, which explains the low backscattering within the first moments. With the pedestrian's increasing distance from the radar, the angle between the actual velocity vector and the radial projection v^D becomes smaller. As a consequence, the v^D velocities rise to 5 m/s. At around 5.5 s, the test person initiates the turn, which can be noticed by a decrease in velocity. The swing of the left hand and, consequently, the upper body's rotation induces the turning movement. At this precise moment in time, at around 5.7 s, the algebraic sign of the left hand's Doppler velocity switches, which is well discernible in the spectrum. This demeanor can be employed to design predictive algorithms in terms of pedestrian safety since the velocity shift can be considered an impending change in the direction of movement (e.g. road crossing). Directly with changing the direction of motion at around 6.0 s, the velocities v^D decline drastically and adopt negative values, since now the radial velocity vector points in the opposite direction. This abrupt fall in the pedestrian's radial velocity may indicate a sudden direction modification towards lateral orientation. With the proceeding trend, v^D becomes smaller until it reaches zero at passing the sensor's longitudinal axis at around 7.0 s. Afterward, v^D increases again with a shifted algebraic sign.

However, all results and distinctive attributes acquired are based exclusively on one subject's movement behavior and cannot be generalized. For this cause, further advanced investigations are needed to develop more adequate approaches concerning safety algorithms.
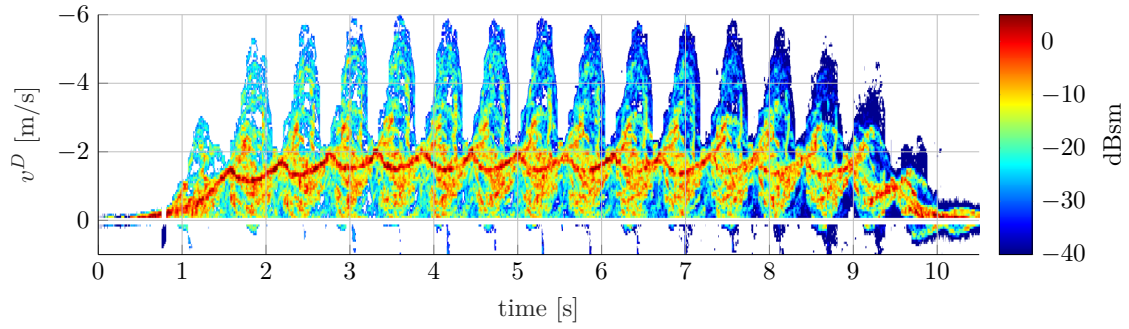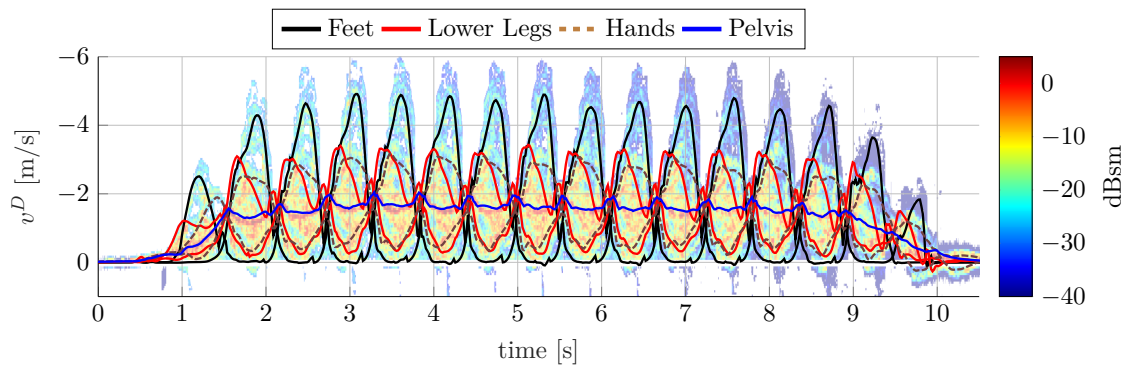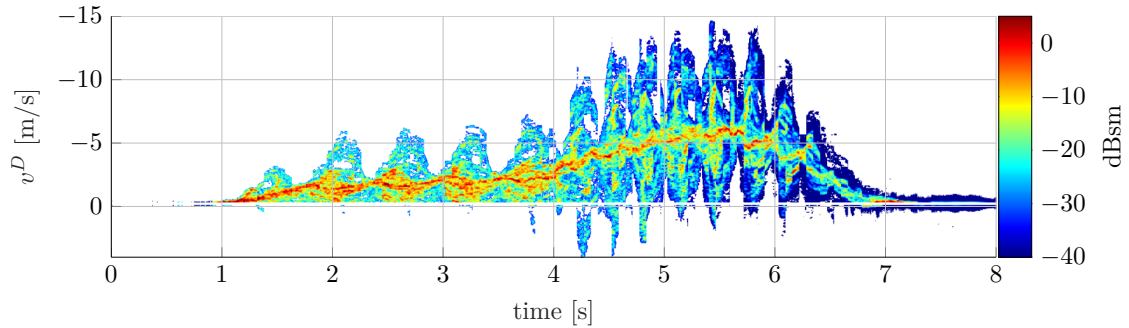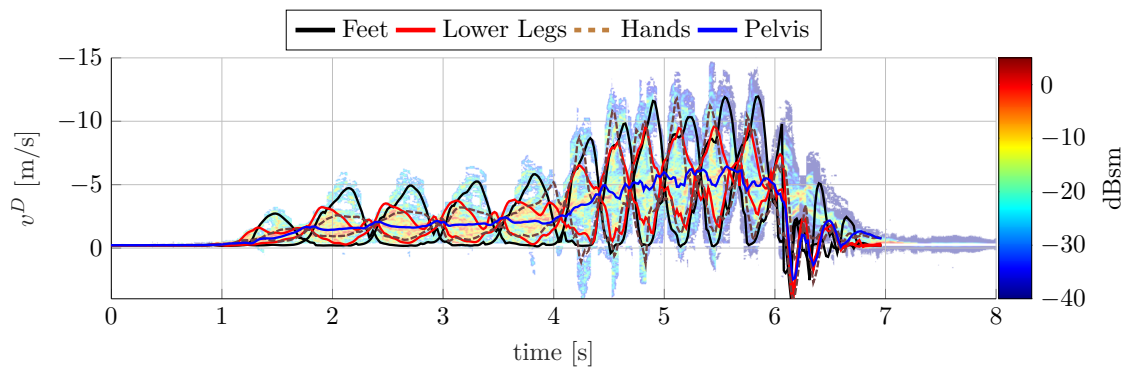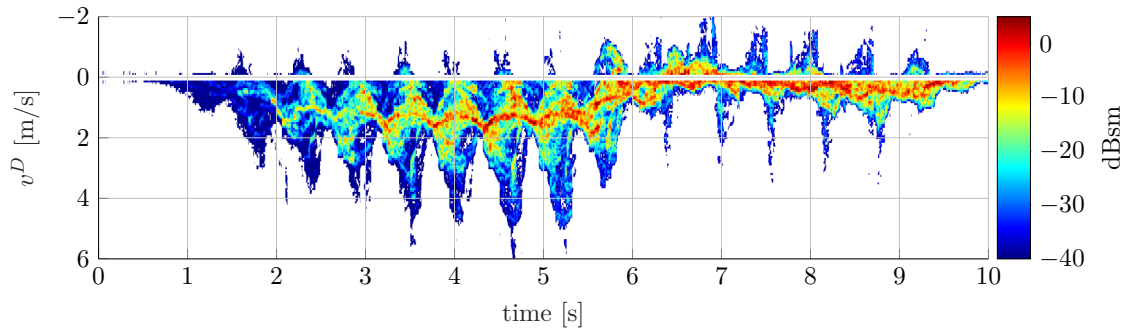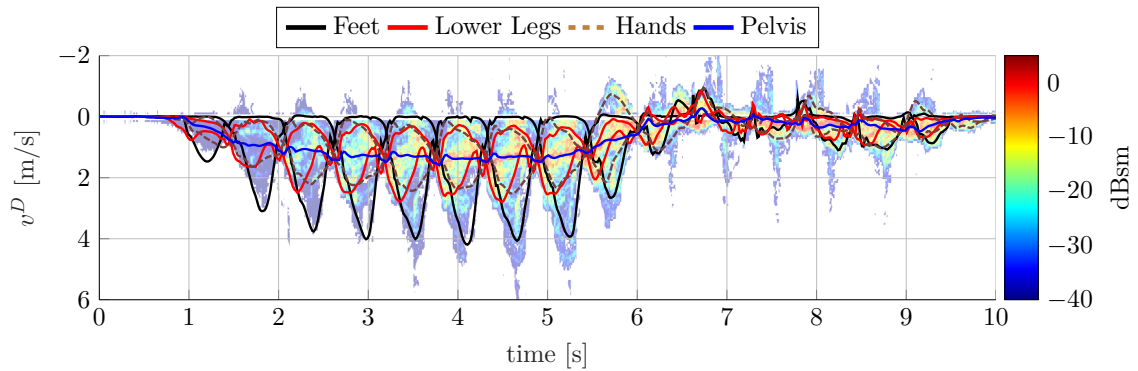
(a) Spectrogram of measured μD radar data.

(b) Spectrogram of measured μD radar data with MoCap data of the feet, hands, lower legs, and pelvis.

(c) Spectrogram of measured μD radar data.

(d) Spectrogram of measured μD radar data with MoCap data of the feet, hands, lower legs, and pelvis.

Figure 4.4: Spectra and corresponding MoCap data of a radially towards the sensor walking/running pedestrian. © 2018 IEEE [50]

Table 4.1: Radar Configurations

Parameter	Value
RF center frequency	77 GHz
RF bandwidth	2 GHz
Range resolution	0.075 m
Sampling frequency	10 MHz
Single chirp duration T	51.2 μs
IF samples per chirp m_s	512
Number of chirps	1024
Chirp repetition interval G	62.5 μs
Doppler resolution	0.03 m/s
Sensor height	0.65 m
Azimuth 3 dB beamwidth (TX)	51°
Azimuth 3 dB beamwidth (RX)	76.5°

(a) Spectrogram of measured μD radar data.

(b) Spectrogram of measured μD radar data with MoCap data of the feet, hands, lower legs, and pelvis.

Figure 4.5: Spectra and corresponding MoCap data of a pedestrian, who first walks longitudinally away from the radar and then crosses laterally the radar's field of view. © 2018 IEEE [50]
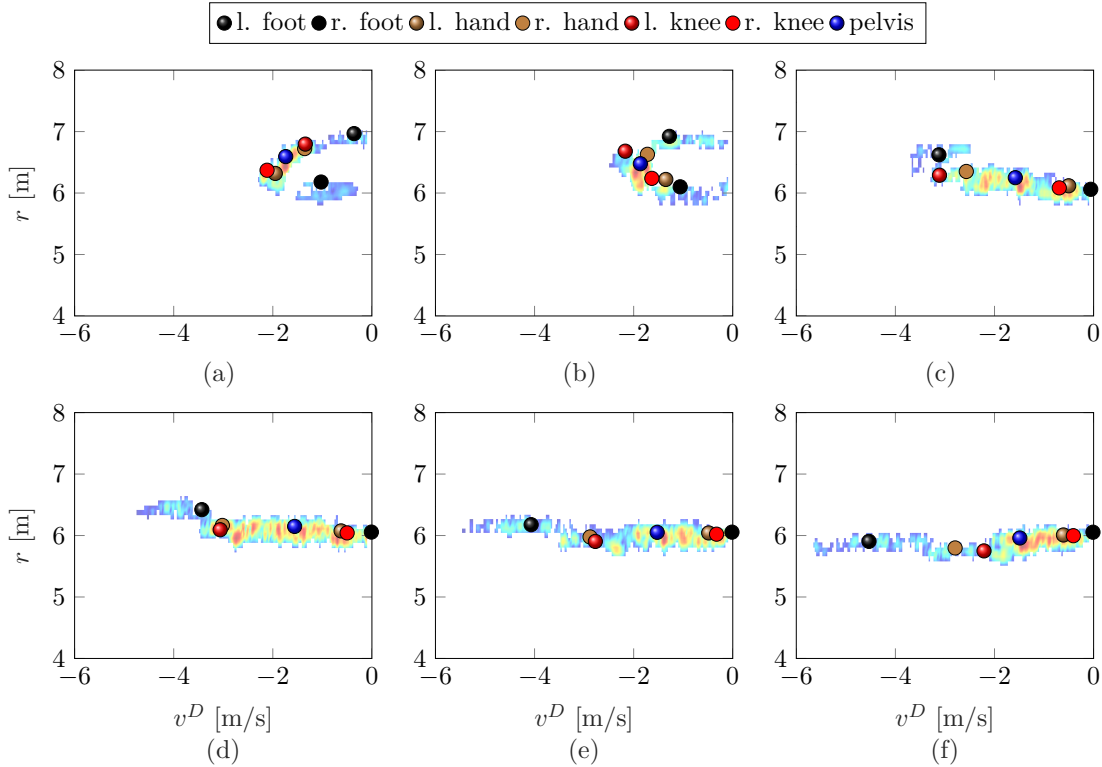
Figure 4.6: Range-Doppler frame sequence of radar and MoCap data during a quarter gait cycle. Adopted from [50] © 2018 IEEE.

4.2 Pedestrian feature extraction

Radar-based feature extraction of pedestrians for automotive applications is still a modern and relatively unresearched topic. Prior automotive radar sensors could not provide the required physical sensor attributes, i.e., resolutions, to detect distinctive features that can be utilized for more advanced procedures beyond mere classification. One of those techniques is the data-based identification of individual limbs while moving. The ability to separate, identify, and extract individual reflection components can be of enormous importance and raises the question of possibilities for a more precise radar-based determination of a pedestrian's intended behavior. So far, there are only a few approaches concerning the frequency-based separation of time-variant limb gestures based on automotive radar. In [81], *Raj et al.* present the first method that decomposes motion data of various body regions using an iterative, non-parametric tracking algorithm in the time-Doppler domain combined with a model of the human gait. However, the presented method is based exclusively on simulated radar reflections lacking the complexity of multi-reflective points of real radar measurements. In [35], *Fogle et al.* present a nonlinear least-squares approach to separate the scattering locations with an expectation-maximization algorithm that assigns the range-Doppler frequencies to scatter centers. By utilizing the range information, the authors achieve the extraction of individual body components of a pedestrian. Since their

research objective is encountered in the surveillance technology, the used radar parameters and the experimental setup strongly differ to the automotive field. In particular, the large elevational angle of incidence due to the sensor's height combined with an ultra-fine range resolution leads to an improved limb separation ability in the range dimension, which is not given in automotive radars. Moreover, both of the mentioned approaches require a large time window leading to a high computational burden making it inapplicable for real-time processing in automotive safety functions.

Another approach is introduced by *Abdulatif et al.* in [2]. Using a machine learning technique based on simulated radar returns of walking humans, they assign scatterers to four classes of body parts. While they achieve satisfying results for the feet and base movement, the assignment of the arms and legs underly high false-positive rates.

To summarize, the proposed procedures are not persuading for automotive applications due to simplified data or radar parameters such as large aspect angles and integration times. Instead, a methodology is desired that takes the automotive parameterization into account. One approach is the frequency-based localization and extraction of body components based on algorithmic segmentation of individual backscatterers at each CPI, presented in the following.

4.2.1 Frequency-based limb separation

The ability of radar-based identification of individual limbs enables the targeted recognition of a pedestrian's movement behavior and lays the foundation for behavioral prediction by detecting gestures. This requires procedures that can separate the spectral velocity components and assign them to the extremities. The following methodology, which is based on [95], aims to identify and extract relevant scatterers emanating from different body parts out of the range-Doppler spectrum to get information about the instantaneous motion state of a pedestrian that can be used for further applications, such as intention recognition. The clustered range-Doppler detections are forwarded to a modified CLEAN algorithm that subdivides the detections into potential body segments. Subsequent DOA estimation is performed, enabling the assignment to corresponding body components, which completes the proposed algorithm.

The approach acts on the range-Doppler spectrum after OS-CFAR power detection and DBSCAN object-clustering according to the provided signal processing in Section 2.3, as shown in Figure 4.7(a) for a instant of time during the stance phase. The used radar parametrization is given in Table 4.1. As can be seen, the limbs are encountered at various distances to the radar during a stride. This extension allows separating body elements according to their frequency-based location in the range-Doppler spectrum. However, the used high resolution comes with difficulties that need to be considered regarding the extraction process. Different body parts exhibit different backscattering intensities due to various sized reflection surfaces. Moreover, range and Doppler information suffer from orientation dependence due to the radial principle and depend intensely on the object's motion direction. To cover the mentioned challenges, an advanced extraction procedure is applied to the range-Doppler spectrum.

The proposed procedure is based on a modified CLEAN algorithm. Initially applied by *Högbom* in the field of radio astronomy to detect luminescent objects in the sky, the

CLEAN algorithm is an iterative deconvolution method that distinguishes between real data structures and data disturbances due to sidelobing in the spectrum of low power targets [58]. Typically, standard Fourier methods suffer from pronounced gating responses, such as sidelobe patterns, due to signal artifacts' gaps in the synthesized beams. The CLEAN algorithm treats the radar spectrum as a collective of independent point sources. The idea is then to exchange individual sidelobe suffering reflections in the Fourier spectrum by a "clean" beam pattern, i.e., an immaculate beam of equal shape but without disturbing sidelobes. This is done by iteratively detecting the maximum scatterer in the image and subtract a small gain convolved with the point spread function (PSF). By doing so, the Fourier spectrum gets decomposed into point targets yielding an increased spectral resolution. Several variants of the original CLEAN have been developed over the years, such as the coherent CLEAN algorithm for microwave applications, which additionally contains target phases, intensities, and directions [103].

Further approaches investigate the integration of clustering procedures within the CLEAN algorithm to define a stop criterion for the iteration [20], [106]. This is advantageous in so far as we are striving for a computationally fast calculation. Since CLEAN first starts the iteration at frequency regions with high amplitude, e.g., torso, these objects are resolved excessively before small intensity objects, e.g., feet, are regarded. For this reason, a clustering procedure is included in the CLEAN algorithm to mask regions that are already sufficiently resolved. The presented cluster-CLEAN execution comprises the following steps:

1) Find scatterer with maximal magnitude in the range-Doppler spectrum.

2) Generate the PSF with the scatterer's frequency and amplitude parameters.

3) Update scatter list (λ_i, μ_i, a_i).

4) Substract γ times PSF with $0 < \gamma < 1$ from the Fourier spectrum.

5) Starting from 1) and iterate until the cluster criterion $|a_i| <$ threshold for a finalized cluster is reached.

6) Mask finalized clusters in the Fourier spectrum

First, the highest amplitude is identified in the complex 2-D range-Doppler spectrum $\boldsymbol{X}_0(\lambda, \mu)$ and its frequency position (λ_1, μ_1) and complex amplitude a_1 are determined. The selected point scatterer is stored in a scatter list. Next, a weighted PSF is subtracted from the range-Doppler spectrum

$$\boldsymbol{X}_1(\lambda, \mu) = \boldsymbol{X}_0(\lambda, \mu) - \gamma \Psi_1, \tag{4.1}$$

where Ψ_1 is the PSF, and γ is the loop gain $0 < \gamma < 1$. Choosing $\gamma = 0.5$, a rapid calculation is achieved. In order to acquire satisfying results when applying CLEAN, the

PSF needs to be modeled precisely. However, in radar measurements, the PSF is usually not determinable entirely. Classically, point targets are modeled based on their frequency components, i.e., sinusoids. The PSF is then the result of an FFT transformed sinusoid, i.e., a multivariate sinc-function, with identical sampling and windowing to the measured data [85]. However, disturbing influences, such as dynamics like accelerations and μD effects, or hardware imperfections, lead to deviations from this idealized model. Analyses have shown that the Fourier transformed window functions used to calculate the range-Doppler spectrum, i.e., Hann windows, represent an adequate PSF for our purpose. Thus, the PSF as the response of the point scatterer can be given by

$$\Psi_1 = a_1 \boldsymbol{W}_\lambda(\lambda - \lambda_1) \boldsymbol{W}_\mu(\mu - \mu_1), \tag{4.2}$$

where $\boldsymbol{W}_\lambda$ and $\boldsymbol{W}_\mu$ are the Fourier transforms of the window functions $\mathbf{w}_\lambda(l_s)$ and $\mathbf{w}_\mu(m_s)$, respectively.

In the original CLEAN algorithm, (4.1) and (4.2) are iteratively repeated until a threshold, e.g., noise floor, is reached. As a modification, we implement a clustering procedure for the scatter list in between the CLEAN deconvolution steps. Various criteria for the definition of a completed cluster were tested. Best results have been achieved for determining an amplitude threshold covering the differences between maximum and minimum intensities inside a cluster. If the amplitude threshold is reached, i.e., a cluster is finalized, the point scatterers of this cluster are removed from the scatter list and the cluster region $\boldsymbol{X}_i(\lambda, \mu)$ is masked for consecutive iterations. Figure 4.7(b) displays the subdivided range-Doppler snapshot as a consequence of the provided cluster-CLEAN extraction algorithm in form of individual clusters comprising various scatterer points. Extracted scatterer points which are not fulfilling the defined cluster density criterion represent unassigned scatterers and hence are not included for further processing.

4.2.2 Extraction of body parts

After completing the clustering procedure using the cluster-CLEAN algorithm in the range-Doppler dimension, the angular frequency spectrum in azimuth direction is computed by means of DOA for each scatterer point belonging to a cluster. A subsequent magnitude maximum search is applied to estimate the clusters' angles yielding the parameters (r, v^D, ϕ). Figure 4.8 visualizes the result of the entire signal processing chain for the same snapshot used for Figure 4.7.

The algorithm is now applied to a sequence of radar snapshots to enhance the body part segmentation over time. The time-dependent μD spectrum of a pedestrian heading towards the radar is depicted in Figure 4.9(a). The section shows the typical μD signature revealing the periodic Doppler frequency oscillations generated by the limb motion. Similar to Figure 4.4(a), the feet responses induce the highest velocities of around $5\,\text{m/s}$ while swinging, and $0\,\text{m/s}$ while standing still. Moreover, the torso region dominates the backscattering intensities due to the enlarged reflection surface. In total, the spectrum expresses a superposition of various body reflections whose velocity trajectories are hard to follow.

For the assignment of scatter points to body components, the pedestrian's body is segmented into five regions along its lateral extent. The torso primarily denotes the body's central
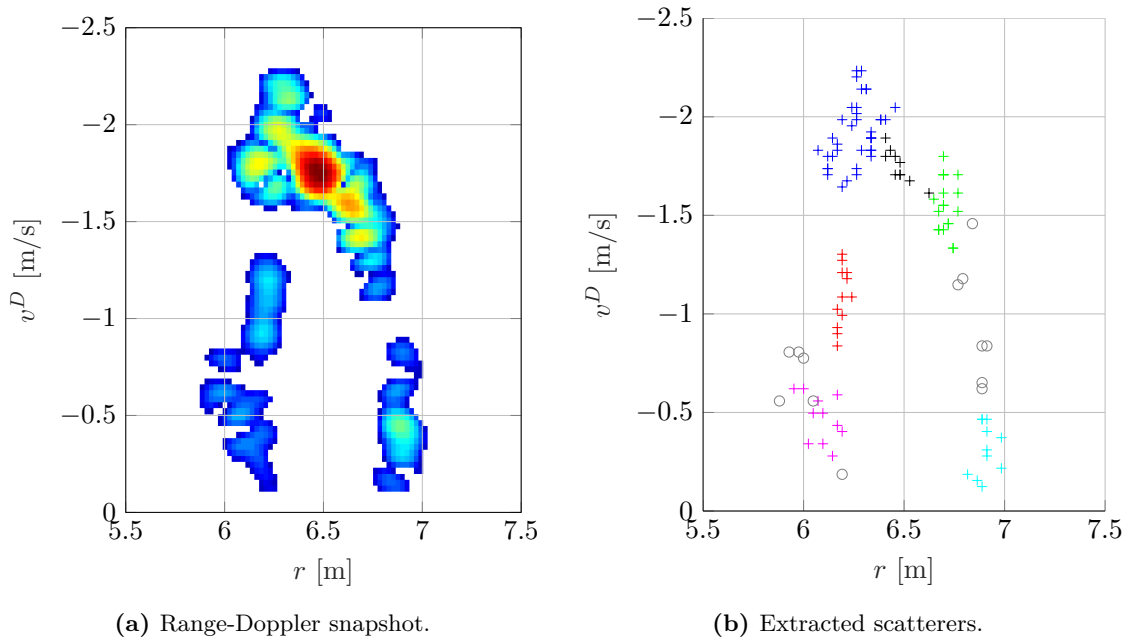
(a) Range-Doppler snapshot.

(b) Extracted scatterers.

Figure 4.7: Result of the cluster-CLEAN scatterer extraction. a) Power detected range-Doppler spectrum of a walking pedestrian. b) After cluster-CLEAN scatterer extraction. Cross marks with identical color form a cluster. Circles represent unassigned scatterers. © 2019 IEEE [95].
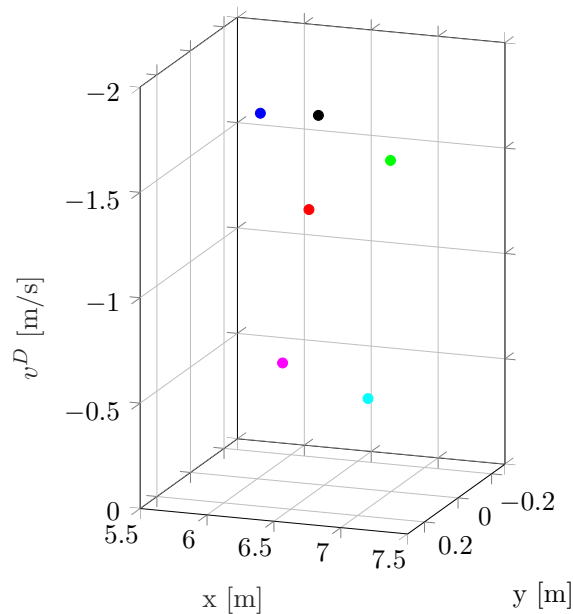

Figure 4.8: Result of the cluster-CLEAN algorithm for a walking pedestrian after angular frequency estimation. Each point represents the angular resolved maximum of the corresponding cluster. The polar coordinates (r, ϕ) are transformed into Cartesian ones (x, y). © 2019 IEEE [95].

region, as shown in Figure 4.9(b). The regions to the center's left and right sides are allocated to the left and right leg, respectively. Their assigned scatterer points are depicted in Figure 4.9(c). At the extreme left and right side of the body segments, the left and right arms are located. The MoCap data of the corresponding body component is additionally depicted in each subfigure to evaluate the extraction and assignment performance.

The torso assigned scatterers exhibit a periodic velocity course with slight amplitude changes for each stride, confirmed by the corresponding motion reference. The comparison between the legs' reference curves and the assigned scatterer points reveals the knees' presence as the second dominant scatter contribution. The upper and lower leg's anatomical connection by the knee joint induces different velocity amplitudes during the swing phase. Both components are fairly noticeable in the extracted radar data. The motion of the extracted arms assumes an almost sinusoidal trend, which is validated by the MoCap data.

4.2.3 Experimental results

The extracted data's periodic behavior with similar cycle times displays the limbs' overall interaction. The asynchronous motion behavior of the respective limbs to each other is fairly recognizable. Thus, the left arm's and left leg's velocity amplitudes behave in the opposite direction to their respective counterparts. Consequently, the extracted radar data describe essential information of the body's movement behavior during locomotion. The extracted data may include noise points and outliers that do not belong to the corresponding body part, as can be primarily seen in the extracted torso in Figure 4.9(b). Besides, there are periods in which data points are missing due to imperfection in the extraction procedure or occlusion by different body parts during the measurement. These outliers could be removed or significantly reduced by further signal processing methods. Overall, however, the extracted data match the reference data very closely, validating the provided extraction technique and demonstrating limb separation potential for automotive safety systems such as behavior prediction.

Besides radial walking trajectories, crossing movements are particularly relevant for future safety applications. Investigations were carried out in a measured crossing scenario proving the provided extraction procedure's fundamental feasibility. Despite the greatly increased complexity from a radar perspective, several body parts can be correctly extracted. However, the significantly increased quantity of data occlusion and considerably smaller Doppler amplitudes due to the body's orientation seriously reduces the separability and hence the number of correctly assigned body components compared to radial movements.
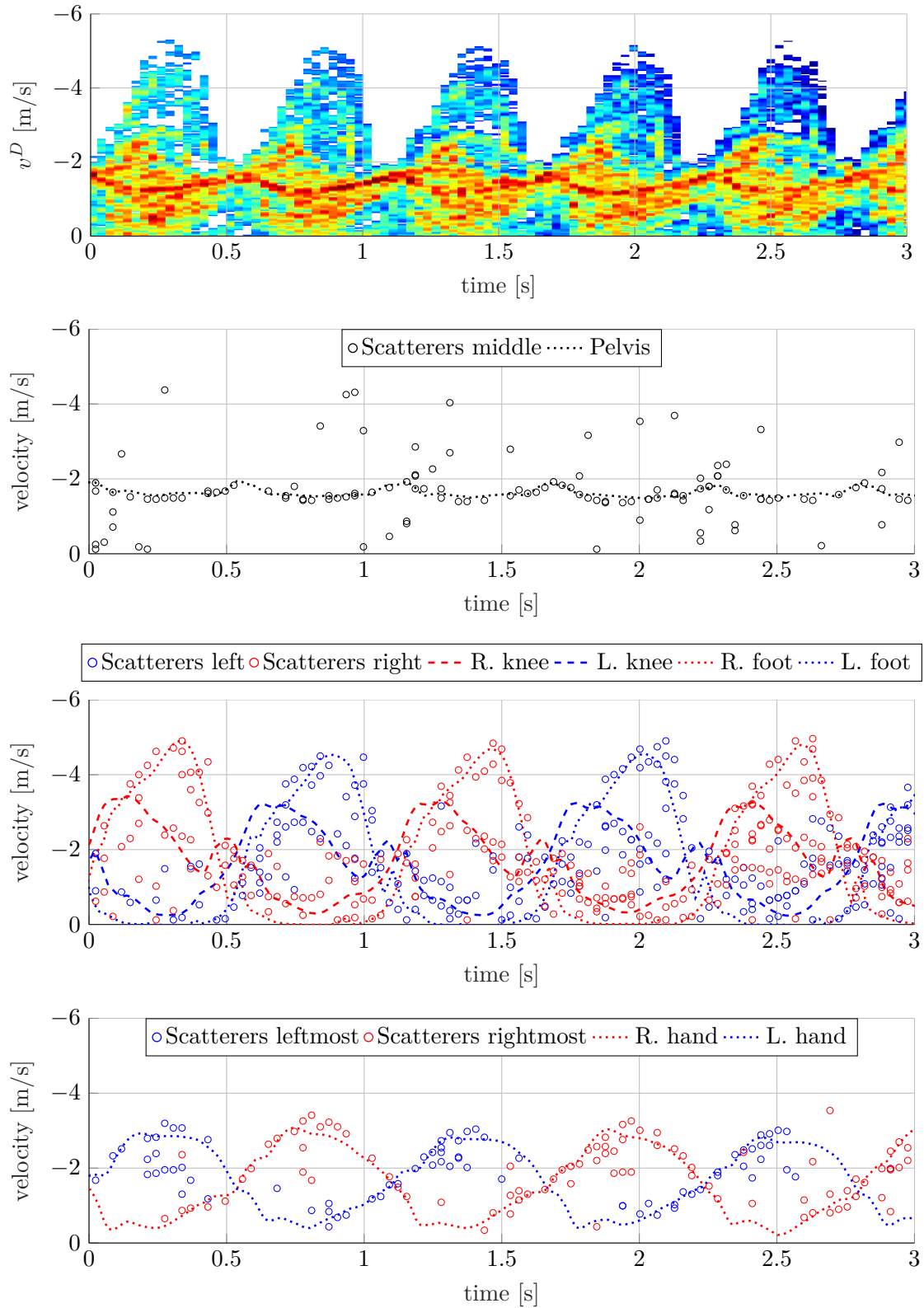
Figure 4.9: Micro-Doppler spectrum and time-dependent velocity of different body parts. (a) Micro-Doppler spectrum (Color-coding is equal to Figure 4.4). (b) Torso. (c) Legs. (d) arms. The marks result from the cluster-CLEAN extraction procedure of the radar measurement in (a). The individual MoCap data of the corresponding body parts are depicted in dashed lines. © 2019 IEEE [95].

4.3 Pedestrian tracking

Detailed environment perception for HAD requires accurate and consistent models that comprise dynamic and static attributes of surrounding targets. The processing of measured detections, acquired from various sources, to reveal and manage state estimates of objects is crucial for self-localization, i.e., mapping the ego-position by stationary landmarks, such as houses and trees, and targeting moving objects, e.g., pedestrians, in the vicinity of an autonomous vehicle. Estimation of stationary and moving objects refers to generating information about their locations, dynamics, orientations, as well as their shape and size. This information is used for interactive navigating and collision avoidance, i.e., to enable safe autonomous driving. In order to perceive this estimated surrounding knowledge, a highly automated vehicle is equipped with exteroceptive sensors, e.g., lidars, cameras, or radars.

In the automotive safety and driver assistance domain, pedestrian tracking is commonly located in computer vision applications [5], [78], [80]. Typical digital image processing techniques, such as feature-based classifiers or deep learning networks, are used for pedestrian classification. Once the object is classified for consecutive frames, it is regarded as a candidate object and forwarded to a tracking system that estimates its motion path using point-based tracking, kernel-based tracking, or silhouette-based tracking [80].

With time, electronic technology is substantially advancing toward an increase in hardware properties, i.e., higher resolution, and an enlarged field of view. Current radar sensors typically resolve multiple detections of an object in their surveillance area resulting in geometric and kinematic extents. Hence, extended-object tracking routines for robust state estimation, capable of handling challenges such as false alarms and missed detections, are required. In extended object tracking, one is more interested in estimating the extended object as a whole, i.e., its utter geometric and kinematic properties, instead of retrieving attributes of single points that cause the detections. Moreover, from a generalized perspective, the number of objects inside the surveillance area is time-varying and cannot be determined in advance due to the limited sensor properties, e.g., field of view, and disturbing influences such as occlusion or false detections. In this work, we investigate radar-based pedestrian tracking methods to improve the overall tracking performance and derive parameters and indicators that give rise to potential behavior prediction. Consequently, we focus on a single pedestrian in the field of view that is mostly free of the mentioned disturbances. However, it is not limited to single-object tracking, as we will see in the following.

In radar-based pedestrian tracking, diverse methods exist concerning the data level on which the tracking is established and the appropriate tracker's preference. Sequential Monte Carlo methods, i.e., particle filters, express a dominant practice for macroscopic pedestrian tracking, where either global signal maxima or even unfiltered measurements are considered [55], [56]. Due to particle filters' heuristic principle, intricate motion models can be usually bypassed, making them very attractive. Nevertheless, they demand remarkable computational effort and therefore appear as unsuitable for automotive real-time functions. Contrary, other approaches apply computationally manageable methods like standard Kalman filters [77], [107]. Kalman filter-based tracking is widely utilized throughout the radar community due to its straightforwardness. However, the simplicity is accompanied

by a loss of data information. In [107], the authors abstain from incorporating the Doppler velocity into the measurement vector due to the consequential non-linearity and resulting need for computationally intensive higher-order filters.

As already presented throughout this thesis, it is precisely the Doppler and μD information induced by the limb movements that provide unique features. Consequently, we will demonstrate that first-order approximations, e.g., the extended Kalman filter, are well suited to integrate μD components into the tracking procedure.

This chapter's overall objective is to present a novel tracking method based on the results obtained so far, which executes the targeted tracking of leg emerging radar reflections. It thus achieves highly accurate results by localizing the feet while walking and enables novel possibilities concerning the behavioral indication of pedestrians in road traffic.

4.3.1 Probabilistic state estimation

An automated vehicle's environment is a dynamically changing system with various attributes, whose information is acquired by the vehicle's sensors. Consequently, the environment is characterized by a collection of the environments' objects related attributes, i.e., states. Some of them comprise dynamic information, e.g., a moving car, where others remain static, e.g., a building's wall. The perceptual interaction results in a gathering of measurements that give information about the environments' momentary states. However, sensor data are typically deteriorated by noise, and they maintain merely partial information about those quantities. State estimation algorithms attempt to recover state variables from the data by computing hypotheses distributions since the evolution of states and measurements can be portrayed by probabilistic laws. In probabilistic robotics, a hypothesis denotes the robot's internal knowledge about the environment's state expressed as a conditional probability distribution that considers each possible hypothesis with respect to the true state. A probability density function (PDF) is utilized to represent the hypothesis by incorporating both the measurements' information as well as the uncertainty processes of the environment. In order to perform the process of state estimation, two models are required that characterize the dynamic system: the system model representing the state's evolution and a measurement model that associates the measurements to the state. The models' interplay is postulated in probabilistic form, which provides an explicit general framework for dynamic state estimation [102].

The Bayesian filtering approach represents the most general formulation for estimating hypotheses based on a sequence of received sensor measurements and control data. The Bayes filter recursively calculates the posterior PDF of a state in each time step by first predicting the prior density forward and then updating it using the latest measurement, as systematically illustrated in Figure 4.10. In the prediction step, the system model generally translates, deforms, and widens the prior density due to unknown system uncertainty. In the update step, the prior density is adjusted, i.e., narrowed, in the light of additional knowledge from recent data. In a nonlinear system or measurement model, the resulting posterior PDF will be non-Gaussian since it assimilates all available statistical information. Consequently, the posterior PDF provides a complete solution to an estimation problem [86].

Let $\mathbf{x}_k \in \mathbb{R}^{n_x}$ be a target state vector with dimension n_x and time index k. Assuming that

the prior hypothesis $\mathbf{x}_0$ at time $k = 0$ has been correctly initialized, the state propagates according to a generalized discrete-time model function [42], [86]

$$\mathbf{x}_k = \boldsymbol{f}_{k-1}\left(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}\right), \tag{4.3}$$

where $\boldsymbol{f}_{k-1}(\cdot)$ denotes a possibly nonlinear function of state $\mathbf{x}_{k-1}$. The calculation of the hypothesis over the state $\mathbf{x}_k$ at time k based on the prior hypothesis over state $\mathbf{x}_{k-1}$ at time $k-1$ indicates the Bayes' recursion in the prediction step and emphasizes the Markov assumption that postulates the state as a complete representation of the history. The measurement equation

$$\mathbf{z}_k = \boldsymbol{h}_k\left(\mathbf{x}_k, \mathbf{e}_k\right) \tag{4.4}$$

portrays the relation between measurements $\mathbf{z}_k \in \mathbb{R}^{n_z}$ with dimension n_z and the target state by a possibly nonlinear measurement function $\boldsymbol{h}_k(\cdot)$. $\mathbf{v}_{k-1}$ and $\mathbf{e}_k$ are considered to be white, mutually independent process and measurement noise with covariances $\mathbf{Q}_{k-1}$ and $\mathbf{R}_k$, respectively.

In a first step, the prior density is obtained by the Chapman-Kolmogorov prediction given by

$$p\left(\mathbf{x}_k|\mathbf{Z}_{1:k-1}\right) = \int p\left(\mathbf{x}_k|\mathbf{x}_{k-1}\right) p\left(\mathbf{x}_{k-1}|\mathbf{Z}_{1:k-1}\right) \mathrm{d}\mathbf{x}_{k-1}. \tag{4.5}$$

Here, the prior density $p\left(\mathbf{x}_k|\mathbf{Z}_{1:k-1}\right)$ is expressed as the marginalization of the previous state at time $k-1$ given measurements $\mathbf{Z}_{1:k-1}$ up to and including time $k-1$ and the transition density $p\left(\mathbf{x}_k|\mathbf{x}_{k-1}\right)$. Since we exploit the Markov assumption we can formulate $p\left(\mathbf{x}_k|\mathbf{x}_{k-1}, \mathbf{Z}_{1:k-1}\right) = p\left(\mathbf{x}_k|\mathbf{x}_{k-1}\right)$.

Given the predicted state density at time k and an accessible measurement $\mathbf{z}_k$, the update stage is performed via the Bayes' rule

$$\begin{aligned} p\left(\mathbf{x}_k|\mathbf{Z}_k\right) &= p\left(\mathbf{x}_k|\mathbf{z}_k, \mathbf{Z}_{1:k-1}\right) \\ &= \frac{p\left(\mathbf{z}_k|\mathbf{x}_k, \mathbf{Z}_{1:k-1}\right) p\left(\mathbf{x}_k|\mathbf{Z}_{1:k-1}\right)}{p\left(\mathbf{z}_k|\mathbf{Z}_{1:k-1}\right)} \\ &= \frac{p\left(\mathbf{z}_k|\mathbf{x}_k\right) p\left(\mathbf{x}_k|\mathbf{Z}_{1:k-1}\right)}{p\left(\mathbf{z}_k|\mathbf{Z}_{1:k-1}\right)}, \end{aligned} \tag{4.6}$$

where the normalizing denominator

$$p\left(\mathbf{z}_k|\mathbf{Z}_{1:k-1}\right) = \int p\left(\mathbf{z}_k|\mathbf{x}_k\right) p\left(\mathbf{x}_k|\mathbf{Z}_{k-1}\right) \mathrm{d}\mathbf{x}_k, \tag{4.7}$$

results from the likelihood function $p\left(\mathbf{z}_k|\mathbf{x}_k\right)$, expressed by the measurement model (4.4). The Bayes update uses $\mathbf{z}_k$ to adjust the prior density resulting in the posterior density and satisfies the conceptual Bayesian filtering recursion.

Unlike traditional model-based programming methods, probabilistic approaches appear to be more robust in complicated real-world scenarios that come with high uncertainty. Moreover, they exhibit weaker requirements regarding the accuracy of the vehicle's sensors and the underlying models. However, the benefits of probabilistic algorithms come at

a price. Since probabilistic algorithms attempt to compute entire probability densities rather than single guesses, they tend to be computationally unmanageable, and hence approximations or suboptimal Bayesian algorithms are needed [102]. Generally, we are striving for compact parametric approximations, e.g., Gaussians, in order to sufficiently resemble the real density.

4.3.2 Gaussian filters

As a part of recursive state estimators, Gaussian filters all share the basic idea of representing true hypotheses by multivariate normal distributions. By doing so, Gaussian filters formulate the earliest possible implementations of the Bayes recursion for continuous spaces. The idea of approximating the posterior by a Gaussian comes along with essential benefits. First of all, Gaussians denote compact models that can be entirely parameterized by a probability distribution's first and second moments, i.e., it's mean and covariance. A general form of the multivariate normal distribution can be given by [102]

$$p(\mathbf{x}) = \det\left(2\pi\mathbf{\Sigma}\right)^{-\frac{1}{2}} e^{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T\mathbf{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})}, \tag{4.8}$$

where the density over the multivariate random variable $\mathbf{x}$ is expressed by the mean, i.e., the expectation of the random variable $\mathbf{x}$ given by

$$\boldsymbol{\mu} = \mathbb{E}\left[\mathbf{x}\right] = \int x p(x)\mathrm{d}x, \tag{4.9}$$

and the quadratic and positive-semidefinite covariance matrix, i.e., the squared expected deviation from the mean given by

$$\mathbf{\Sigma} = \mathrm{Cov}\left[\mathbf{x}\right] = \mathbb{E}\left[\mathbf{x} - \mathbb{E}\left[\mathbf{x}\right]\right]^2. \tag{4.10}$$

Moreover, the unimodal Gaussians, i.e., single maximum distributions, match the typical posteriors of numerous tracking applications in which the posterior models the true state with some uncertainty, and hence not many distinct hypotheses contribute to the true posterior.

4.3.3 The Kalman filter

Originally invented by *Swerling* and *Kalman* in 1958 and 1960, respectively, the Kalman filter represents the most researched and established technique for implementing Bayesian
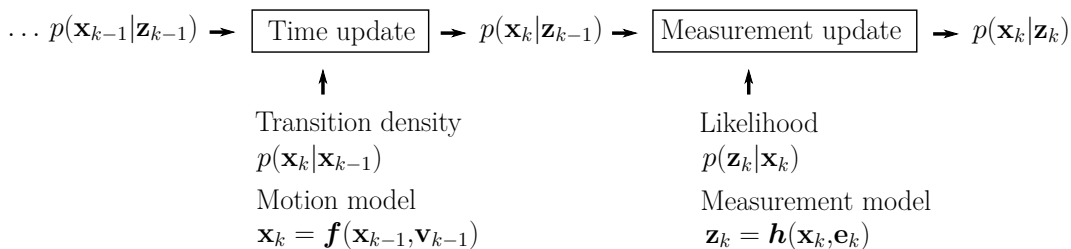

Figure 4.10: A systematic description of the Bayes recursion [42].

filtering in linear Gaussian systems [12]. In order to do so, the Kalman filter uses moments parametrization, i.e., mean $\boldsymbol{\mu}_k$ and covariance matrix $\mathbf{P}_k$ at time k, to calculate hypotheses. It can be proved that if the prior state probability density, $p\left(\mathbf{x}_{k-1}|\mathbf{Z}_{1:k-1}\right)$, appears to be Gaussian, the resulting posterior density, $p\left(\mathbf{x}_k|\mathbf{Z}_k\right)$, is Gaussian as well, presumed that the following three conditions hold: [57], [86], [102]

1) $\mathbf{v}_{k-1}$ and $\mathbf{e}_k$ are normally distributed with known parameters.

2) $\boldsymbol{f}_{k-1}\left(\mathbf{x}_{k-1}, \mathbf{v}_{k-1}\right)$ must be a known linear function (or a locally linearized function) in its arguments $\mathbf{x}_{k-1}$ and $\mathbf{v}_{k-1}$.

3) $\boldsymbol{h}_k\left(\mathbf{x}_k, \mathbf{e}_k\right)$ must be a known linear function (or a locally linearized function) in its arguments $\mathbf{x}_k$ and $\mathbf{e}_k$.

Those three properties ensure that the posterior hypothesis appears to be Gaussian at any point in time k. Consequently, (4.3) and (4.4) can be rewritten as [86]

$$\mathbf{x}_k = \mathbf{F}_{k-1}\mathbf{x}_{k-1} + \mathbf{v}_{k-1} \tag{4.11}$$

$$\mathbf{z}_k = \mathbf{H}_k\mathbf{x}_k + \mathbf{e}_k. \tag{4.12}$$

Now, the former nonlinear state and measurement functions are expressed in linear representation with matrix $\mathbf{F}_{k-1}$ (of dimension $n_\mathbf{x} \times n_\mathbf{x}$) and matrix $\mathbf{H}_k$ (of dimension $n_\mathbf{z} \times n_\mathbf{x}$). Considering the Bayes prediction and update rules in (4.5) and (4.6), the Kalman filter algorithm follows a recursive relation

$$p\left(\mathbf{x}_{k-1}|\mathbf{Z}_{k-1}\right) = \mathcal{N}\left(\mathbf{x}_{k-1}; \hat{\mathbf{x}}_{k-1|k-1}, \mathbf{P}_{k-1|k-1}\right) \tag{4.13}$$

$$p\left(\mathbf{x}_k|\mathbf{Z}_{k-1}\right) = \mathcal{N}\left(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}\right) \tag{4.14}$$

$$p\left(\mathbf{x}_{k-1}|\mathbf{Z}_k\right) = \mathcal{N}\left(\mathbf{x}_k; \hat{\mathbf{x}}_{k|k}, \mathbf{P}_{k|k}\right), \tag{4.15}$$

where $\mathcal{N}\left(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}\right)$ denotes a normally distributed density with variable $\mathbf{x}$, mean $\boldsymbol{\mu}$, and covariance matrix $\boldsymbol{\Sigma}$ according to the defined normal distribution in (4.8). The complete Kalman filter algorithm for linear normally distributed state transitions and measurements reads

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{F}_{k-1}\hat{\mathbf{x}}_{k-1|k-1} \tag{4.16}$$

$$\mathbf{P}_{k|k-1} = \mathbf{Q}_{k-1} + \mathbf{F}_{k-1}\mathbf{P}_{k-1|k-1}\mathbf{F}_{k-1}^T \tag{4.17}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\left(\mathbf{z}_k - \mathbf{H}_k\hat{\mathbf{x}}_{k|k-1}\right) \tag{4.18}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^T \tag{4.19}$$

where

$$\mathbf{S}_k = \mathbf{H}_k\mathbf{P}_{k|k-1}\mathbf{H}_k^T + \mathbf{R}_k, \tag{4.20}$$

denotes the covariance of the innovation $\boldsymbol{\varepsilon} = \mathbf{z}_k - \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}$, and

$$\mathbf{K}_k = \mathbf{P}_{k|k-1} \mathbf{H}_k^T \mathbf{S}_k^{-1}, \tag{4.21}$$

is the Kalman gain.

Equations (4.16) and (4.17) represent the calculation of the predicted hypothesis $\hat{\mathbf{x}}_{k|k-1}$ and $\mathbf{P}_{k|k-1}$ by considering the recursion of the state and covariance propagation. Due to the quadratic appearance of $\mathbf{P}_{k|k-1}$, the linear transition matrix $\mathbf{F}_{k-1}$ is multiplied twice into the covariance $\mathbf{P}_{k-1|k-1}$.

The predicted hypothesis is now converted into the absolute hypothesis by including the measurement $\mathbf{z}_k$ in (4.18). The Kalman gain $\mathbf{K}_k$ determines to what degree the measurement influences the updated state estimate by weighting the innovation in (4.18). Finally, the updated posterior hypothesis is calculated in (4.19) by incorporating the adjusted measurement information.

The Kalman filter represents the optimal solution to any tracking problem if the mentioned conditions hold. Moreover, its computation is reasonably efficient due to simple matrix additions and multiplications. The inversion complexity of the innovation's covariance matrix in (4.21) can be approximately given with $\mathcal{O}\left(d^{2.4}\right)$ for a matrix of size $d \times d$ and hence determines the lower bound of each iteration [102].

4.3.4 The extended Kalman filter

In reality, typical tracking applications may be very complex and are rarely linear. Consequently, they appear nonlinear and non-Gaussian, whose dynamic behavior cannot be described by linear state transitions. As a result, the linear Kalman filter's unimodal hypothesis assumption is only approximately applicable to most tracking problems.

A variety of suboptimal solutions have been proposed over the years tackling the need for nonlinear approximations. Some of them strive for an analytic approximation, such as the extended Kalman filter (EKF)) [3], [86], [102]. Others try to solve the non-linearity via grid-based numerical integration, such as the Viterbi algorithm [105]. Gaussian sum filters, e.g., static and dynamic multiple-model estimators, follow the idea to approximate the posterior hypothesis by a weighted sum of normally distributed densities [12], [86]. The last group refers to a sampling approach where the posterior approximation represents a deterministically sampled Gaussian density, such as the unscented Kalman filter (UKF) [86], [102].

A computational efficient but yet elegant procedure comes with the generalized, i.e., extended and non-closed-form solution given by the EKF. The EKF relaxes the linearity assumption in Kalman filtering by utilizing a nonlinear generalized Gaussian approximation. Instead of calculating the exact hypothesis, the EKF estimates its mean and covariance. This is achieved by the use of linearization, which manifests the key concept in EKF.

The generalized nonlinear systems in (4.3) and (4.4) are specified for the EKF approach with additive noise given by

$$\mathbf{x}_k = \boldsymbol{f}_{k-1}\left(\mathbf{x}_{k-1}\right) + \mathbf{v}_{k-1} \tag{4.22}$$

$$\mathbf{z}_k = \boldsymbol{h}_k\left(\mathbf{x}_k\right) + \mathbf{e}_k. \tag{4.23}$$

The EKF algorithm mechanics appear to be almost equivalent to its linear counterpart introduced in Subsection 4.3.3 except for the linearization procedure. The hypothesis propagation is repeated for convenience [86]

$$\hat{\mathbf{x}}_{k|k-1} = \boldsymbol{f}_{k-1}\left(\hat{\mathbf{x}}_{k-1|k-1}\right) \tag{4.24}$$

$$\mathbf{P}_{k|k-1} = \mathbf{Q}_{k-1} + \hat{\mathbf{F}}_{k-1}\mathbf{P}_{k-1|k-1}\hat{\mathbf{F}}_{k-1}^T \tag{4.25}$$

$$\hat{\mathbf{x}}_{k|k} = \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k\left(\mathbf{z}_k - \boldsymbol{h}_k\left(\hat{\mathbf{x}}_{k|k-1}\right)\right) \tag{4.26}$$

$$\mathbf{P}_{k|k} = \mathbf{P}_{k|k-1} - \mathbf{K}_k\mathbf{S}_k\mathbf{K}_k^T, \tag{4.27}$$

with

$$\mathbf{S}_k = \hat{\mathbf{H}}_k\mathbf{P}_{k|k-1}\hat{\mathbf{H}}_k^T + \mathbf{R}_k \tag{4.28}$$

$$\mathbf{K}_k = \mathbf{P}_{k|k-1}\hat{\mathbf{H}}_k^T\mathbf{S}_k^{-1}. \tag{4.29}$$

The local linearization of $\boldsymbol{f}_{k-1}$ and $\boldsymbol{h}_k$ by their first order Taylor expansion remains the Gaussian assumptions to be valid. The resulting Jacobians, $\hat{\mathbf{F}}_{k-1}$ and $\hat{\mathbf{H}}_k$, evaluated at $\mathbf{x}_{k-1|k-1}$ and $\mathbf{x}_{k|k-1}$, respectively, are defined by

$$\hat{\mathbf{F}}_{k-1} = \left[\nabla_{\mathbf{x}_{k-1}}\boldsymbol{f}_{k-1}^T\left(\mathbf{x}_{k-1}\right)\right]^T|_{\mathbf{x}_{k-1}=\hat{\mathbf{x}}_{k-1|k-1}} \tag{4.30}$$

$$\hat{\mathbf{H}}_k = \left[\nabla_{\mathbf{x}_k}\boldsymbol{h}_k^T\left(\mathbf{x}_k\right)\right]^T|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k-1}}, \tag{4.31}$$

where

$$\nabla_{\mathbf{x}_k} = \left[\frac{\partial}{\partial\,\mathbf{x}_k[1]}\cdots\frac{\partial}{\partial\,\mathbf{x}_k[n_{\mathbf{x}}]}\right]^T, \tag{4.32}$$

with $\mathbf{x}_k[i]$, $i = 1,\ldots,n_{\mathbf{x}}$ being the ith component of vector $\mathbf{x}_k$. The elements of $\hat{\mathbf{H}}_k$ can be given by

$$\hat{\mathbf{H}}_k[i,j] = \frac{\partial\,\boldsymbol{h}_k[i]}{\partial\,\mathbf{x}_k[j]}|_{\mathbf{x}_k=\hat{\mathbf{x}}_{k|k-1}}, \tag{4.33}$$

where $\boldsymbol{h}_k[i]$ being the ith component of vector $\boldsymbol{h}_k(\mathbf{x}_k)$.

Figure 4.11 demonstrates the principle of a nonlinear transformation on a normally distributed random variable $\mathbf{x} \sim \mathcal{N}\left(\mathrm{x};\mu,\Sigma\right)$ with mean μ and variance Σ. We assume that $\mathbf{x}$ is passed through the highly nonlinear function f resulting in a non-Gaussian, heavily skewed distributed density due to the nonlinearities in f [102]. The probability densities of the initial random variable $\mathbf{x}$ and the transformed random variable, $\mathbf{y} = f(\mathbf{x})$, are marked by the gray area in the lower right and upper left plot, respectively. As already stated, the linearization procedure in EKF approximates the function f by its first derivative at a single point, i.e., the tangent function to f at the Gaussian's mean (dashed line in the upper right plot). By doing so, a projection of the Gaussian density is achieved (dashed line in the upper left plot). Note that the linearization is also applied to an involved measurement function h in order to retain the Gaussian nature in the posterior density. As a reference, the mean and variance from an exact Monte-Carlo estimate is depicted additionally (solid

line in the upper left plot) [102]. The deviation between the Gaussian projection and the Gaussian reference reveals the error caused by linearizing f. In general, the more local nonlinearities arise in f, the poorer the quality of approximations. There are EKF-based approaches in the literature trying to overcome linearization errors. Some of them refer to the approximation's extension by incorporating further terms in the Taylor expansion [74]. Others perform the linearization on the updated measurement model instead of on the predicted state, such as the iterated EKF [12]. However, their increased computational demand restricted their overall usage.

Another aspect that impacts the approximation result is the degree of uncertainty in state estimation. A higher prior covariance yields less accurate estimations of the projected mean and covariance due to the sole regard of the mean by the Taylor expansion. An approach that might handle both the degree of local nonlinearity and the degree of uncertainty better than the EKF while still highly efficient is the UKF. The UKF appears to be a derivative-free filter that uses weighted deterministically determined samples to resemble the projected Gaussian density [102].

Despite all, we will see later in this work that both mentioned challenges can be tackled by keeping the functions approximated at moderate nonlinearity and choosing specific initialization parameters. Overall, the crucial benefit of the EKF lies in its simplicity and its computational efficiency. The calculation of the posterior hypothesis depends on the state and measurement vectors' dimensions and generally comes with low computing times of $\mathcal{O}\left(n_{\mathbf{z}}^{2.4} + n_{\mathbf{x}}^2\right)$ [102]. Since we are striving for a tracking methodology that is eventually applicaple in automotive safety functions, the EKF provides a thoroughly attractive approach.

4.3.5 Multiple-object tracking

Multiple-object tracking (MOT), as the generalization of object tracking, denotes the joint estimation problem of effectively determining the number of objects and their states based on noisy sensor measurements. Traditionally, MOT has been used for far-field applications, e.g., radar-based air surveillance, where an object generates at most one detection due to the sensor's resolution. Hence, classical MOT approaches assume independently evolving objects following the point object assumption (multiple-point object tracking (MPOT)) and thus do not incorporate any object's spatial or kinematic extent. Near-field applications, such as HAD, demand detailed environment perception and provide high-resolution sensors, leading to extended objects. In real-world tracking applications, extended objects, e.g., cars, bicycles, or humans, generate various spatially and kinematically distributed reflections and expand classical MOT to a multiple-extended object tracking (MEOT) phenomenon [42].

Usually, in extended object tracking, a sparse point cloud is available at each time step, which can significantly deviate from the previous state due to sensor characteristics and object motion. Therefore, the objects' attributes, e.g., their shape, cannot be extracted based on a single sensor scan. Instead, several measurements comprising the (unknown) objects' motion are needed to derive specific target-related information.
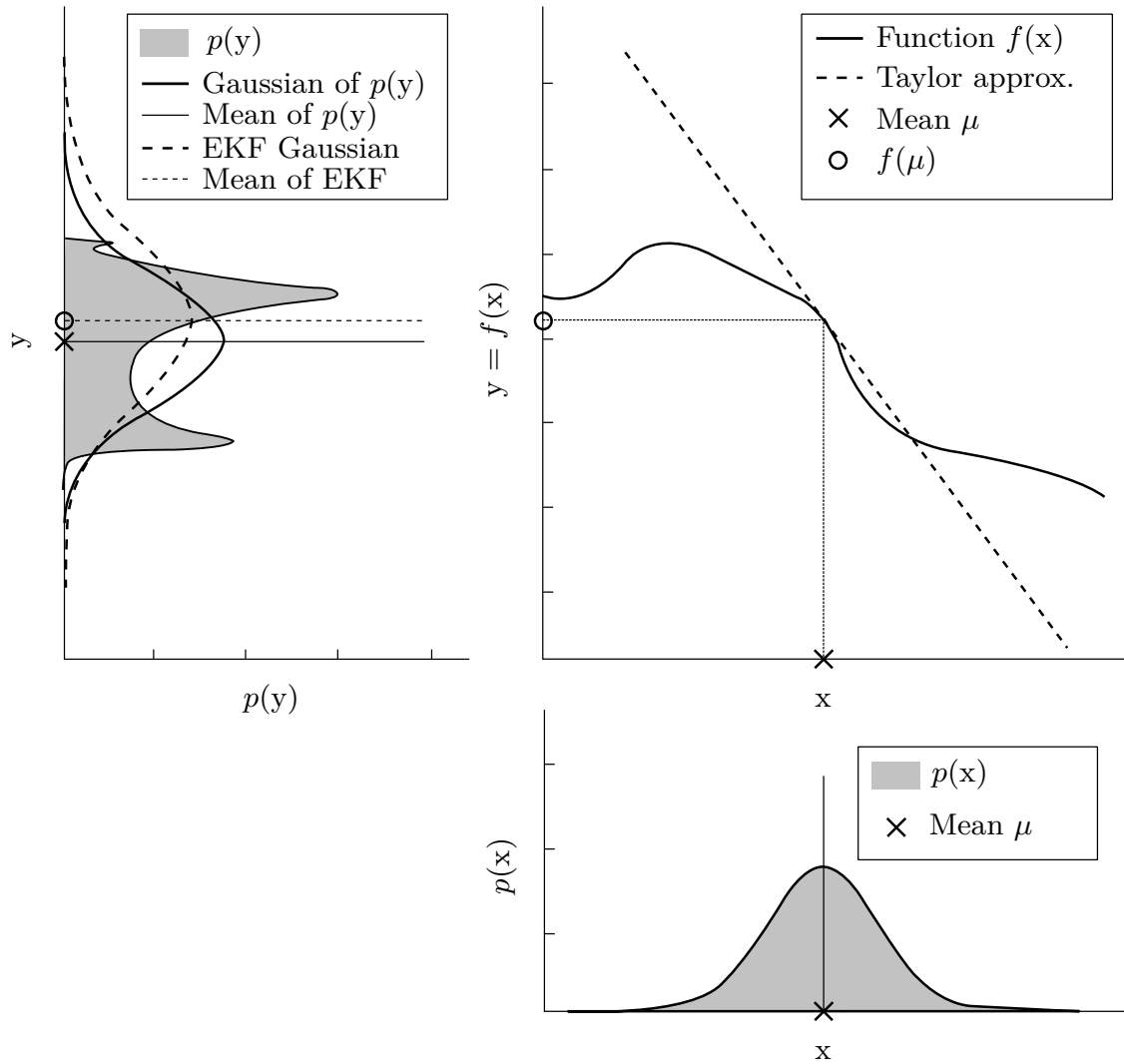
Figure 4.11: Representation of the EKF's linearization principle based on a Gaussian projection through a highly nonlinear function. The EKF passes a Gaussian through a first-order Taylor approximated function, i.e., the tangent to f at the mean of the initial Gaussian, rather than employing the nonlinear function f. The upper left graph depicts the resulting Gaussian (dashed line), as well as a reference Gaussian calculated from the exact Monte-Carlo estimation (solid). The mismatch between the Gaussians highlights the approximation error that results from linearization [102].

Regardless of the object type to be tracked, MOT comes with several challenges: [43]

- Unknown and time-varying number of objects.

- Missed detections, i.e., reduced or no received measurements of an object at a time.

- Unknown and varying number of object detections.

- Clutter measurements, i.e., non-target-originating measurements.

- Unknown measurement origins, i.e., unknown soure of each measurement.

A changing number of objects in the sensor's surveillance region is referred to as data handling during the tracking process. Usually, a time-varying object number can be either caused by *track birth* and *track death*, i.e., the emergence of new objects and the deletion of old objects, or by missed detection effects.

Missing detections occur by temporal blindness, often caused by occlusion by other objects, e.g., cars, or intra-object occlusion, e.g., hidden car wheels or human limbs. Hence, the detection probability of an object is always less than one since we do not know whether or not an object generated any detection.

As already stated, extended objects give rise to a varying number of detections depending on the sensors' resolution. In addition to the explicit resolution capabilities, an object's actual reflective surface can vary and influence the emerging detections.

Extended object tracking typically comes with false alarm detections known as clutter detections caused by the sensor itself or by environmental influences. This implies that the detection origin is generally unknown, meaning that there is no information on whether a detection is caused by clutter or by an actual object. Besides, it is not apparent which object caused which detection.

Let n_k be the number of objects present at time k, and let $\mathbf{x}_k^i$ be the state of object i at time k. Consequently, $\mathbf{X}_k$ defines all present object states at time k given by [11], [42]

$$\mathbf{X}_k = \left\{ \mathbf{x}_k^i \right\}_{i=1}^{n_k}. \tag{4.34}$$

Let m_k be the number of detections $\mathbf{z}_k^j$, then the set of detections at time k is

$$\mathbf{Z}_k = \left\{ \mathbf{z}_k^j \right\}_{j=1}^{m_k}. \tag{4.35}$$

Typically, the object states are assumed to be initially independent. Hence, the transition density for the n states describing the evolution from $\mathbf{x}_{k-1}^i$ to $\mathbf{x}_k^i$

$$p_k \left(\mathbf{X}_k | \mathbf{X}_{k-1} \right) = p_k \left(\mathbf{x}_k^1, \mathbf{x}_k^2, \ldots, \mathbf{x}_k^i, \ldots, \mathbf{x}_k^n \middle| \mathbf{x}_{k-1}^1, \mathbf{x}_{k-1}^2, \ldots, \mathbf{x}_{k-1}^i, \ldots, \mathbf{x}_{k-1}^n \right),$$

can be formulated as the product of the joint densities

$$p_k\left(\mathbf{X}_k|\mathbf{X}_{k-1}\right) = \prod_{i=1}^{n_k} p_k\left(\mathbf{x}_k^i|\mathbf{x}_{k-1}^i\right). \tag{4.36}$$

In case of assumed density filtering, e.g., Gaussian approximations for the object state densities [43] to perform the Bayes recursion with a predictable complexity, the initial prior at $k = 0$ can be described as a product of Gaussian densities with mean vectors $\boldsymbol{\mu}_0^i$ and covariance matrices $\boldsymbol{\Sigma}_0^i$ given by

$$p\left(\mathbf{X}_0\right) = \prod_{i=1}^{n_0} \mathcal{N}\left(\mathbf{x}_0^i; \boldsymbol{\mu}_0^i, \boldsymbol{\Sigma}_0^i\right). \tag{4.37}$$

The goal in MOT is the estimation of the object states $\mathbf{X}_k$ as a recursive computation of the posterior density $p\left(\mathbf{X}_k|\mathbf{Z}_k\right)$ given the measurements $\mathbf{Z}_k$. Note that the set of measurements $\mathbf{Z}_k$ may comprise not only object originated detections but also clutter detections $\mathbf{C}_k$. For the sake of completeness, we expand the set of measurements at time k as a random permutation according to

$$\mathbf{Z}_k = \sqcap\left(\mathbf{O}_k, \mathbf{C}_k\right), \tag{4.38}$$

where $\mathbf{O}_k = \left[\mathbf{o}_k^1, \ldots, \mathbf{o}_k^i, \ldots, \mathbf{o}_k^{n_k}\right]$ denotes the set of object detections and $\mathbf{C}_k = \left[\mathbf{c}_k^1, \ldots, \mathbf{c}_k^{m_k^c}\right]$ denotes the set of clutter detections where m_k^c is the number of clutter detections at time k. Note that the vectors of $\mathbf{C}_k$ are independent and identical distributed. Typically, the occurrence of clutter is modeled as a Poisson point process (PPP) and hence assumed to be Poisson distributed throughout the sensor's field of view according to $m_k^c \sim \mathrm{Po}\left(\bar{\lambda}_c\right)$ where $\bar{\lambda}_c$ denotes the mean and variance of the Poisson distribution $\mathrm{Po}(\cdot)$ [11].

Overall, the width of $\mathbf{Z}_k$ is random in every time step, and we do not know which, if any, detection in $\mathbf{Z}_k$ is an object originated detection. Hence, the problem of unknown measurement origins, i.e., all possible associations between measurements and objects, arises, which is arguably the most challenging aspect of MOT.

4.3.6 Data association

The objective of this chapter is to develop a kinematic-based framework for tracking the leg movements of a walking pedestrian based on high-resolution automotive radar data. To ensure an effective assignment of the detections to their sources, and thus an identification of the left and right leg, the implementation of a data association procedure is mandatory whose algorithmic context is introduced in this subsection. Figure 4.12 illustrates the lower body of a walking human and exemplarily visualizes the leg emanating detections as well as potential clutter detections that have to be associated to the respective source in each sensor cycle in order to enable the leg tracking approach.

MOT constitutes increased complexity as the problem of data association or correspondence problem arises. In data association, each detection has to be associated with one of the detection emanating sources., i.e., either to an actual object or clutter, to define the measurements' origins. In the case of MPOT, at most one detection can be caused by
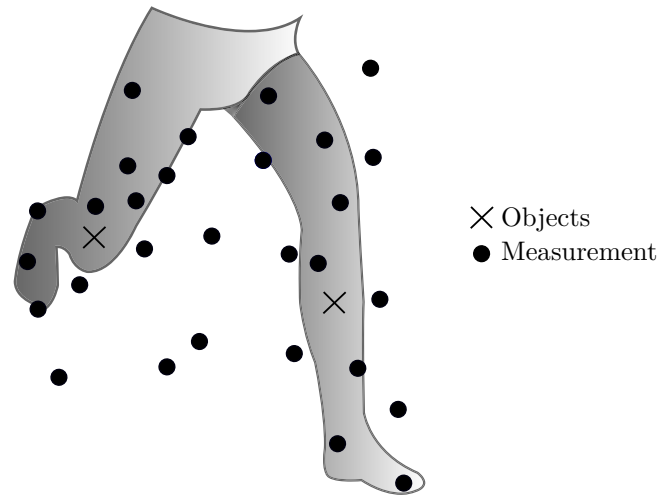
Figure 4.12: Exemplary depiction of the leg emanating detections and potential clutter detections of a walking human.

each object due to the nature of a point object. Figure 4.13(a) illustrates a simple MPOT scenario where we assume to have known data associations. The diagram visualizes the objects' positions over three consecutive time steps. In the case of data handling, decisions about the measurements' belonging regarding newly appeared and already existing objects have to be made additionally, which can rapidly lead to a tremendous and hence intractable computational effort. In the given case, where we assume to have known data associations, the corresponding problem becomes trivial. Normally, we do not have any prior information about the origin of the detections. Hence, we cannot tell which detections are from already detected objects, which detections are from newly appeared objects, and which detections are false detections. Figure 4.13(b) images the real scenario with unknown data association certainties. In this case, any measurement at any time step could be from an object that was already detected in a previous time step, or a newly appeared object, or a false detection. Hence, each detection is labeled either object-originating or clutter in every step, resulting in a vast number of possible association events. Accurate data association is significant because an erroneous association solution can result in fatal filtering performance. There are several solutions to MPOT regarding the consideration of their measurement-to-track hypotheses, such as the global nearest neighbor (GNN) algorithm [13], multiple hypothesis tracking (MHT) [19], [83], and the joint probabilistic data association (JPDA) filter [11], [13], [37].

  Contrary to point objects, an extended object causes more than a single detection in each sensor cycle and thus a data association, i.e., an association from measurement to source, is generally achieved in two parts: [43]

1) **Partition**: The partition of a set of measurements can be regarded as a division of the set's elements into non-empty subsets, called cells. The division is carried out such that every element of a subset is assigned to one and only one cell, i.e., the cell's measurements all belong to the same source, which can be an object or clutter.
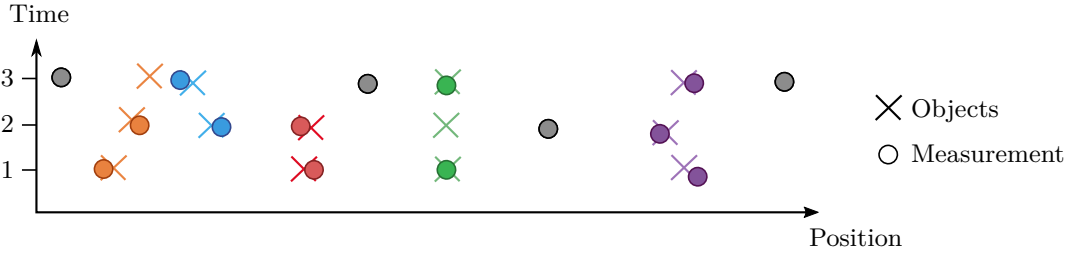
(a) Trivial case with known data associations: The colored crosses indicate the objects' positions over time and the measurements are color-coded according to the known associations in each time step. The gray measurements represent clutter or false detections, respectively.
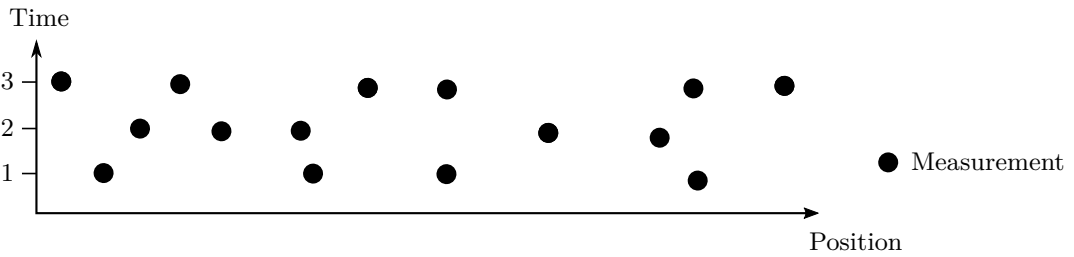
(b) Real case with unknown data associations: No prior information about the measurements' origins. Hence each measurement can be labeled either object-originating or clutter in each time step.

Figure 4.13: Exemplary illustration of the data association uncertainty using a 1-D scenario comprising a time varying number of evolving objects (time step 1: 4 objects, time step 2: 5 objects, time step 3: 4 objects) in position over time.

2) **Cell association**: The partitioned cells are associated to a measurement source, which can be either one of the objects or clutter.

Concerning Bayes's optimality, evaluating all feasible data associations in the MEOT update step is essential. Hence, the consideration of all partitions of a measurements' set is necessary, as well as each partition's feasible cell associations. In non-vanishingly small numbers of measurements and objects, the partitioning quickly takes on dimensions that cannot be realized due to numerous possible partitions and cell associations whose likelihood functions all have to be incorporated in the measurement update. Assuming a set of n measurements, the number of feasible partitioning ways is defined by the nth logarithmically convex Bell number $B(n)$ [43], [90]. For a measurement number $n = 3$ there are $B(3) = 5$ feasible partitions, as exemplarily illustrated in Figure 4.14. Duplicating the number of measurements ($n = 6$) yields $B(6) = 203$ feasible partitions, and for $n = 90$ we get $B(90) > 10^{100}$ feasible partitions. From those numbers, it gets apparent that the consideration of all feasible partitions is computationally unmanageable, and simplifications are needed.

Different approaches exist regarding the measurement modeling in MEOT procedures. Typically, physical modeling is applied where specific geometric shapes, e.g., an ellipse or
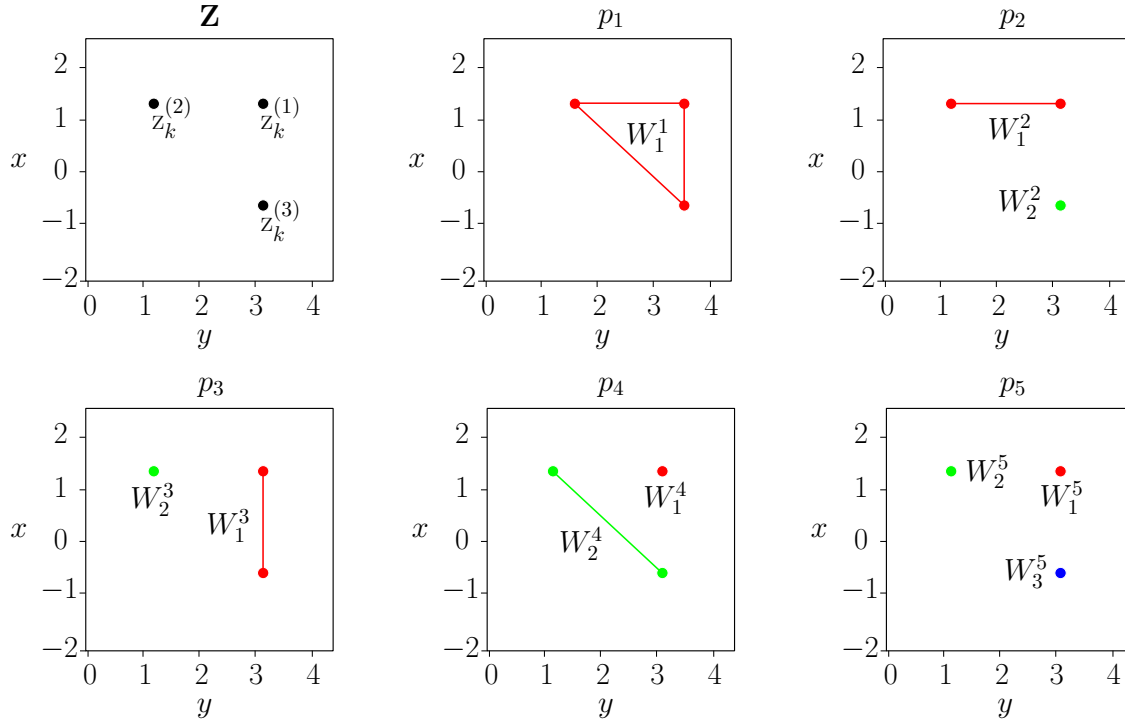
Figure 4.14: Partition principle using three arbitrarily ordered measurements $z_k^{(1)}$, $z_k^{(2)}$, and $z_k^{(3)}$ at time k that are partitioned in five possible ways, where p_j refers to the jth partition and W_i^j refers to the ith cell at partition j. The procedure comprises partitions with various numbers of cells, e.g., $p_1 = \{|W_i^1| = 1\}$, $p_2 = \{|W_i^2| = 2\}$, $p_5 = \{|W_i^5| = 3\}$ [43].

a rectangle are used to model the objects' extensions. Additionally, the objects' surfaces can be equipped with a set of reflection points generating the detections with certain probabilities. However, this requires data association between the target detections and the reflection points [43]. Another approach is to use an inhomogeneous Poisson Point Process (PPP) that models the number of detections and clutter detections as Poisson distributed around the target. By that, the computational demanding associations are avoided. However, the tracking performance may suffer from the high sensitivity to the state dependent Poisson rate [41].

The use of Random finite set (RFS) methods relaxes the necessity for solving the described data association task [73]. RFS based tracking approaches model both the cardinality of objects that are present in the surveillance area and the set of measurements as random variables. The update step in RFS is based on the multi-object likelihood function, which avoids explicit data association by averaging over all possible association hypotheses. Computationally tractable RFS-based implementations include the probability hypothesis density (PHD) [72] filter and the cardinalized probability hypothesis density (CPHD) filter [71].

Typically, RFS-based approaches assume time-varying numbers of present object detections, i.e., the appearance and disappearance of objects in the surveillance area. However, for the

present case where we are striving for the tracking of two objects, i.e., the human legs, we refrain from using complex procedures that take into account the mentioned data handling task. Moreover, we simplify the overall measurement modeling by transferring the present MEOT problem to the assumption of an MPOT problem. It should be explicitly stated that the assumption of point targets for the present tracking application corresponds to a considerable simplification of the measurement modeling. By doing so, computationally attractive MPOT methods, such as the JPDA filter, become applicable. However, it should be noted that generally we would expect an improvement of the tracking performance by using one of the mentioned physical modeling techniques or an sophisticated RFS approach. Pursuing the simplifying point object assumption, partitioning and cell association procedures become needless. Instead, we calculate joint association probabilities, i.e., the probability that the object is the origin of the measurement considering each measurement-object-pair. By doing so, a data association variable, θ, is defined as the association event that measurement j originated from object i with state $\mathbf{x}_k^i$ at time k according to [11]

$$\theta_k^i = \begin{cases} j & \text{if object } i \text{ is associated to measurement } j \\ 0 & \text{if object } i \text{ is undetected.} \end{cases} \tag{4.39}$$

The association for n objects is defined by the association vector $\boldsymbol{\theta}_k = \left[\theta_k^1, \theta_k^2, \ldots, \theta_k^i, \ldots, \theta_k^n\right]$ that consists of the associations for each individual object. Specifying the assumed multiple-point object assumption for $\boldsymbol{\theta}_k \in \boldsymbol{\Theta}_k$, where $\boldsymbol{\Theta}_k$ is the set of all valid associations at time k, the following must hold:

1) Each object must be either detected or miss-detected

$$\theta_k^i \in \{0, \ldots, m_k\}, \forall\, i \in \{1, \ldots, n_k\}.$$

2) Any pair of detected objects cannot be associated to the same measurement

$$\forall\, i, i' \in \{1, \ldots, n\},\, i \neq i',\, \text{if } \theta_k^i \neq 0,\, \theta_k^{i'} \neq 0 \Rightarrow \theta_k^i \neq \theta_k^{i'}.$$

The conditions ensure that the number of associated measurements is at most n in each association event $\boldsymbol{\theta}_k \in \boldsymbol{\Theta}_k$. Given measurements $\mathbf{Z}_k$ and association events $\boldsymbol{\theta}_k$ we can derive

- The association to object i

$$\mathbf{o}_k^i = \begin{cases} \mathbf{z}_k^{\theta_k^i} & \text{if } \theta_k^i \neq 0 \\ \emptyset & \text{if } \theta_k^i = 0. \end{cases}$$

- Which $\mathbf{z}_k^j$ are assigned to the clutter detections $\mathbf{C}_k$

$$j \in \{1, \ldots, m_k\}\ :\ \nexists i \in \{1, \ldots, n_k\},\, \theta_k^i = j.$$

- The number of actual object detections m_k^o, and the number of clutter detections m_k^c

in event $\boldsymbol{\theta}_k$

$$m_k^o = \sum_{i \in \{1,\dots,n_k\}:\theta_k^i \neq 0} 1 \quad \text{and} \quad m_k^c = m_k - m_k^o.$$

The number of ways to associate m_k measurements to n_k objects can be given by

$$N_{\boldsymbol{\theta}_k}(m_k, n_k) = \sum_{m_k^o=0}^{\min(m_k, n_k)} \binom{n_k}{m_k^o} \binom{m_k}{m_k^o} m_k^o!. \tag{4.40}$$

A prevalent method to further reduce the data association events is the use of gating [9], [10], [11]. Gating reduces feasible measurement-to-object associations by comparing the individual measurements to the objects' tracking predictions and identifies unlikely associations. The validation gate acts as an association threshold, where each measurement-to-object association needs to comply with a metric of "acceptance" or is rejected otherwise. For extended objects, the validation gates must consider several object attributes, such as its position, kinematic, and state uncertainties. For linear Gaussian systems, the hyper-ellipsoid of probability concentration is an optimal choice for the validation gate and can be found in numerous applications in literature [9], [10], [66], [111].

Suppose we have a linear observation model with additive noise for an MOT application [7]

$$\mathbf{z}_k = \mathbf{H}_k \mathbf{x}_k + \mathbf{e}_k, \tag{4.41}$$

with Gaussian state density $p(\mathbf{x}_k^i | \mathbf{Z}_{k-1}) = \mathcal{N}(\mathbf{x}_k^i; \hat{\mathbf{x}}_{k|k-1}^i, \mathbf{P}_{k|k-1}^i)$ and Gaussian zero-mean sensor noise $p(\mathbf{e}_k | \mathbf{x}_k) = \mathcal{N}(\mathbf{e}_k; 0, \mathbf{R}_k)$ and time index k. The ellipsoidal gating distance, i.e., the squared Mahalanobis distance[1], from measurement $\mathbf{z}_k^j \in \mathbf{Z}_k$ to the predicted object detection i then reads

$$d_{i,j}^2 = \underbrace{\left(\mathbf{z}_k^j - \hat{\mathbf{z}}_{k|k-1}^i\right)^\top}_{\varepsilon_k^{i,j}} \left(\mathbf{S}_k^i\right)^{-1} \underbrace{\left(\mathbf{z}_k^j - \hat{\mathbf{z}}_{k|k-1}^i\right)}_{\varepsilon_k^{i,j}}, \tag{4.42}$$

where $\varepsilon_k^{i,j}$ is the innovation of detection $\mathbf{z}_k^j$ and predicted object detection $\hat{\mathbf{z}}_{k|k-1}^i = \mathbf{H}_k \mathbf{x}_{k|k-1}^i$, and $\mathbf{S}_k^i$ is the innovation covariance given in (4.20). The Gaussian gating approach assumes that the measurements are Gaussian distributed around their emanating target predictions according to $p(\mathbf{z}_k | \mathbf{x}_k^i) = \mathcal{N}(\mathbf{z}_k; \mathbf{H}_k \hat{\mathbf{x}}_{k|k-1}^i, \mathbf{S}_k^i)$. Measurments are considered to be valid if

$$d_{i,j}^2 \leq \gamma, \tag{4.43}$$

where γ is the gating threshold. If γ becomes small, then the probability that the object measurement is located outside the gate increases. Generally, we strive for a sufficiently

---

1   Let $x = (x_1, x_2, \dots, x_N)^T$ be an observation vector and let $\mu = (\mu_1, \mu_2, \dots, \mu_N)^T$ be a set of observations with mean values, and let S be a covariance matrix, then the Mahalanobis distance can be expressed as: $d_M(x) = \sqrt{(x-\mu)^T S^{-1}(x-\mu)}$. [9]

small probability that the object measurement is outside the gate while trying to keep the data association computation at a moderate level. In order to do so, we introduce a gate probability parameter that specifies the probability that the validation gate contains the true measurement, which is given by

$$P_{\mathrm{G}} = \mathrm{Pr}\left[d_{i,j}^2 < \gamma\right]. \tag{4.44}$$

One can show that the ellipsoidal gating distance follows

$$d_{i,j}^2 \sim \chi^2(n_{\mathbf{z}}), \tag{4.45}$$

where $\chi^2(n_{\mathbf{z}})$ is the chi-squared distribution with $n_{\mathbf{z}}$ degrees of freedom [11]. A common strategy to obtain γ is to set a desired value for P_{G}, e.g., 99.5%, and subsequently use the cumulative distribution of $\chi^2(n_{\mathbf{z}})$. The square roots of the eigenvalues of $\gamma \mathbf{S}_k^i$ denote the semiaxes of the ellipsoid. If the measurement $\mathbf{z}_k^j$ falls outside the gate of $\hat{\mathbf{z}}_{k|k-1}^i$, it is disregarded as a valid data association in the further processing. By doing so, the number of local hypotheses in the posterior n_k object density is significantly reduced.

In nonlinear systems, the ellipsoidal validation gate may reject a higher number of correct associations and be statistically unprecise. There are approaches in the literature that provide nonlinear extensions for the linear gating concept, such as incorporating generic Monte Carlo solutions [7]. However, as can be seen in this work's further course, the use of ellipsoidal gates on the nonlinear tracking methods achieves acceptable results and is applied in the following MOT procedure.

4.3.7 Joint probabilistic data association

In the previous subsection, we have seen that data association can quickly be overwhelming, and methods for alleviating complexity, such as gating, are needed. Instead of computing cell associations, association probabilities are calculated that reveal the probability that the measurement is originated from an object. In probabilistic data association (PDA), posterior association probabilities are determined for each valid measurement candidate lying within a validation region revealing the probability that the measurement is originated from the object or from clutter. This Bayesian information is further incorporated as a weighted sum of innovations in the update step in a tracking filter, such as the PDA filter for a single target in clutter [10], [11].

The basic PDA assumption is that each object is strictly isolated from others without any occurring interference between them. In the automotive context, the majority of real-world scenarios include multiple interfering sources in very high-clutter environments. Hence, measurements from several nearby targets give rise to persistent derogation, which complicates the association task.

The general expression for the normalized n object posterior density at time k can be given by [9], [13]

$$p_{k|k}\left(\mathbf{X}_k\right) = \sum_{\boldsymbol{\theta}_{1:k} \in \boldsymbol{\Theta}_{1:k}} \mathbf{w}_{k|k}^{\boldsymbol{\theta}_{1:k}} p_{k|k}^{\boldsymbol{\theta}_{1:k}}\left(\mathbf{X}_k\right) = \sum_{\boldsymbol{\theta}_1 \in \boldsymbol{\Theta}_1} \cdots \sum_{\boldsymbol{\theta}_k \in \boldsymbol{\Theta}_k} \mathrm{Pr}\left[\boldsymbol{\theta}_{1:k}|\mathbf{Z}_{1:k}\right] p\left(\mathbf{X}|\boldsymbol{\theta}_{1:k}, \mathbf{Z}_{1:k}\right). \tag{4.46}$$

As can be seen from (4.46), the posterior at time k represents a sum over data association sequences from time $1 : k$ where each prior mixture component is conditioned on the corresponding data association sequence where the normalized weights

$$\mathbf{w}_{k|k}^{\boldsymbol{\theta}_{1:k}} = \frac{\tilde{\mathbf{w}}_{k|k}^{\boldsymbol{\theta}_{1:k}}}{\sum_{\boldsymbol{\theta}'_{1:k}} \tilde{\mathbf{w}}_{k|k}^{\boldsymbol{\theta}'_{1:k}}} = \frac{\prod_{i=1}^{n} \prod_{t=1}^{k} \tilde{\mathbf{w}}^{\theta_t^i|\boldsymbol{\theta}_{1:t-1}^i}}{\sum_{\boldsymbol{\theta}'_1} \cdots \sum_{\boldsymbol{\theta}'_k} \prod_{i=1}^{n} \prod_{t=1}^{k} \tilde{\mathbf{w}}^{\theta_t^{i'}|\boldsymbol{\theta}_{1:t-1}^{i'}}}, \tag{4.47}$$

with

$$\tilde{\mathbf{w}}^{\boldsymbol{\theta}^i} \begin{cases} \int \left(1 - P^{\mathrm{D}}(\mathbf{x}^i)\right) p^i(\mathbf{x}^i) \mathrm{d}\mathbf{x}^i & \text{if } \boldsymbol{\theta}^i = 0 \\ \int \frac{P^{\mathrm{D}}(\mathbf{x}^i) p(\mathbf{z}^{\boldsymbol{\theta}^i}|\mathbf{x}^i)}{\lambda_c(\mathbf{z}^{\boldsymbol{\theta}^i})} p^i(\mathbf{x}^i) \mathrm{d}\mathbf{x}^i & \text{if } \boldsymbol{\theta}^i \neq 0, \end{cases} \tag{4.48}$$

denote the probabilities of the data associations $\Pr[\boldsymbol{\theta}_{1:k}|\mathbf{Z}_{1:k}]$ assuming a constant detection probability P^{D} for the correct measurement, and a clutter Poisson intensity function $\lambda_c(\mathbf{z}^{\boldsymbol{\theta}^i})$ for the detections associated to clutter.

Due to the rapidly increasing number of mixture components, the exact posterior is intractable, and approximations are necessary. While MHT and RFS usually assume that the number of targets is a time-varying discrete random variable by means of data handling, the JPDA filter deals with an a priori known and constant number of objects in clutter. By doing so, the handling of the data association hypotheses becomes somewhat easier. Initially developed by *Bar-Shalom* and *Fortmann*, the JPDA filter manifests a suboptimal approach for MOT in high-clutter environments. The conceptual idea in JPDA is to approximate the posterior in (4.46) as a Gaussian density by combining the different hypotheses in the JPDA posterior density into a single one in order to get the simplicity of a single-hypothesis algorithm while maintaining information from all hypotheses. This is achieved by determining an association between the tracks and a set of weighted innovations utilized as an approximation of the optimal association in the update step [40]. As in PDA, the JPDA filter computes the probabilities of association in a non-back scan approach, i.e., only the measurements at the current time are considered. However, contrary to the PDA algorithm in the single object case, where the association probabilities are computed separately, the JPDA algorithm computes them jointly across all present objects and clutter.

In JPDA, the exact posterior in (4.46) is approximated by the JPDA density [11], [13]

$$p_{k|k}^{\mathrm{JPDA}}(\mathbf{X}_k) = p_{k|k}^{\boldsymbol{\beta}_{1:k}}(\mathbf{X}_k), \tag{4.49}$$

paramterized by the individual object densities

$$p_{k|k}^{i,\boldsymbol{\beta}_{1:k}}(\mathbf{x}_k^i), \quad \text{for } i = 1, 2, \ldots, n_k \tag{4.50}$$

where $\boldsymbol{\beta}_{1:k}$ denotes the sequence of marginal association probabilities computed in each time step according to

$$\boldsymbol{\beta}_1, \quad \boldsymbol{\beta}_2|\boldsymbol{\beta}_1, \quad \ldots \quad \boldsymbol{\beta}_k|\boldsymbol{\beta}_{1:k-1}.$$

The marginal probabilities are computed jointly for all objects as the sum of the weights $\mathbf{w}^{\theta_k}$ for which $\theta_k^{i,j} = j$ based on the set of valid association events, Θ_k, given by (see Appendix C)

$$\beta_k^{i,j} = \mathrm{Pr}\left(\theta_k^i = j | \mathbf{Z}_{1:k}\right) = \sum_{\theta_k \in \Theta_k : \theta_k^i = j} \mathbf{w}^{\theta_k} \propto \sum_{\theta_k \in \Theta_k : \theta_k^i = j} \tilde{\mathbf{w}}^{\theta_k}. \tag{4.51}$$

Consequently, the marginal probability that object i is not detected at time k reads

$$\beta_k^{i,0} = \mathrm{Pr}\left(\theta_k^i = 0 | \mathbf{Z}_{1:k}\right) = 1 - \sum_j \beta_k^{i,j} \sum_{\theta_k \in \Theta_k : \theta_k^i = 0} \mathbf{w}^{\theta_k} \propto \sum_{\theta_k \in \Theta_k : \theta_k^i = 0} \tilde{\mathbf{w}}^{\theta_k}. \tag{4.52}$$

The marginal posterior for object i is determined using the computed marginal association probabilities according to

$$p_{k|k}^{i,\beta_{1:k}}\left(\mathbf{x}_k^i\right) = \beta_k^{i,0} p_{k|k-1}^{i,\beta_{1:k-1}}\left(\mathbf{x}_k^i\right) + \sum_{j=1}^{m_k} \beta_k^{i,j} p_{k|k-1}^{i,\beta_{1:k-1},j}\left(\mathbf{x}_k^i\right), \tag{4.53}$$

where $p_{k|k-1}^{i,\beta_{1:k-1},j}\left(\mathbf{x}_k^i\right)$ is the posterior density that results from updating the prior of object i

$$p_{k|k-1}^{i,\beta_{1:k-1}}\left(\mathbf{x}_k^i\right) = \int p\left(\mathbf{x}_k^i | \mathbf{x}_{k-1}^i\right) p_{k-1|k-1}^{i,\beta_{1:k-1}}\left(\mathbf{x}_{k-1}^i\right) \mathrm{d}\mathbf{x}_{k-1}^i, \tag{4.54}$$

with measurement j.

Next, the marginal posterior for each object is merged into a single density. This is achieved by using moment matching, i.e., matching the mean and covariance with respect to the minimization of the Kullback-Leibler divergence [11]. The resulting merged Gaussian density of object i then reads

$$p_{k|k}^{i,\mathrm{JPDA}}\left(\mathbf{x}_k^i\right) = \mathcal{M}\left(p_{k|k}^{i,\beta_{1:k}}\left(\mathbf{x}_k^i\right)\right) = \mathcal{N}\left(\mathbf{x}_k^i; \bar{\mathbf{x}}_{k|k}^{i,\mathrm{JPDA}}, \mathbf{P}_{k|k}^{i,\mathrm{JPDA}}\right) \tag{4.55}$$

where $\mathcal{M}$ denotes the merging function, and $\bar{\mathbf{x}}_{k|k}^{i,\mathrm{JPDA}}$ and $\mathbf{P}_{k|k}^{i\,\mathrm{JPDA}}$ are the mean and covariance of the merged density, respectively. Figure 4.15 exemplary visualizes the explained concept of hypotheses merging in the JPDA algorithm using a scenario with two propagating objects and applied gating.
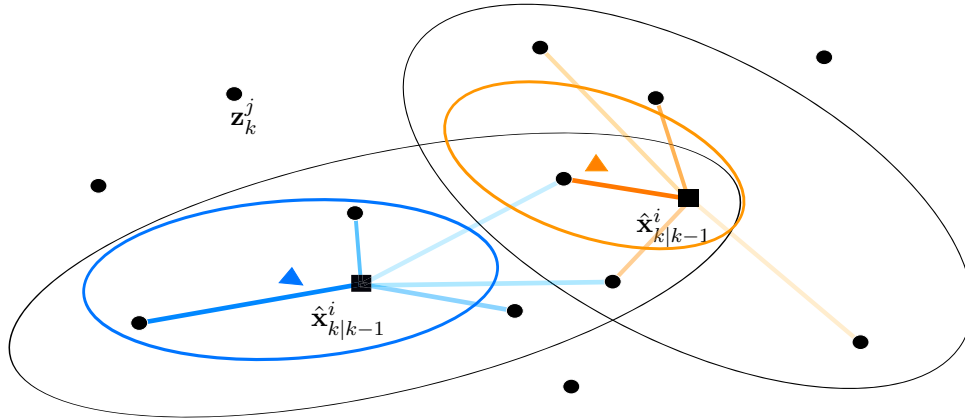
(a) Two propagating objects at time k with m_k surrounding measurements $\mathbf{z}_k^j$ for $j = 1, \ldots, m_k$. The black ellipses describe the validation regions (gates) centered at the predicted object detections $\hat{\mathbf{x}}_{k|k-1}^1$ and $\hat{\mathbf{x}}_{k|k-1}^2$, respectively. Measurements outside the validation regions are disregarded. Measurements inside a validation region could originate from the corresponding object or clutter. Measurements that lie at the intersection of the ellipses could originate from one of the objects or clutter. The color intensities of the measurement-to-track hypotheses refer to the marginal association probabilities $\beta_k^{i,j}$, i.e., the probability that object $\hat{\mathbf{x}}_{k|k-1}^i$ is associated to measurement $\mathbf{z}_k^j$ after applying gating. The blue and orange ellipses, respectively, exemplary represent the merged mean (triangles) and covariance of the updated marginal posterior in the sense of JPDA.
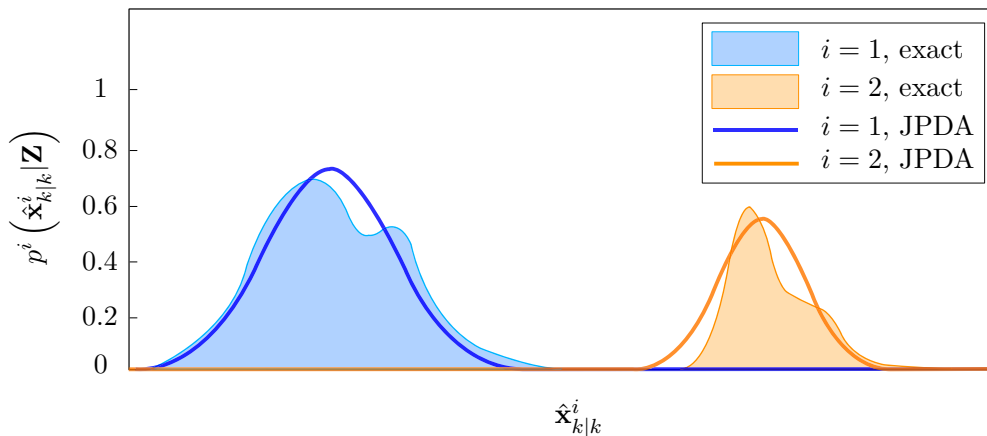
(b) Marginal posterior density functions of the two objects using JPDA. The marginal posterios are multi-modal distributions due to the incorporation of weighted marginal association hypotheses. The idea in JPDA is to approximate the marginal posteriors with Gaussians by merging all hypotheses into a single one.

Figure 4.15: Exemplary illustration of the hypotheses merging and density approximation in the JPDA algorithm.

4.4 Kinematic-based pedestrian tracking

The results of radar-based pedestrian detection presented so far in this thesis illustrate the tremendous potential of μD information concerning a targeted function for pedestrians' enhanced state estimation in road traffic. In contrast to previous approaches, which consider the pedestrian as a holistic object, this work pursues the goal of a segmented limb specific tracking based on high-resolution radar point clouds. The preceding subchapters lay the mathematical foundations for developing a tracking algorithm in multiple-object scenarios. The following approach's core idea consists of designing a motion model that purposefully maps the complex kinematic behavior of a body segment in several dimensions and enables its tracking through statistical filtering. More precisely, we approximate the feets' spatial locomotion using a kinematic model that variably adapts to the average walking velocity. The reason for tracking the foot motion instead of other limbs lies in its overall significance. The interaction of the stride length and the stride frequency regulates the average walking velocity of a pedestrian [59]. Consequently, in disparity to other limbs, the foot behavior determines the pedestrians' overall macroscopic motion and indicates initiating movement changes through the stride parameters and foot positioning. Moreover, as seen in this work, the foot motion reveals the largest Doppler amplitudes and the highest spatial extension during the walking phase, making the foot movement best distinguishable from other scatter clusters. Particularly in critical situations when the distance between the sensor carrying vehicle and an approaching pedestrian becomes small, the major part of measurable reflections emanate from the lower body. Besides the low installation heights of SRRs, this is mainly due to their antenna characteristics. While their azimuthal beamwidths are typically enlarged, they show relatively narrow elevational beamwidths leading to a concentrated perception of targets with similar height.

The usage of an elevation-resolving radar antenna aperture allows the perpendicular segmentation of the human body and, thus, the extraction of the leg movement's scattered points. The reflections are subsequently assigned to the respective leg using a suitable data association procedure yielding the position, velocity, and kinematic knowledge of the left and right foot, respectively. Contrary to what is known from the literature, we will demonstrate that suboptimal first-order approximations, such as the EKF, and the integration of the μD detections into the measurement tracking vector are well suited to describe nonlinear kinematics. To the author's best knowledge, this is the first approach to use human kinematics to track microscopic motion based on μD data of a high-resolution automotive short-range radar to reveal behavior indications during single measurement cycles. The approach provides a computationally efficient system that is suitable for the embedded use in automotive series sensors.

4.4.1 Kinematic modeling

This subsection draws on the human kinematics given in Subsection 4.1.1. As mentioned, the body's motion possesses a periodically repeating locomotion in the gait cycle, which consists of a stance and swing phase where the legs show acceleration and deceleration motion patterns. This work's motion model development for the approximation of the foot movement is established on the *global human walking model* of *Boulic*, *Thalmann*,

and *Thalmann* [21] which was adapted and implemented by *Chen* [25]. They presented a mathematical parameterization emanated from empirical biomechanical data. The proposed model produces spatial, kinematic, and joint parameters of a person's average walking behavior for any point in time. By doing so, the model represents a universally applicable motion model that appears free of any personalized movement characteristics. The global human walking model is designed in analogy to the Denavit-Hartenberg (DH) convention. The DH convention is a standard kinematic expression that defines the transfer of spatial coordinate systems within kinematic chains [47]. The DH convention is widely employed in robotics and simplifies the forward kinematics of spatial joint parameters. Each leg of the human walking model consists of three joints connected by rigid links. These possess the hip, knee, and ankle joints whose DH connections are systematically pictured in Figure 4.16. Each joint coordinate system is designed based on the following beliefs:

1) The $\breve{z}_i$ axis points along the axis of rotation of the ith joint.

2) The $\breve{x}_i$ axis goes along the common normal of the two joint axes $\breve{z}_{i-1}$ and $\breve{z}_i$.

3) The $\breve{y}_i$ axis completes the right handed coordinate system.

The DH parameters are emanated from the common normal $\breve{a}_i$ between successive $\breve{z}$ axes. The resulting $\breve{x}_i$ axis is collinear with the common normal. With these joint axes, four parameters specify the joint-to-joint transformation: $\breve{r}$, $\breve{\psi}$, $\breve{a}$, $\breve{\alpha}$. Here, $\breve{\psi}$ is the angle about the previous $\breve{z}$ to align its $\breve{x}$ with the new source. $\breve{a}_i$ is the length along the rotated $\breve{x}_{i-1}$ axis. Finally, $\breve{\alpha}$ spins around the new $\breve{x}$ axis to set $\breve{z}$ in its expected direction. Parallel $\breve{z}$
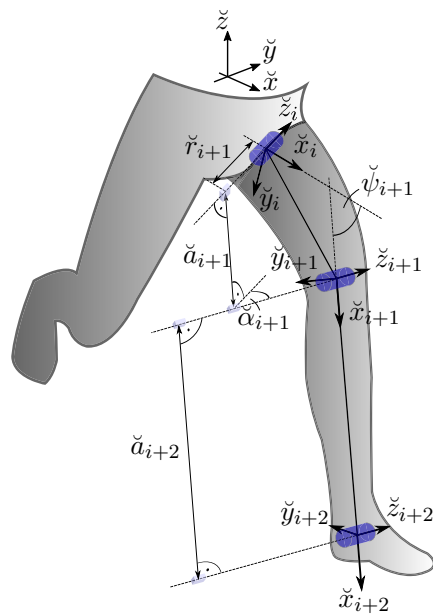
Figure 4.16: A systematic description of the motion relationships of a human leg (translations and rotations in the forward movement direction) between adjacent joints connected by rigid links in accordance with the DH convention.

axes formulate a special case. Since parallel $\check{z}$ axes have an infinite number of common normals, one can select any suitable $\check{r}$ value. The other parameters are determined as before. $\check{\alpha}$ is zero in this case.

Accordingly, a human body is modeled as joints and corresponding rigid links providing six degrees of freedom per element, i.e., three Cartesian position coordinates in 3-D space and three Euler angles of rotation. Accurately, 12 trajectories, each relying on the average walking velocity, represent the motion of one gait cycle for the complete body [21]. Six trajectories are expressed by sinusoidal functions (one of them as piecewise function) comprising the translational and rotational movement of the hip, and six trajectories are defined by cubic spline functions based on extremities' control points describing the fexing of joints.

Given an average walking velocity $\bar{v}$, the relative length of a gait cycle can be empirically formulated by [21], [59]

$$r_c = 1.346 \cdot \sqrt{\bar{v}}. \tag{4.56}$$

From that, the gait can be divided in the time duration of single-leg support t_s, i.e. the swing phase, and in the time duration of double support t_{ds}, i.e., the stance phase according to

$$t_s = 0.752t_c - 0.143 \tag{4.57}$$

$$t_{ds} = 0.252t_c - 0.143, \tag{4.58}$$

where $t_c = r_c/\bar{v}$ denotes the cycle duration. Assuming a body-fixed coordinate system centered at the spine's origin, the translational motion trajectories are:

1) Vertical translation: the vertical offset from the spine's origin

$$tr_{\text{vertical}} = -a_v + a_v \sin\left(2\pi(2t_r - 0.35)\right), \tag{4.59}$$

   where $a_v = 0.015\bar{v}$, and t_r is the normalized relative time.

2) Lateral translation: the lateral oscillation from the spine's origin

$$tr_{\text{lateral}} = -a_l \sin\left(2\pi(t_r - 0.1)\right), \tag{4.60}$$

   where

$$a_l = \begin{cases} -0.128\bar{v}^2 + 0.128\bar{v} & \text{if } \bar{v} < 0.5 \\ -0.032 & \text{if } \bar{v} > 0.5. \end{cases}$$

3) Forward/backward translation: the acceleration and deceleration processes occurring while beginning and finishing a step

$$tr_{\text{F/B}} = -a_{\text{F/B}} \sin\left(2\pi(2t_r + 2\varphi_{\text{F/B}})\right), \tag{4.61}$$

   where

$$a_{\text{F/B}} = \begin{cases} -0.084\bar{v}^2 + 0.084\bar{v} & \text{if } \bar{v} < 0.5 \\ -0.021 & \text{if } \bar{v} > 0.5, \end{cases}$$

and $\varphi_{\mathrm{F/B}} = 0.625 - t_{\mathrm{s}}$.

The trajectories describing the hip's rotational movements are:

1) Forward/backward rotation: the back's flexing movement relative to the hip beafore each step

$$ro_{\mathrm{F/B}} = -ar_{\mathrm{F/B}} + ar_{\mathrm{F/B}} \sin\left(2\pi(2t_{\mathrm{r}} - 0.1)\right), \qquad (4.62)$$

where

$$ar_{\mathrm{F/B}} = \begin{cases} -8\bar{\boldsymbol{v}}^2 + 8\bar{\boldsymbol{v}} & \text{if } \bar{\boldsymbol{v}} < 0.5 \\ 2 & \text{if } \bar{\boldsymbol{v}} > 0.5. \end{cases}$$

2) Left/right rotation: the flexing movement that rotates the hip on the side of the swinging leg. The piecewise function is given by

$$ro_{\mathrm{L/R}} = \begin{cases} -ar_{\mathrm{L/R}} + ar_{\mathrm{L/R}} \cos\left(2\pi(10t_{\mathrm{r}}/3)\right) & \text{for } 0 \le t_{\mathrm{r}} < 0.5 \\ -ar_{\mathrm{L/R}} - ar_{\mathrm{L/R}} \cos\left(2\pi(10(t_{\mathrm{r}} - 0.15)/7)\right) & \text{for } 0.15 \le t_{\mathrm{r}} < 0.5 \\ -ar_{\mathrm{L/R}} - ar_{\mathrm{L/R}} \cos\left(2\pi(10(t_{\mathrm{r}} - 0.5)/3)\right) & \text{for } 0.5 \le t_{\mathrm{r}} < 0.65 \\ -ar_{\mathrm{L/R}} + ar_{\mathrm{L/R}} \cos\left(2\pi(10(t_{\mathrm{r}} - 0.65)/7)\right) & \text{for } 0.65 \le t_{\mathrm{r}} < 1, \end{cases}$$

where $ar_{\mathrm{L/R}} = 1.66\bar{\boldsymbol{v}}$.

3) Torsion rotation: the hip's rotation relative to the spine

$$ro_{\mathrm{Torsion}} = -ar_{\mathrm{Torsion}} \cos\left(2\pi t_{\mathrm{r}}\right), \qquad (4.63)$$

where $ar_{\mathrm{Torsion}} = 4\bar{\boldsymbol{v}}$.

Three of the six flexing trajectories characterize the lower body's motion. They are constructed by cubic spline functions fitted to the joints' control points [21]. Figure 4.17 depicts the lower body motion trajectories for one gait cycle, including the given translational, rotational, and flexing functions of the hip, knee, and ankle, respectively. Those trajectories are subsequently utilized to compute the location of reference points on the human body in 3-D space. For the lower body, this is achieved by recursively employing the Euler angles rotation matrix to every joint up to the hip at each frame. By that, each flexing motion directly influences the previously calculated flexing motion [21], [25]. The result reveals the human lower body's integrated spatial movement for a non-moving model (with respect to the environment) on which the following kinematic-based leg tracking approach is based.

4.4.2 Tracking motion model

Motion models express the state transition probability $p\left(\mathbf{x}_k | \mathbf{x}_{k-1}\right)$ in the prediction step and hence are, in in combination with measurement models, crucial elements concerning the implementation of any Bayesian filtering algorithm. Generally, a reasonable probabilistic model's goal may appear to accurately model the distinct kinds of uncertainty in the system's actuation and perception.
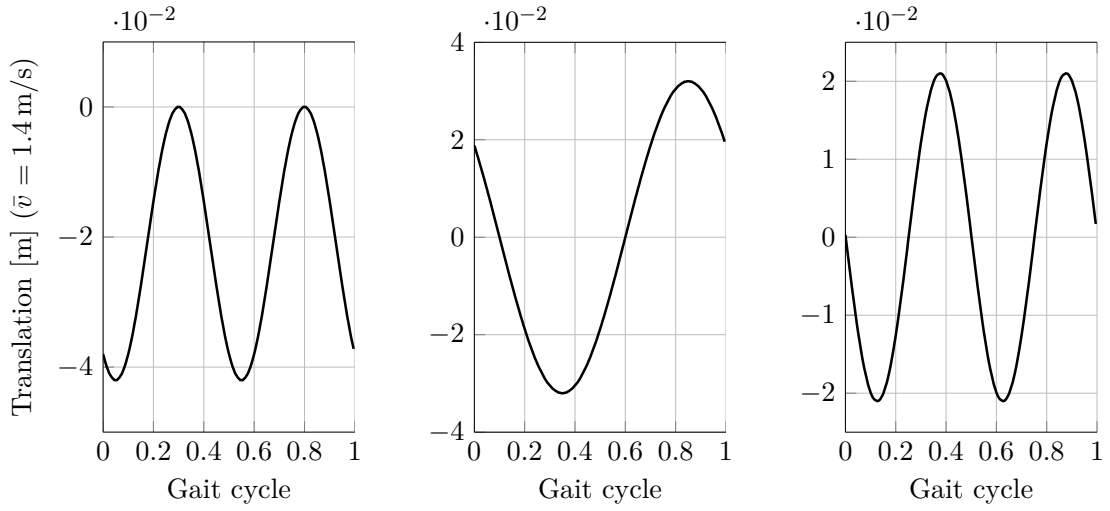
(a) Vertical translation offset from the spine's origin.

(b) Lateral translation offset from the spine's origin.

(c) Forward/backward translation offset from the spine's origin.
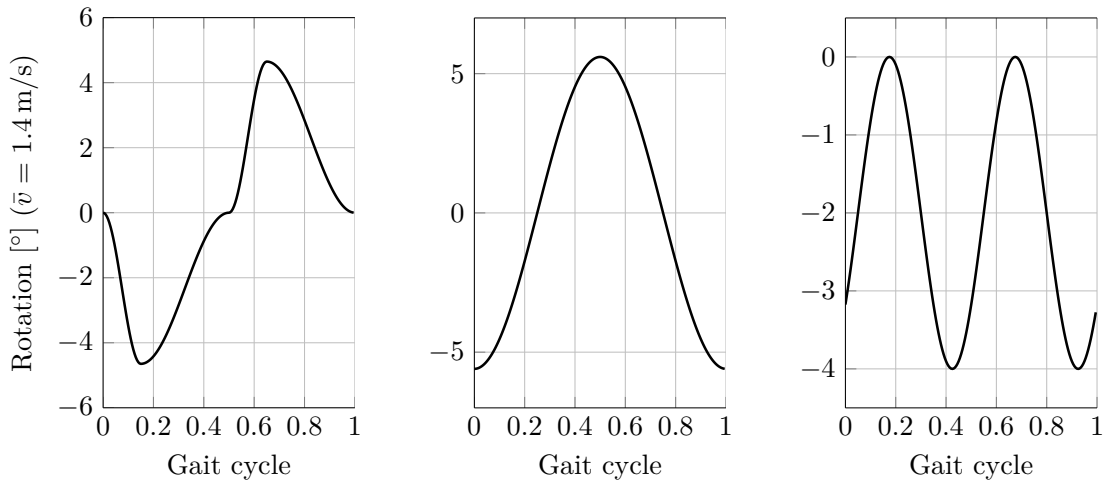
(d) Rotation left/right.

(e) Rotation forward/backward.

(f) Torsion rotation.
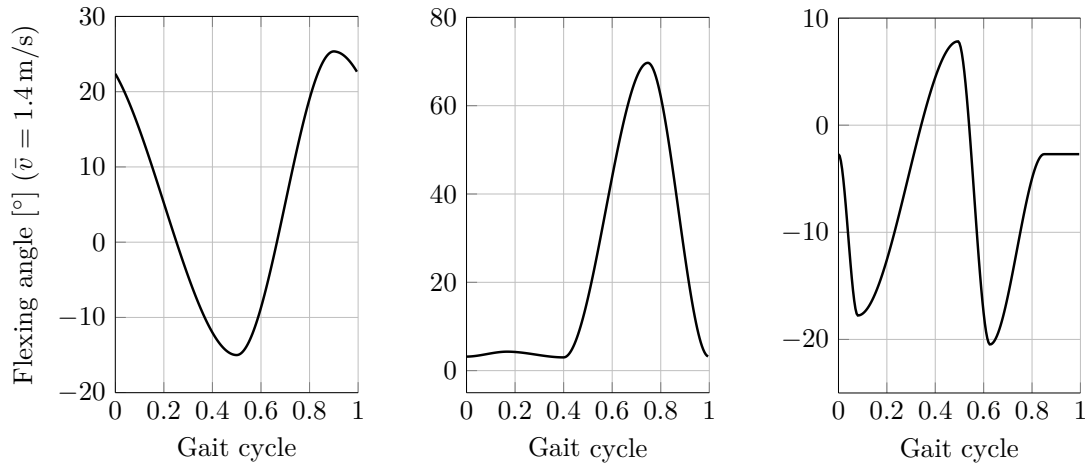
(g) Flexing at the hip.

(h) Flexing at the knee.

(i) Flexing at the ankle.

Figure 4.17: Lower body motion trajectories of the human walking model [25].

The motion model's development is carried out in accordance to the leg's motion behavior, i.e., the ankle's position displacement within a gait cycle established on the global human walking simulation presented in the previous Subsection 4.4.1. Following up the integrated spatial movement of the lower body, the displacement functions of one ankle in $\breve{x}$ and $\breve{y}$ direction are from now on denoted as $f_{\mathrm{Ankle}}^{\breve{x}}(t)$ and $f_{\mathrm{Ankle}}^{\breve{y}}(t)$, respectively, and illustrated in Figure 4.18. Note that the average walking velocity $\bar{\boldsymbol{v}}$ complies to $1.4\,\mathrm{m/s}$. The right-handed body coordinate system is centered at the spine's origin with directions shown in Figure 4.16. As shown in Figure 4.18, the modeled displacement functions of the ankle describe periodic curves with apparently trigonometric signal components. A convenient way to sufficiently approximate $f_{\mathrm{Ankle}}^{\breve{x}}(t)$ and $f_{\mathrm{Ankle}}^{\breve{y}}(t)$ can be realized by Fourier synthesis. The Fourier series describes a synthesized periodic function consisting of weighted harmonic sinusoids whose general form can be given by

$$F_{N_\kappa}(t) := \frac{a_0}{2} + \sum_{\kappa=1}^{N_\kappa} \left(a_\kappa \cos\left(\kappa t\right) + b_\kappa \sin\left(\kappa t\right)\right) \quad \text{for } a_\kappa, b_\kappa \in \mathbb{R}. \tag{4.64}$$

If we apply the Fourier synthesis process to each of the displacement functions we get the following approximations

$$f_{\mathrm{Ankle}}^{\breve{x}}(t) \approx \underbrace{\frac{a_0^{\breve{x}}}{2} + \sum_{\kappa=1}^{N_\kappa} \left(a_\kappa^{\breve{x}} \cos\left(\kappa t\right) + b_\kappa^{\breve{x}} \sin\left(\kappa t\right)\right)}_{\mathbf{f}_{\mathrm{Ankle}}^{\breve{x}}(t)} \tag{4.65}$$

$$f_{\mathrm{Ankle}}^{\breve{y}}(t) \approx \underbrace{\frac{a_0^{\breve{y}}}{2} + \sum_{\kappa=1}^{N_\kappa} \left(a_\kappa^{\breve{y}} \cos\left(\kappa t\right) + b_\kappa^{\breve{y}} \sin\left(\kappa t\right)\right)}_{\mathbf{f}_{\mathrm{Ankle}}^{\breve{y}}(t)}. \tag{4.66}$$

With three superimposed harmonic terms, $\kappa = 3$, each approximation function possesses seven Fourier coefficients in total. As can be seen in Figure 4.18, the three-term series expansion displays the actual displacement functions appropriately well. More precisely, the approximation $f_{\mathrm{Ankle}}^{\breve{x}}(t)$ depicts both the leg's initial sinusoidal-like extension up to about $0.55\,\mathrm{s}$ and the subsequent almost linear stance phase until the end of the cycle. The same applies for $f_{\mathrm{Ankle}}^{\breve{y}}(t)$ where the overall course is satisfactorily represented. The overall root mean square error (RMSE) between the approximations and the simulated functions created with $\bar{\boldsymbol{v}} = 1.4\,\mathrm{m/s}$ are $0.011\,\mathrm{m}$ in $\breve{x}$, and $0.003\,\mathrm{m}$ in $\breve{y}$ direction, respectively. The fact that the three-term Fourier series already match with high compliance is especially advantageous regarding low computational demand.

Since $f_{\mathrm{Ankle}}^{\breve{x}}(t)$ and $f_{\mathrm{Ankle}}^{\breve{y}}(t)$ depend on the average walking velocity, we strive for a variable motion model that takes kinematic deviations into account. Since the average walking velocity denotes a product of the stride length and the stride frequency, a suitable motion model must consider these parameters in the best possible way. For that, we add additional global parameters to the approximations that raise the model's variability greatly. The first parameter is a global amplitude, $\tilde{a}$, that adjusts the model to modifications in walking speed

and specifically manages the step length in $\breve{x}$ direction. Secondly, a frequency parameter, $\tilde{\omega}$, is introduced to reflect the relationship between walking velocity and stride frequency. Finally, a third parameter, $\tilde{\varphi}$, is applied to control the foot movement's overall phase position in $\breve{x}$ direction. By that, we derive a single motion model that describes the motion of both feet simultaneously by a relative to each other shifted phase.

Since the time vector t is subject to a linear propagation belief depending on the average walking velocity to which the model was initially approximated, an additional term in the form of the sampling interval, $T = t_k - t_{k-1}$, with time index k is incorporated into the phase, which entitles the model to incrementally shift as a function of time. Then, the final discretized motion model propagation reads

$$
\begin{aligned}
\mathbf{f}_k^{\breve{x}} = \tilde{a}_{k-1} \Bigg[& \sum_{\kappa=1}^{N_\kappa} \bigg(a_\kappa^{\breve{x}} \cos\left(\tilde{\omega}_{k-1}\kappa\left(t_k - \tilde{\varphi}_{k-1}\right)\right) \\
& + b_\kappa^{\breve{x}} \sin\left(\tilde{\omega}_{k-1}\kappa\left(t_k - \tilde{\varphi}_{k-1}\right)\right) \bigg) + \frac{a_0^{\breve{x}}}{2} \Bigg]
\end{aligned}
\tag{4.67}
$$

$$
\begin{aligned}
\mathbf{f}_k^{\breve{y}} = \tilde{a}_{k-1} \Bigg[& \sum_{\kappa=1}^{N_\kappa} \bigg(a_\kappa^{\breve{y}} \cos\left(\tilde{\omega}_{k-1}\kappa\left(t_k - \tilde{\varphi}_{k-1}\right)\right) \\
& + b_\kappa^{\breve{y}} \sin\left(\tilde{\omega}_{k-1}\kappa\left(t_k - \tilde{\varphi}_{k-1}\right)\right) \bigg) \Bigg] \pm \frac{a_0^{\breve{y}}}{2}.
\end{aligned}
\tag{4.68}
$$

Note that $\tilde{a}$ has no influence on the vertical displacement factor $\pm a_0^{\breve{y}}/2$ in (4.68). Contrary to (4.67), (4.68) shall display this average displacement in the case of $\bar{\boldsymbol{v}} = 0$ to assure the ankle's initial location relative to the body's vertical axis. The algebraic sign $\pm$ marks whether $\mathbf{f}_k^{\breve{y}}$ is referring to the left or the right leg due to axial symmetry.

For validation purposes, nonlinear least-squares analyses are performed to analyze the model's adaption ability regarding different average walking velocities. The global parameters, $\tilde{\boldsymbol{\gamma}} = (\tilde{a}, \tilde{\omega}, \tilde{\varphi})^T$, of the motion models are selected in such a way that they approximate the simulated displacement functions, $f_{\text{Ankle}}^{\psi,\eta}(t)$, in the least-squares sense, meaning that the sum of residual squares is minimized. The minimization procedure corresponds to (3.6) in Subsection 3.2.1. Here, ψ and η indexing the motion dimension and the selected average walking velocities of the simulated functions, respectively. The parameters that minimize $\chi(\tilde{\boldsymbol{\gamma}})$ are supposed to be the best fit parameters. Table 4.2 displays the RMSE between the approximations, $\mathbf{f}^{\psi,\eta}(t)$, and the functions, $f_{\text{Ankle}}^{\psi,\eta}(t)$, approximated for diverse walking velocities in $\breve{x}$ and $\breve{y}$ dimension. With selected velocities, $\bar{\boldsymbol{v}}$, from 0.6 to 2.6 m/s the average range of pedestrian walking speeds of 1.20 to 1.50 m/s is abundantly covered [67]. With a median RMSE for all velocities of 0.014 m in $\breve{x}$ and 0.009 m in $\breve{y}$ direction, respectively, the suggested approach emphasizes its suitability as a kinematically variable motion model for the leg tracking approach.
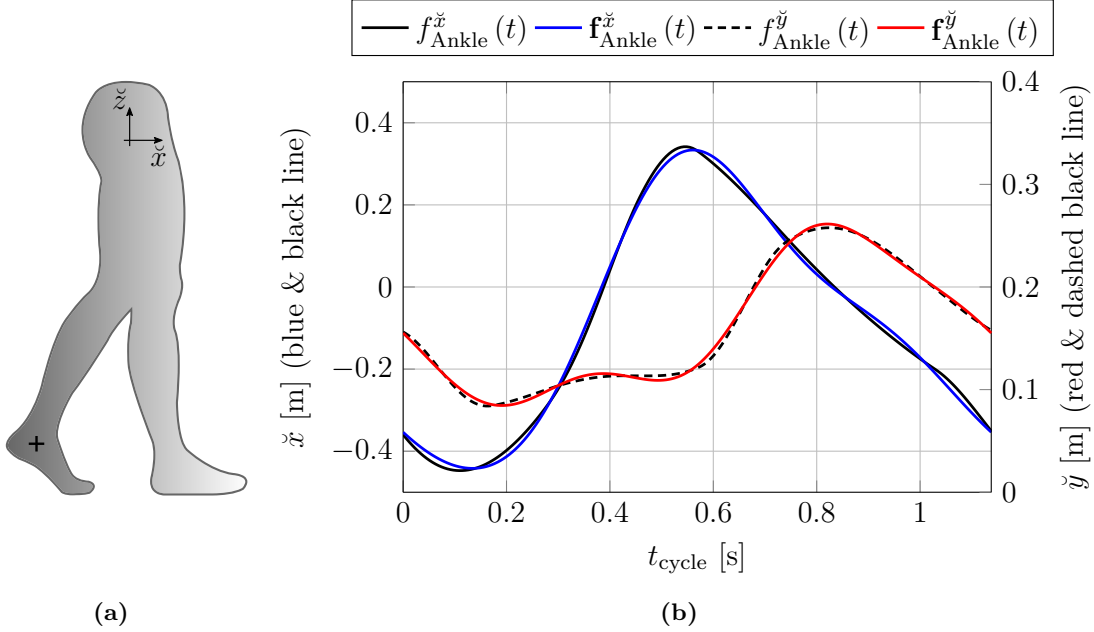
Figure 4.18: Integrated spatial movement displacement of the left ankle. (a) Relative cycle starting position of the left ankle. (b) Movement displacement (——) in $\breve{x}$ and (- - -) in $\breve{y}$ direction of one gait cycle with $\bar{\boldsymbol{v}} = 1.4\,\text{m/s}$ within the body coordinate system with the corresponding three-term Fourier series approximations (——) in $\breve{x}$ and (——) in $\breve{y}$, respectively.

Table 4.2: Results of the motion model evaluation.

$\bar{\boldsymbol{v}}$ in [m/s]	0.6	1.0	1.4	1.8	2.2	2.6
$\text{RMSE}_{\mathbf{f}^{\breve{x}}}$ in [m]	0.012	0.011	0.011	0.017	0.029	0.042
$\text{RMSE}_{\mathbf{f}^{\breve{y}}}$ in [m]	0.021	0.013	0.003	0.014	0.027	0.041

4.4.3 4-D radar point cloud

Given the progressive development of HAD, short-range applications are becoming increasingly crucial for current ADAS research. Especially in confusing urban areas, a detailed environmental perception of the vehicle is necessary. Previous radar systems in the automotive field usually did not provide elevation measuring ability, mainly due to cost and sensor size reasons. Moreover, coming from a classical long-range domain, horizontal angular information outranks vertical information in most radar-based applications [29]. However, next-generation radar sensors have advanced antenna structures that significantly improve environmental detection for active safety systems. Particularly in pedestrians, the elevation dimension allows exploiting the entire μD signature characteristics and lays the foundation for sophisticated limb tracking approaches for advanced pedestrian safety.

Up to this point, only the azimuth angle was considered in all experimental results, which contain angular information. Now, we will extend the dimensionality of the detections by the elevation angle utilizing an elevation-resolving aperture, including horizontally and

diagonally aligned receiving antennas that provide information of range r, Doppler v^D, azimuth angle ϕ, and elevation angle ϵ for each detection as shown in Figure 4.19(a).

The used radar-frontend comprises 16 receiving (RX) antennas with spacing $d = \lambda/2$ subdivided into two arrays of 12 and 5 elements, respectively, whereas both arrays utilize the overlapping antenna. Digital beamforming is performed for both arrays separately according to the principle of DOA given in Subsection 2.2.3. Subsequently, the position (range r and angle Ψ) of the most energetic scatterer is determined. Note that the resulting angular spectrum of the diagonally aligned antenna elements may possess both azimuth and elevation target information. In order to determine the actual height of a target, z_T, the geometry of the antenna structure is exploited. As can be seen from the antenna positions, the two array planes are rotated by 45° to each other, which means that the calculated target positions are located on two planes which are rotated by 45°. As shown in Figure 4.19(b), the positions are $y_1 = r \cdot \sin(\Psi_1)$, and $y_2 = r \cdot \sin(\Psi_2)$. Note that y_1 already corresponds to the target's y-position, y_T, in the radar's coordinate system. z_T is calculated with the help of y_2, where $\sqrt{2}$ results from the rotation by 45°. Thus, y_2 can be regarded as one side of a square whose diagonal equals to $\sqrt{2}y_2$. From the illustration it gets apparent that $z_T = y_1 - \sqrt{2}y_2$. Finally, the target's x-coordinate, x_T, can be calculated using $x_T = r \cdot \cos(\Psi_1)$.

4.4.4 Tracking implementation

In the following, the radar-based pedestrian tracking approach is particularized in detail. The algorithm consists of a nested tracking structure comprising upper body (ub) and lower body (lb) tracking procedures. For this intent, in each frame, the pedestrian's detections resulting from the described signal processing steps are assigned to the upper and lower body, respectively. Detections whose height is above the sensor height, are assigned to the upper body's detections $\mathbf{Z}^{ub}$, and detections whose height is below the sensor height are assigned to the lower body's detections $\mathbf{Z}^{lb}$. An architectural description of the pedestrian
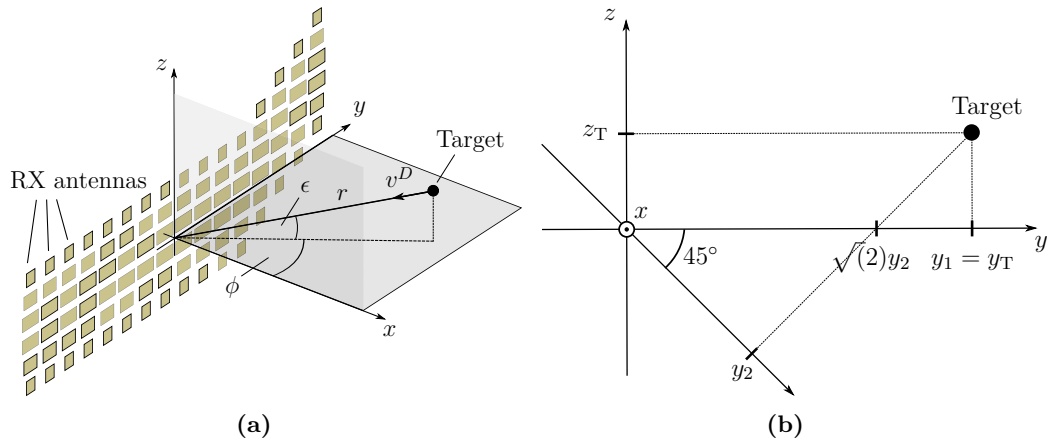
Figure 4.19: Physical representation of the measuring principles of range r, Doppler v^D, azimuth angle ϕ, and elevation angle ϵ. (a) Uniform linear antenna array including horizontally and diagonally aligned RX antennas. (b) Principle of the height determination.

tracking system is given in Figure 4.20. As can be seen, the entire upper body's radar detections $\mathbf{Z}^{\mathrm{ub}}$ including arm, torso, and head reflections are forwarded to the upper body Kalman filter (Upper body KF). In contrast, the lower body detections $\mathbf{Z}^{\mathrm{lb}}$ are available for the data association module that computes the marginal association probabilities in JPDA's sense resulting in probabilistical weights used in the lower body trackers' update steps. The lower body tracking consists of two parallely propagating EKFs using the developed kinematic-based motion model from Subsection 4.4.2. Upper body tracking is performed using a linear Kalman filter with a constant-acceleration (CA) model. The upper body's prediction is subsequently added to the local feet movement's prediction in the EKF of the left (Left foot EKF) and right foot (Right foot EKF), respectively, to display the actual locomotion relatively to the environment. JPDA assigns the lower body's detections to the respective foot, which enables the tracking of both feet separately and allows the specific recognition of behavioral indications initiated by the leg motion in addition to the precise estimation of the foot's position, velocity, and kinematic parameter. In the following, the mathematical implementation of the pedestrian tracking algorithm is provided by covering the essential tracking steps of prediction, data association, and update.

1) *Prediction step*: Assuming a linear Gaussian system with mean $\mathbf{x}_{k-1|k-1}^{\mathrm{ub}}$ and covariance $\mathbf{P}_{k-1|k-1}^{\mathrm{ub}}$ at time $k-1$, the upper body's state and measurement functions propagate in the sense of a linear Kalman filter which reads

$$\hat{\mathbf{x}}_{k|k-1}^{\mathrm{ub}} = \mathbf{F}_{k-1}^{\mathrm{ub}}\hat{\mathbf{x}}_{k-1|k-1}^{\mathrm{ub}} + \mathbf{v}_{k-1}^{\mathrm{ub}} \tag{4.69}$$

$$\mathbf{z}_k^{\mathrm{ub}} = \mathbf{H}_k^{\mathrm{ub}}\hat{\mathbf{x}}_{k|k-1}^{\mathrm{ub}} + \mathbf{e}_k^{\mathrm{ub}}, \tag{4.70}$$

where the state vector

$$\hat{\mathbf{x}}_{k|k-1}^{\mathrm{ub}} = [x^{\mathrm{ub}}, v_x^{\mathrm{ub}}, a_x^{\mathrm{ub}}, y^{\mathrm{ub}}, v_y^{\mathrm{ub}}, a_y^{\mathrm{ub}}]^T, \tag{4.71}$$

follows a CA model in the x-y plane of the sensor coordinate system as shown in Figure 4.19 and comprises estimated information about the upper body's position, velocity, and acceleration. The measurement vector

$$\mathbf{z}_k^{\mathrm{ub}} = [\bar{r}_k^{\mathrm{ub}}\cos(\bar{\phi}_k^{\mathrm{ub}}) \quad \bar{r}_k^{\mathrm{ub}}\sin(\bar{\phi}_k^{\mathrm{ub}})]^T, \tag{4.72}$$

contains the averaged Cartesian coordinates of the upper body detections $\mathbf{Z}_k^{\mathrm{ub}}$ as functions of averaged range $\bar{r}_k$ and averaged azimuth angle $\bar{\phi}_k$. The idea is to overcome the nonlinearity of the measured variables r_k and ϕ_k by using so-called pseudo-coordinates. The measurement noise variance matrix is empirically set to be

$$\mathbf{R}_k^{\mathrm{ub}} = \begin{pmatrix} \sigma_x^2 & 0 \\ 0 & \sigma_y^2 \end{pmatrix} \quad \text{with} \quad \sigma_x = \sigma_y = 0.5\,m. \tag{4.73}$$

The EKFs of the left and right foot, respectively, evolve according to the discretized

state transition model

$$
\hat{\mathbf{x}}_{k|k-1}^{i,\mathrm{lb}} = \begin{pmatrix} x^{i,\mathrm{lb}} \\ y^{i,\mathrm{lb}} \\ v_x^{i,\mathrm{lb}} \\ v_y^{i,\mathrm{lb}} \\ \tilde{\omega}^i \\ \tilde{\varphi}^i \\ \tilde{a}^i \end{pmatrix} = \begin{pmatrix} \xi_{k|k-1}^{x,\mathrm{ub}} + \mathbf{f}_{k-1|k-1}^{i,x} \\ \xi_{k|k-1}^{y,\mathrm{ub}} + \mathbf{f}_{k-1|k-1}^{i,y} \\ \mathrm{d}x^{i,\mathrm{lb}}/\mathrm{d}T \\ \mathrm{d}y^{i,\mathrm{lb}}/\mathrm{d}T \\ \tilde{\omega}_{k-1|k-1}^i \\ \tilde{\varphi}_{k-1|k-1}^i \\ \left(v_x^{\mathrm{ub}^2} + v_y^{\mathrm{ub}^2} \right)^{1/2} / v_{\mathrm{init}} \end{pmatrix}, \tag{4.74}
$$

with

$$
\xi_{k|k-1}^{x,\mathrm{ub}} = x^{\mathrm{ub}} + v_x^{\mathrm{ub}}T + a_x^{\mathrm{ub}}\frac{T^2}{2} \tag{4.75}
$$

$$
\xi_{k|k-1}^{y,\mathrm{ub}} = y^{\mathrm{ub}} + v_y^{\mathrm{ub}}T + a_y^{\mathrm{ub}}\frac{T^2}{2}, \tag{4.76}
$$

and

$$
\mathbf{f}_{k-1|k-1}^{i,x} = \mathbf{f}_{k-1|k-1}^{i,\breve{x}} \cos(\alpha_k^{\mathrm{ub}}) - \mathbf{f}_{k-1|k-1}^{i,\breve{y}} \sin(\alpha_k^{\mathrm{ub}}) \tag{4.77}
$$

$$
\mathbf{f}_{k-1|k-1}^{i,y} = \mathbf{f}_{k-1|k-1}^{i,\breve{y}} \cos(\alpha_k^{\mathrm{ub}}) + \mathbf{f}_{k-1|k-1}^{i,\breve{x}} \sin(\alpha_k^{\mathrm{ub}}), \tag{4.78}
$$

expressing the Cartesian coordinates of the local foot propagation transformed from the local body coordinate system, $(\breve{x}, \breve{y})$, into the sensor coordinate system, (x, y), where

$$
\alpha_k^{\mathrm{ub}} = \tan^{-1} \left(\frac{v_y^{\mathrm{ub}}}{v_x^{\mathrm{ub}}} \right) \big|_{\alpha_k = \alpha_{k|k-1}}, \tag{4.79}
$$

is the predicted orientation of the upper body. Note that the upper body's linear propagation in (4.75) and (4.76) is added to the lower body's transformed local position in (4.74) to realize actual locomotion with respect to the environment. The velocities $v_x^{i,\mathrm{lb}}$ and $v_y^{i,\mathrm{lb}}$ are obtained by differentiating $x^{i,\mathrm{lb}}$ and $y^{i,\mathrm{lb}}$ with respect to T. The kinematic parameters $\tilde{\omega}^i$ and $\tilde{\varphi}^i$ are not subject to a specified dependence. Preferably, the most convincing results are accomplished by keeping them variable throughout the tracking routine. The global amplitude $\tilde{a}^i$ relies on the normalized velocity, where v_{init} is a normalization factor referring to the velocity the motion model was approximated initially. The covariance predictions for both EKFs are performed as in (4.25) according to the state vector's local linearization principle in (4.30) given by

$$
\mathbf{P}_{k|k-1}^{i,\mathrm{lb}} = \hat{\mathbf{F}}_{k-1}^{i,\mathrm{lb}} \mathbf{P}_{k-1|k-1}^{i,\mathrm{lb}} \hat{\mathbf{F}}_{k-1}^{i,\mathrm{lb},T} + \mathbf{Q}_{k-1}^{i,\mathrm{lb}}, \tag{4.80}
$$

where

$$
\hat{\mathbf{F}}_{k-1}^{i,\mathrm{lb}} = \frac{\partial \boldsymbol{f}_{k-1}^{i,\mathrm{lb}}(\mathbf{x}_{k-1}^{i,\mathrm{lb}})}{\partial\left(x, y, v_x, v_y, \tilde{\omega}, \tilde{\varphi}, \tilde{a} \right)} \big|_{\mathbf{x}_{k-1}^{i,\mathrm{lb}} = \hat{\mathbf{x}}_{k-1|k-1}^{i,\mathrm{lb}}}, \tag{4.81}
$$

denotes the Jacobian matrix containing the first-order partial derivatives of $\boldsymbol{f}_{k-1}^{i,\mathrm{lb}}(\mathbf{x}_{k-1}^{i,\mathrm{lb}})$ with respect to $\mathbf{x}_{k-1}^{i,\mathrm{lb}}$. The measurement vector $\mathbf{z}_k^{\mathrm{lb}}$ at time k provides the lower body's detections and reads

$$\mathbf{z}_k^{\mathrm{lb}} = \left[r_k^{\mathrm{lb}}, \phi_k^{\mathrm{lb}}, v_k^{D,\mathrm{lb}}\right]^T . \tag{4.82}$$

The corresponding measurement covariance matrix $\mathbf{R}_k^{\mathrm{lb}}$ takes the form

$$\mathbf{R}_k^{\mathrm{lb}} = \begin{pmatrix} \sigma_r^2 & 0 & 0 \\ 0 & \sigma_\phi^2 & 0 \\ 0 & 0 & \sigma_{v^D}^2 \end{pmatrix} . \tag{4.83}$$

The measurement function $\boldsymbol{h}_k^{i,\mathrm{lb}}(\hat{\mathbf{x}}_{k|k-1})$ maps the predicted state $\hat{\mathbf{x}}_{k|k-1}$ into the corresponding measurement space according to

$$\begin{pmatrix} r^{\mathrm{lb}} \\ \phi^{\mathrm{lb}} \\ v^{D,\mathrm{lb}} \end{pmatrix} \xleftarrow{\boldsymbol{h}_k^{i,\mathrm{lb}}(\mathbf{x}_k)} \begin{pmatrix} x^{i,\mathrm{lb}} \\ y^{i,\mathrm{lb}} \\ v_x^{i,\mathrm{lb}} \\ v_y^{i,\mathrm{lb}} \end{pmatrix} \qquad \boldsymbol{h}_k^{i,\mathrm{lb}}(\mathbf{x}_k) = \begin{pmatrix} \sqrt{x^{2,\,i,\mathrm{lb}} + y^{2,\,i,\mathrm{lb}}} \\ \tan^{-1}\left(\frac{y^{i,\mathrm{lb}}}{x^{i,\mathrm{lb}}}\right) \\ \frac{x^{i,\mathrm{lb}}v_x^{i,\mathrm{lb}} + y^{i,\mathrm{lb}}v_y^{i,\mathrm{lb}}}{\sqrt{x^{2,\,i,\mathrm{lb}} + y^{2,\,i,\mathrm{lb}}}} \end{pmatrix} \Big|_{\mathbf{x}_k = \hat{\mathbf{x}}_{k|k-1}} . \tag{4.84}$$

The local linearization of $\boldsymbol{h}_k^{i,\mathrm{lb}}(\hat{\mathbf{x}}_{k|k-1})$ yields the Jacobian $\hat{\mathbf{H}}_k^{i,\mathrm{lb}}$ evaluated at $\hat{\mathbf{x}}_{k|k-1}$

$$\hat{\mathbf{H}}_k^{i,\mathrm{lb}} = \frac{\partial \boldsymbol{h}_k^{i,\mathrm{lb}}(\mathbf{x}_k^{i,\mathrm{lb}})}{\partial\left(x, y, v_x, v_y, \tilde{\omega}, \tilde{\varphi}, \tilde{a}\right)} \Big|_{\mathbf{x}_k^{i,\mathrm{lb}} = \hat{\mathbf{x}}_{k|k-1}^{i,\mathrm{lb}}} . \tag{4.85}$$

2) *Data association*: To reduce the association computation for the lower bodys' detections, we first prune hypotheses with small weights, i.e., ellipsoidal gating for Gaussian densities is used as introduced in Subsection 4.3.6. By doing so, unlikely data associations are disregarded, which reduces the number of hypotheses in the EKFs posterior object densities. The ellipsoidal gating distance for detection $\mathbf{z}_k^j \in \mathbf{Z}_k^{\mathrm{lb}}$ and object i under hypothesis h can be expressed by [11]

$$d_{i,j,h}^2 = \underbrace{\left(\mathbf{z}_k^j - \hat{\mathbf{H}}_k^{i,h,\mathrm{lb}}\hat{\mathbf{x}}_{k|k-1}^{i,h,\mathrm{lb}}\right)}_{\varepsilon_k^{i,j}}^T (\mathbf{S}_k^{i,h,\mathrm{lb}})^{-1} \underbrace{\left(\mathbf{z}_k^j - \hat{\mathbf{H}}_k^{i,h,\mathrm{lb}}\hat{\mathbf{x}}_{k|k-1}^{i,h,\mathrm{lb}}\right)}_{\varepsilon_k^{i,j}} . \tag{4.86}$$

Consequently, if $d_{i,j,h}^2 \leq \gamma$, where γ denotes the gating threshold, $\mathbf{z}_k^j$ is considered as possible detection from object i under hypothesis h, and pruned otherwise. Based on the valid detections the marginal association probabilities are computed as presented in Subsubsection 4.3.7.

3) *Update step*: The update is performed in accordance to JPDA. For that, an expected innovation, $\boldsymbol{\varepsilon}_k$, is calculated as the sum over all valid detections, m_k, of the individual

innovations, $\boldsymbol{\varepsilon}_k^{i,j}$, weighted by the marginal association probabilities, $\beta_k^{i,j}$, given by

$$\boldsymbol{\varepsilon}_k = \sum_{j=1}^{m_k} \beta_k^{i,j} \boldsymbol{\varepsilon}_k^{i,j}. \tag{4.87}$$

The updated merged mean and covariance for object i include the expected innovation and can be expressed by the Kalman update

$$\hat{\mathbf{x}}_{k|k}^{i,\mathrm{lb}} = \hat{\mathbf{x}}_{k|k-1}^{i,\mathrm{lb}} + \mathbf{K}_k^{i,\mathrm{lb}} \boldsymbol{\varepsilon}_k \tag{4.88}$$

$$\mathbf{P}_{k|k}^{i,\mathrm{lb}} = \beta_k^{i,0} \mathbf{P}_{k|k-1}^{i,\mathrm{lb}} + \left(1 - \beta_k^{i,0}\right) \bar{\mathbf{P}}_k^{i,\mathrm{lb}} + \tilde{\mathbf{P}}_k^{i,\mathrm{lb}}, \tag{4.89}$$

with

$$\bar{\mathbf{P}}_k^{i,\mathrm{lb}} = \mathbf{P}_{k|k-1}^{i,\mathrm{lb}} - \mathbf{K}_k^{i,\mathrm{lb}} \left(\hat{\mathbf{H}}_k^{i,\mathrm{lb}} \mathbf{P}_{k|k-1}^{i,\mathrm{lb}} \hat{\mathbf{H}}_k^{i,\mathrm{lb}^T} + \mathbf{R}_k^{i,\mathrm{lb}} \right) \left(\mathbf{K}_k^{i,\mathrm{lb}} \right)^T \tag{4.90}$$

$$\tilde{\mathbf{P}}_k^{i,\mathrm{lb}} = \mathbf{K}_k^{i,\mathrm{lb}} \left(\left[\sum_{j=1}^{m_k} \beta_k^{i,j} \boldsymbol{\varepsilon}_k^{i,j} \left(\boldsymbol{\varepsilon}_k^{i,j} \right)^T \right] - \boldsymbol{\varepsilon}_k \boldsymbol{\varepsilon}_k^T \right) \left(\mathbf{K}_k^{i,\mathrm{lb}} \right)^T. \tag{4.91}$$

Note that $\mathbf{P}_{k|k}^{i,\mathrm{lb}}$ denotes a sum of three parts: The first part is the prior covariance weighted with the probability of no association $\beta_k^{i,0}$. The second part is the covariance, $\bar{\mathbf{P}}_k^{i,\mathrm{lb}}$, resulting from updating with the detections weighted with the probability of associating to any detections $1 - \beta_k^{i,0}$. The final component is the covariance, $\tilde{\mathbf{P}}_k^{i,\mathrm{lb}}$, describing the spread of innovations using the gated detections.

The updated mean and covariance for object i represent the mean and covariance of the marginal posterior mixture density in (4.53). Finally, the covariance update of the upper body tracker is carried out as in the standard Kalman filter given in Subsection 4.3.3, which completes the proposed model-based pedestrian tracking algorithm cycle.

4.4.5 Experimental results

Radar measurements are conducted with a microwave radar sensor (INRAS radarlog) with flexible chirp sequence parametrization in Table 4.3 and raw data access. The parametrization complies with current series radars for short-range applications. Furthermore, the Xsens MTw Awinda MoCap sytstem is used to reference the proposed tracking algorithm applied in two different movement scenarios. The following results are achieved utilizing the entire model-based pedestrian tracking approach provided in the previous subsections.

Longitudinal trajectory

Figure 4.21 illustrates the time-range-dependent μD spectrum of a pedestrian walking radially towards the radar. The depicted radar detections correspond to the lower body's reflections that mainly emanate from the legs. The periodic behavior of the feet motion is clearly observable and hence confirms the origin's assumption. While (a) displays mostly
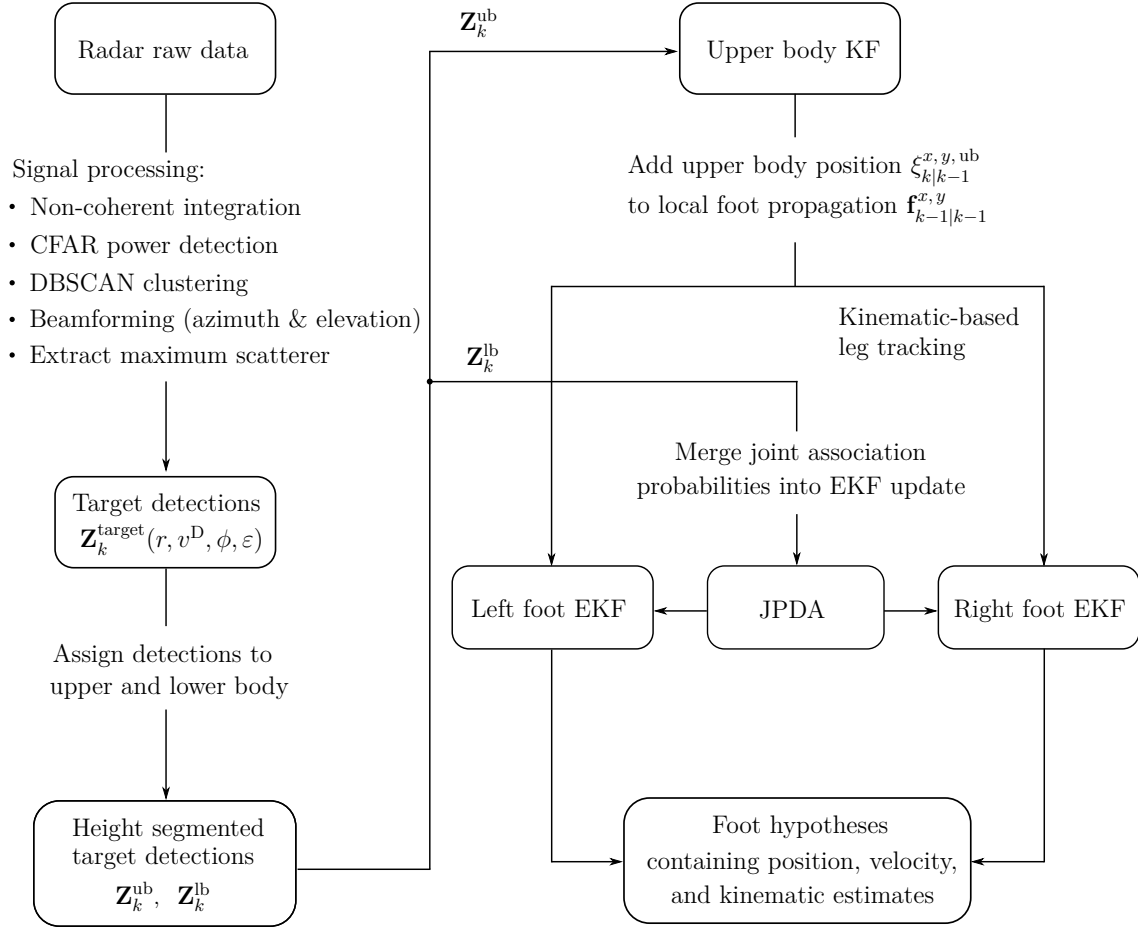
Figure 4.20: An architectural description of the pedestrian tracking system with time index k comprising a linear Kalman filter (Upper body KF) for the upper body's detections, two extended Kalman filters for the left foot (Left foot EKF) and the right foot (Right foot EKF), respectively, and a joint probabilistic data association (JPDA). Output are foot hypotheses containing position, velocity, and kinematic estimates.

the strides' Doppler amplitudes with magnitudes between $0\,\mathrm{m/s}$, and $5\,\mathrm{m/s}$, in (b), the stance and swing phases of a gait cycle become recognizable. Besides the backscattering, the feet references (Ref LF and Ref RF) recorded by the motion capture sensors and the tracking results for both feet (EKF LF and EKF RF) are shown.

Similarly to the results in Subsection 4.1.3, the perceptible discrepancy between the feet's reference amplitudes and the Doppler amplitude can be explained both by the positioning of the accelerometers and due to occurring signal processing effects for extended objects. The feet's tracking output, however, matches the course of the measured radar detections almost entirely. At the end of each swing phase, small pronounced periodical divergences are discernible, where the trackers assume positive velocities for a moment. This effect can be explained with the human motion simulation on which the model function is based. By adjusting the tracker's process noise, the overall divergence extent can be reduced. Table

4.4 displays the prevalent root mean square errors (RMSE) between the tracker outputs and the feets' MoCap data in each step of time of the longitudinal walking scenario from Figure 4.21 with given sensor resolutions in radial range and Doppler for comparability. Note that the tracker outputs incorporate extent information due to the objects' extended appearance. Hence, the RMSE as a standard metric for the performance evaluation of point targets may not possess a superior measure capability in the present instance [43].

Crossing trajectory with sudden turn

Regarding the motivation of this work, it is of particular interest to offer a procedure that, in addition to solely tracking feet motions and their consequential much-improved localization, possesses indications for a pedestrian's behavioral change. Sudden and spontaneous behavior changes may rapidly lead to possible collisions and must be detected as premature as feasible. Consequently, the proposed methodology covers a highly dynamic maneuver in which a pedestrian first approaches the sensor in a diagonal orientation and then abruptly changes his direction by 90°. Generally, lateral walking poses challenges for the tracking approach. Due to the Doppler effect's orientation dependence, non-radial movements are accompanied by smaller Doppler amplitudes that can drastically deviate from the actual velocities. Also, occlusion effects arise due to partially hidden leg surfaces while walking. For these reasons, the selected trajectory is greatly suited for evaluating the proposed approach's overall tracking performance.

Figure 4.22(a) shows the crossing trajectory tracking with a sudden turn depicted in the Cartesian x-y sensor plane. Here, the foot tracking results and the corresponding MoCap references of the respective foot sensors are given. As can be seen, the pedestrian steadily approaches the sensor along a diagonally crossing trajectory. After a momentary transient response phase, the trackers converge towards the references' course. The stance and swing phases identify the legs' stride behavior. While a continuous forward propagation indicates the swing phase, the stance phases show an increased concentration on the same spot.

Table 4.3: Radar Configurations

Parameter	Value
RF center frequency	76.5 GHz
RF bandwidth	1 GHz
Range resolution	0.15 m
Sampling frequency	10 MHz
Single chirp duration T	51.2 μs
IF samples per chirp m_s	512
Number of chirps	512
Chirp repetition interval G	60 μs
Doppler resolution	0.06 m/s
Sensor height	0.37 m
Azimuth 3 dB beamwidth (TX)	51°
Azimuth 3 dB beamwidth (RX)	76.5°
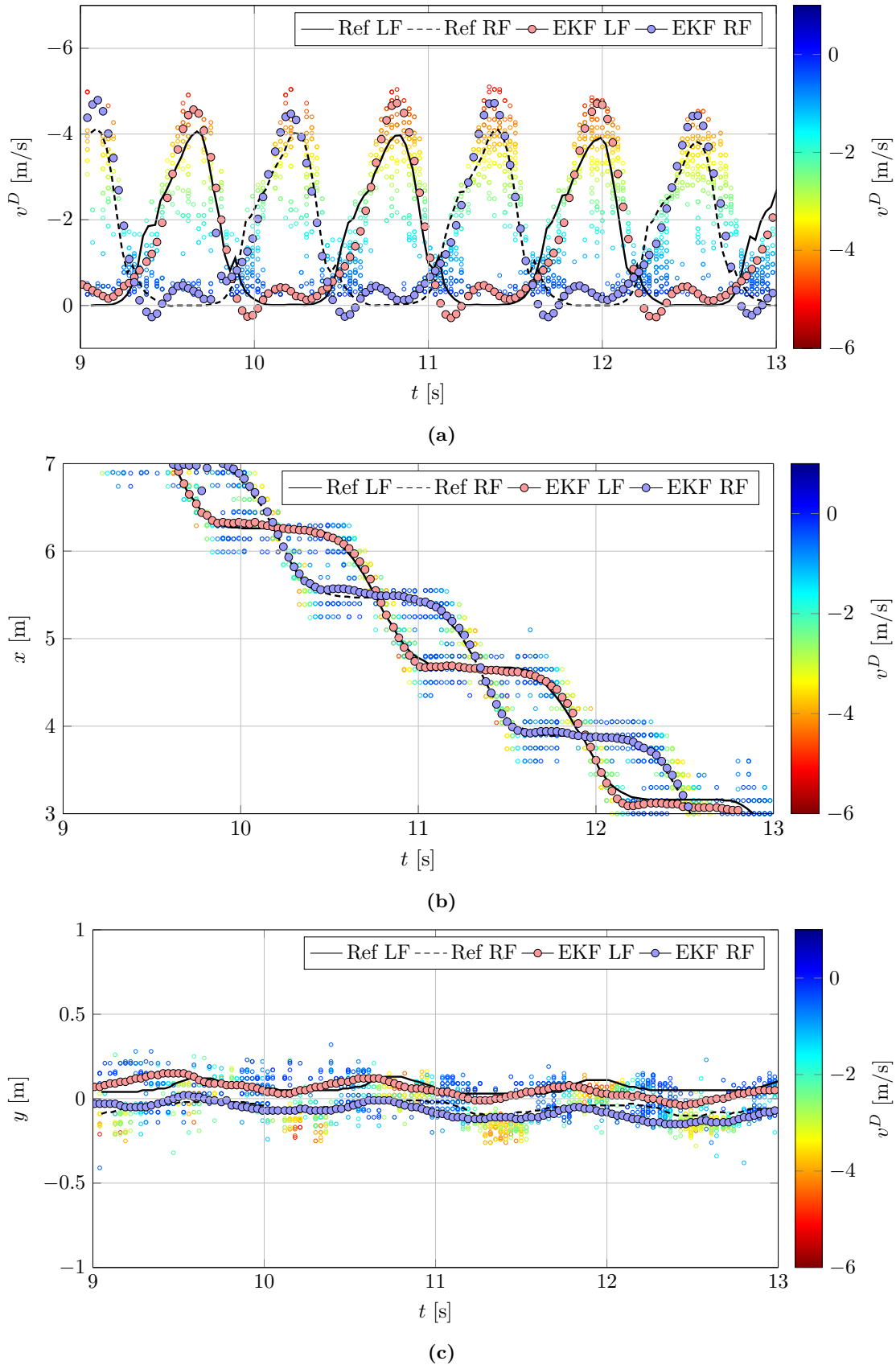
Figure 4.21: Time-dependent lower body radar spectrum of Doppler v^D (a), longitudinal distance x (b), and lateral distance y (c) with foot motion reference and foot tracking results for a radially walking pedestrian.

Table 4.4: RMSE of the longitudinal trajectory.

RMSE	$x\,[\mathrm{m}]$	$y\,[\mathrm{m}]$	$v_x\,[\mathrm{m/s}]$	$v_y\,[\mathrm{m/s}]$
Left foot	0.11	0.20	0.46	0.84
Right foot	0.15	0.20	0.53	0.91

Range resolution	$0.15\,[\mathrm{m}]$
Doppler resolution	$0.064\,[\mathrm{m/s}]$

Momentarily after traversing the radar's vertical line-of-sight, the upcoming change of movement is initiated by reducing the left leg's stride length. The abrupt turn is executed with the right leg rotating around the standing left leg. It is striking to note that the MoCap reference precisely depicts the leg rotation and confirms its accuracy. Next, the right foot is put on the ground, and the left foot accelerates towards the changed direction and completes the turn. Finally, the pedestrian keeps moving while departing from the radar sensor.

The excellent detectability of the right leg's turn motion raises the question of the extent to which this initial movement change can be utilized for sophisticated safety functions. Therefore, we will take a closer look at the right leg's behavior while turning, shown in Figure 4.22(b). The color coding corresponds to the velocity vector v resulting from the components v_x and v_y of the tracker. One possibility to detect changes in direction is to exploit the curvature of the resulting motion path. Since we are dealing with discrete vertices, the standard definition of curvature as the local fitting circle's reciprocal radius needs to be approximated.

Let $\zeta(s) = [y(s), x(s)]^T$ be a parametric representation of a twice differentiable plane curve with vertices s, the curvature can be expressed as

$$\Upsilon(s) = \frac{\mid x'(s)\,y''(s) - x''(s)\,y'(s) \mid}{\left(x'(s)^2 + y'(s)^2\right)^{3/2}}, \tag{4.92}$$

where the primes refer to the discrete differences with respect to s.

In the following, we only regard vertices whose velocities are at least $1.5\,\mathrm{m/s}$. The reason to do so is that a change of movement always accompanies a leg motion. Hence, introducing a velocity threshold appears to be legitimate in order to distinguish between a random motion and an actual indication. Finally, Figure 4.22(b) displays the occurring curvatures as blue arrows in accordance with the definition of (4.92). For reasons of clarity, only every second vertex is considered. The arrows' magnitudes equal to the arising curvatures at the individual vertices. As can be seen in the figure, the curvature correlates with the rotational leg motion and hence reaches its maximum at about $[y = 1.5\,\mathrm{m}, x = 5.1\,\mathrm{m}]^T$. As opposed to this, the curvature converges towards zero during walking phases without directional changes. This discriminability qualifies the curvature behavior to be employed as a distinctive leg feature to characterize a change in motion. Moreover, the proposed
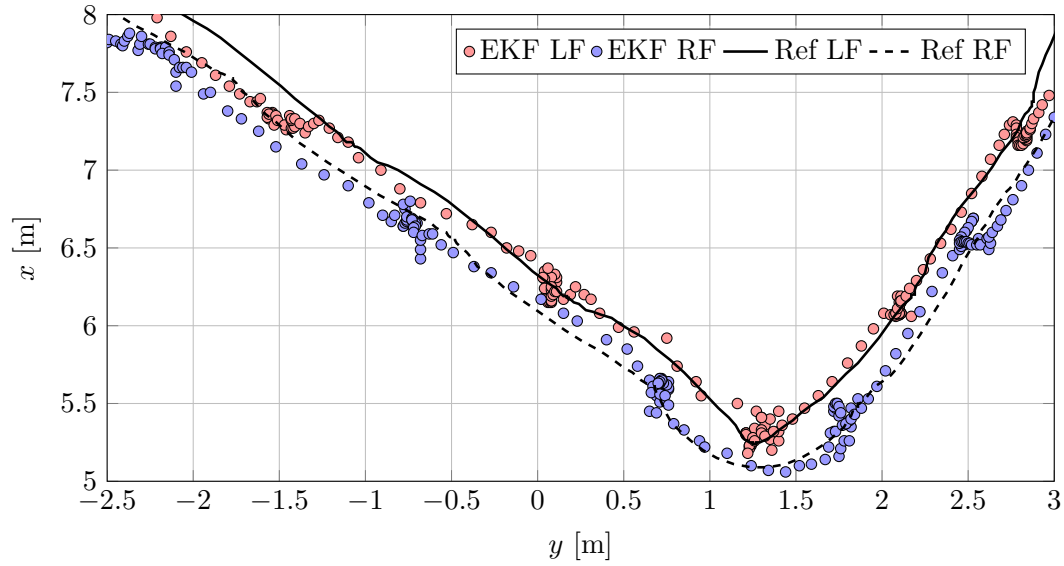
(a)
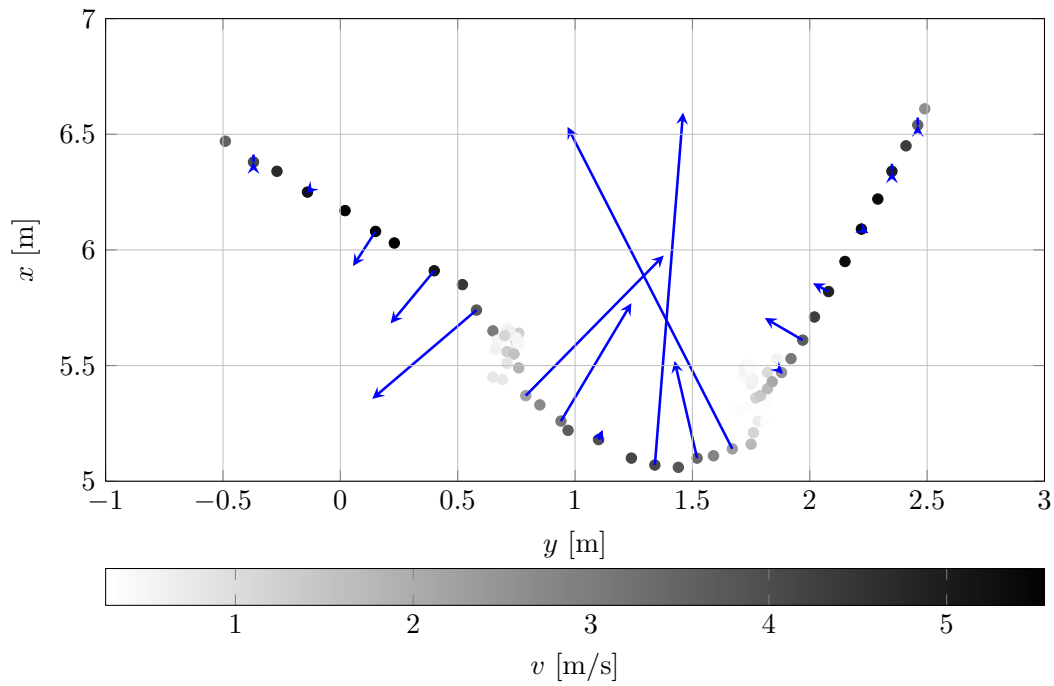
(b)

Figure 4.22: Tracking results of a crossing motion trajectory with sudden turn. (a) Tracking trajectory with left and right foot as well as motion capture reference. (b) Analysis of the turning motion of the right foot using curvature determination as directional change indication. The blue arrows indicate the radius of curvature for every second vertex possessing a velocity larger than 1.5 m/s.

computational effort is extremely manageable and can be performed in the current sensor cycle under the two previous tracking states' exclusive consideration, making it fairly suitable for automotive use.

4.5 Conclusion

The presented examinations analyzed high-resolution radar signatures of the human gait and provided specific μD-based algorithms for the potential use in automotive safety functions. By simultaneously using MoCap sensors during the radar measurement, the motion behavior matching reference could be provided. A detailed reconstruction of a pedestrian's movement behavior considering the individual limbs' motions during the beginning of a cycle could be exposed by combining range and Doppler data. Particular attention was paid to realistic, highly relevant movement scenarios such as traversing a road or a sudden difference in the pedestrian's velocity. The resulting rapidly altering Doppler components and their impacts on pedestrians' reliable detection could be demonstrated. Thereby, essential aspects for potential motion prediction, such as characteristic limb movements, could be obtained.

Proceeding from the μD spectra analysis insights, an extraction procedure for pedestrians' human body components has been proposed. A signal processing chain was introduced, which decomposes a pedestrian's range-Doppler spectrum into relevant scatterers. The concept is based on a modified CLEAN algorithm that incorporates a clustering procedure into the decomposing routine to identify scatterer groups. The approach was applied to a 77 GHz radar measurement of a walking pedestrian in chirp sequence mode. The extracted scatterers were assigned to their local positions based on a limb segmented body region and visualized over time. Synchronized MoCap data display the time-dependent velocity behavior of individual body parts and confirm the provided extraction procedure.

The procedure's limitations motivated a more advanced approach that uses elevation measurement capability to extract the reflections emanating from the legs' movements first and subsequently tracks their propagations over time based on a specific motion model. The model is based on human kinematics and displays the foot's movement with a high degree of variability and low computational effort during the tracking process. A tracking framework for multiple extended objects consisting of linear and extended Kalman filters has been designed to simultaneously estimate the macroscopic and microscopic motions of the upper and lower body, respectively. The arising data association problem is solved by incorporating joint probabilities in JPDA's sense in the Bayesian update step. The proposed technique enables the estimation of each leg's spatial and dynamic properties that can be utilized for integrated safety systems. The tracking implementation was evaluated in different moving scenarios, including a longitudinal path and a complex non-radially motion with sudden turn. MoCap sensors confirm tracking results' accuracy even in highly critical maneuvers such as abrupt motion transitions. The approach possesses valuable object information such as real-time turning detection and motion planning that can enormously enhance the future's pedestrian detection systems.

4.6 Discussion and outlook

The proposed method's key idea is to model the human walking kinematic as the foundation for a straightforward pedestrian tracking framework for automotive near-field applications. Incorporating the μD effect induced by the swinging legs' velocity distribution is fundamental for the presented technique. Spectral analyses with a high-resolution short-range radar have shown that μD signatures can be detected up to a distance of approximately 12-15 m, which qualifies the presented method as a specific near-field application that might trigger low-level reactions, such as a warning beep for the driver or a honk to warn the pedestrian. Regarding the progressive introduction of electric vehicles, which hardly emit any engine noise, this form of warning signal could focus increasingly. The maximum distance of detectability of μD signals depends significantly on the radiation pattern of the antennas as well as on the realizable SNR. More extensive ranges can be realized by a more bundled lobe shape of the radiated power.

The proposed framework groups the μD detections first and predicts the upper body motion, which is then added to the local foot prediction to realize locomotion. Accordingly, foot tracking's performance correlates to the accuracy of the upper body tracking. Consequently, if few or even no upper body detections emerge due to occluding, the leg tracking approach may fail. As mentioned, the average walking velocity can alternatively be estimated using the kinematic relationship between stride length and stride frequency. Note that the stride frequency is one of the tracker's direct outputs, whereas the stride length can be approximated in every time step as the resulting amplitude of the motion model by propagating the course with the current kinematic parameters. Regardless, more promising results are obtained with the presented approach. However, the stride kinematic utilization needs to be further studied to assure satisfying tracking performances in poor conditions.

Furthermore, the filters' initialization possesses challenges. For satisfactory tracking performance, the phase parameter of the feet must be chosen quite accurately. Besides, the best results are obtained when the objects are most distinguishable, e.g., at the beginning of a swing phase. If the objects are too close to each other, erroneous data assignments may occur, leading to the legs' confusion or a total blend of the filters. A potential solution is to use the average walking velocity resulting from the upper body tracking or, as pointed out, straight from the stride kinematic. The walking velocity behavior, including acceleration and deceleration, might be used to infer the respective leg's corresponding phase position.

CHAPTER 5

Summary and outlook

In this work, novel techniques for precise and robust radar-based perception and application of VRU were developed and analyzed concerning their potential for automotive safety functions. Previous radar systems are limited in the depth of resolution regarding extended objects. Within this thesis's scope, the objects were viewed in high resolution, and resulting features such as characteristic μD signatures were revealed. On this basis, approaches for the perception of behavioral indications and trajectory predictions were derived. Conclusions of the developed algorithms are drawn in Section 5.1, and an outlook to possible future work is provided in Section 5.2.

5.1 Developed algorithms

5.1.1 Adaptive pedaling extraction

For robust radar-based intend estimation of cyclists, an adaptive pedaling extraction procedure has been presented. The provided technique utilizes the physical principles that lead to extended ellipsoidal μD distributions in order to separate the wheels' components from the pedaling motion. The pedaling motion gives rise to detailed behavior indications through changes in the pedaling frequency detectable in real-time. By iteratively fitting ellipses to the front and rear wheels' range-Doppler detections, respectively, the pedaling motion is adaptively revealed and appears extractable. The approach enables the extraction of the pedaling movements for a series of consecutive radar frames. It could be shown that the extracted time-varying spectral components describe sinusoidal curves picturing the rotating pedals and parts of the cyclist's leg motions. The proposed technique has been validated by comparing with radar measurements of the same scenario comprising non-pedaling movements. The results demonstrate the significance regarding intend estimations of cyclists for short-range automotive safety applications.

5.1.2 Normalization of the orientation dependence

Extended radar targets are subject to orientation dependence affecting their geometric and kinematic extent due to the radar's radial measurement principle. This might affect the performance of classification procedures and reduces exploitable object information. A normalization methodology for the orientation-dependent range and Doppler amplitudes of a cyclist has been provided. By projecting the radar detections onto an approximation of the bicycle's actual longitudinal axis in Cartesian coordinates, the measured radial range ambiguities can be compensated. Simultaneously, the created axis approximates the bicycle's orientation, whereby the Doppler amplitudes can be normalized accordingly. The measured and normalized range-Doppler spectra have been compared and analyzed.

The results address standardization of range-Doppler spectra in the field of supervised and unsupervised machine learning techniques where classification rates may rise considerably using the provided method. The presented normalization approach depicts μD features such as pedaling motions with improved accuracy and forwards cyclists' intention estimation that may be manifesting on pedaling changes.

5.1.3 Model-based pedestrian tracking

A procedure for the targeted identification and extraction of pedestrians' human body components based on high-resolution radar data has been proposed by utilizing a decomposing spectral routine in the sense of a modified CLEAN algorithm. The results illustrate the procedure's fundamental feasibility to separate microscopic motions.

However, the approach faces limitations in more complex scenarios due to the body's orientation influences, such as data occlusion and considerably smaller Doppler amplitudes, which prevents its vigorous usability for safety-relevant functions. Further research has shown that behavioral indications of a pedestrian, such as incipient changes in the direction of movement, can be most reliably and feasibly detected via the legs' motion. Consequently, an advanced approach has been derived that incorporates elevation information to identify leg emanating detections and completes the μD signature. The provided method denotes the first approach based on automotive radar data that separates individual extremities and tracks their propagation over time. A computational simple but sophisticated motion model has been developed that approximates the lower body's nonlinear locomotion and displays the foot's movement comprising large degrees of variability. A tracking framework for multiple objects consisting of linear and extended Kalman filters has been designed to simultaneously estimate the upper and lower body's macroscopic and microscopic motions, respectively. By incorporating the joint probabilities in the sense of JPDA, the arising data association problem is tackled with simplified assumptions. The proposed technique estimates each leg's spatial and dynamic properties, which have been evaluated in complex maneuvers covering abrupt motion transitions. It could be shown that the tracked motion states possess valuable target information such as real-time turning detection and motion planning that contributes significantly to pedestrian detection systems of the future.

5.2 Outlook

A variety of extensions are conceivable for the proposed model-based pedestrian tracking approach. One possibility is the fusion of several radar sensors of a vehicle. If two radar sensors of different positioning illuminate an extended target, the object segmentation can be enlarged by integrating the geometric and kinematic extension. By doing so, its complete motion state can be determined by combining the detected point clouds. As a result, this can lead to more robust limb identification, allowing for a shorter initialization phase. Also, viewing from different angles can reduce occlusion and object loss, thereby increasing the overall tracking performance. Furthermore, fusion based on detection lists denotes an efficient interface and appears to be feasible regarding the electronic control unit's computational cost in a vehicle setup.

Besides the fusion of homogeneous sensor data, the combination of camera and radar data might be a valuable extension. In computer vision, learning-based methods exist that

estimate the pose of a walking human in real time. However, its accuracy is based on the quality of the sensor data. The complementary fusion of both technologies reveals advantages in the limb extraction procedure by providing rich semantic context information and establishes system-relevant redundancy regarding safety aspects.

Furthermore, an application of the tracking model to different motion patterns is conceivable in order to guarantee stable tracking performance. Due to their high agility, pedestrians tend to spontaneously change their movement behavior and thus place high demands on the vehicle systems. An example is a transition from walking to running, common among children and adolescents in urban areas. By storing different motion models inside an interacting multiple model (IMM) filter, the appropriate model's situation-adaptive choice could be executed based on the detected motion behavior. The IMM filter resolves the target motion uncertainty in the Bayesian framework by using various motion models at a time and switches between the deposited models according to their updated weights. The change between different motion patterns could be triggered by the detectable change of a pedestrian's velocity profile, i.e., a distinctive change of the motion state. In this way, the applicability and accuracy of leg tracking can be further increased.

In conclusion, the trend in the field of ADAS is towards an increasingly granular and reliable interpretation of the vehicle environment. At the same time, the use of high-resolution radars is still in the early stages of development. Technological advances in microelectronics are expected to significantly increase the angular, range, and velocity resolution of future automotive radar sensors. Sophisticated designing processes such as advanced antenna geometries and improved frequency modulation hardware directly impact the radar's resolving capacity. Similarly, the future allocation of much higher automotive frequency bands cannot be ruled out. The techniques presented in this work would directly benefit from resolution-enhancing developments and would experience an increase in performance and accuracy. Thus, pedestrian classification and behavioral indication algorithms are expected to gain importance for near-field applications, and the methods developed in this work will contribute significantly to the reliability of ADAS in series production.

Bibliography

1. A. Fischler, Martin and Robert C. Bolles: 'Random Sample Consensus: A Paradigm for Model Fitting with Applications To Image Analysis and Automated Cartography'. *Communications of the ACM* (June 1981), vol. 24: pp. 381–395 (cit. on p. 40).

2. Abdulatif, S., F. Aziz, B. Kleiner, and U. Schneider: 'Real-time capable micro-Doppler signature decomposition of walking human limbs'. *2017 IEEE Radar Conference (RadarConf)*. 2017: pp. 1093–1098 (cit. on pp. 51, 58).

3. Anderson, B.D.O. and J.B. Moore: *Optimal Filtering*. Dover Books on Electrical Engineering. Dover Publications, 2005 (cit. on p. 69).

4. Andres, M., K. Ishak, W. Menzel, and H.-L. Bloecher: 'Extraction of Micro-Doppler Signatures using Automotive Radar Sensors'. *Journal of RF-Engineering and Telecommunications*. 2012: pp. 371–377 (cit. on pp. 22, 51).

5. Arndt, R., R. Schweiger, W. Ritter, D. Paulus, and O. Lohlein: 'Detection and Tracking of Multiple Pedestrians in Automotive Applications'. *2007 IEEE Intelligent Vehicles Symposium*. 2007: pp. 13–18 (cit. on p. 64).

6. Ayhan, S., S. Scherr, A. Bhutani, B. Fischbach, M. Pauli, and T. Zwick: 'Impact of Frequency Ramp Nonlinearity, Phase Noise, and SNR on FMCW Radar Accuracy'. *IEEE Transactions on Microwave Theory and Techniques* (2016), vol. 64(10): pp. 3290–3301 (cit. on p. 27).

7. Bailey, T., B. Upcroft, and H. Durrant-Whyte: 'Validation Gating for Non-Linear Non-Gaussian Target Tracking'. *2006 9th International Conference on Information Fusion*. 2006: pp. 1–6 (cit. on pp. 79, 80).

8. Banzhaf, H., D. Nienhüser, S. Knoop, and J. M. Zöllner: 'The future of parking: A survey on automated valet parking with an outlook on high density parking'. *2017 IEEE Intelligent Vehicles Symposium (IV)*. 2017: pp. 1827–1834 (cit. on p. 2).

9. Bar-Shalom, Y., W.D. Blair, and Los Angeles. University Extension University of California: *Multitarget-multisensor Tracking: Applications and Advances*. Artech House radar library. Artech House, 1990 (cit. on pp. 79, 80, 131).

10. Bar-Shalom, Y., F. Daum, and J. Huang: 'The probabilistic data association filter'. *IEEE Control Systems Magazine* (2009), vol. 29(6): pp. 82–100 (cit. on pp. 79, 80).

11. Bar-Shalom, Y., T.E. Fortmann, P. Howlett, and A. Torokhti: *Tracking And Data Association by Yaakov BarShalom and Thomas E Fortmann*. Mathematics in Science and Engineering. Elsevier Science, 1988 (cit. on pp. 73–75, 78–82, 95).

12. Bar-Shalom, Y., X.R. Li, and T. Kirubarajan: *Estimation with Applications to Tracking and Navigation: Theory Algorithms and Software*. Wiley, 2004 (cit. on pp. 68, 69, 71).

13. Bar-Shalom, Y., P.K. Willett, and X. Tian: *Tracking and Data Fusion: A Handbook of Algorithms*. YBS Publishing, 2011 (cit. on pp. 75, 80, 81, 131).

14. Barrick, D.E. and Environmental Research Laboratories (U.S.): *FM/CW Radar Signals and Digital Processing*. FM/CW Radar Signals and Digital Processing Bd. 55. Environmental Research Laboratories, 1973 (cit. on pp. 16, 19).

15. Bartsch, A., F. Fitzek, and R. H. Rasshofer: 'Pedestrian recognition using automotive radar sensors'. *Advances in Radio Science* (2012), vol. 10: pp. 45–55 (cit. on pp. 3, 49).

16. Belgiovane, D. and C. Chen: 'Micro-Doppler characteristics of pedestrians and bicycles for automotive radar sensors at 77 GHz'. *2017 11th European Conference on Antennas and Propagation (EUCAP)*. Mar. 2017: pp. 2912–2916 (cit. on pp. 35, 36).

17. Belgiovane, D. and C. C. Chen: 'Micro-Doppler characteristics of pedestrians and bicycles for automotive radar sensors at 77 GHz'. *2017 11th European Conference on Antennas and Propagation (EUCAP)*. 2017: pp. 2912–2916 (cit. on p. 51).

18. Belgiovane, D. J., C. Chen, S. Y. -. Chien, and R. Sherony: 'Surrogate Bicycle Design for Millimeter-Wave Automotive Radar Pre-Collision Testing'. *IEEE Transactions on Intelligent Transportation Systems* (Sept. 2017), vol. 18(9): pp. 2413–2422 (cit. on p. 35).

19. Blackman, S.S. and R. Popoli: *Design and Analysis of Modern Tracking Systems*. Artech House radar library. Artech House, 1999 (cit. on p. 75).

20. Bose, R.: 'Lean CLEAN: Deconvolution Algorithm for Radar Imaging of Contiguous Targets'. *IEEE Transactions on Aerospace and Electronic Systems* (2011), vol. 47(3): pp. 2190–2199 (cit. on p. 59).

21. Boulic, Ronan, Nadia Magnenat-thalmann, and Daniel Thalmann: 'A Global Human Walking Model with Real-Time Kinematic Personification'. *The Visual Computer* (1990), vol. 6: pp. 344–358 (cit. on pp. 85–87).

22. Braun, M., S. Krebs, F. Flohr, and D. M. Gavrila: 'EuroCity Persons: A Novel Benchmark for Person Detection in Traffic Scenes'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2019), vol. 41(8): pp. 1844–1861 (cit. on p. 2).

23. Chen, V. C.: 'Analysis of radar micro-Doppler with time-frequency transform'. *Proceedings of the Tenth IEEE Workshop on Statistical Signal and Array Processing*. 2000: pp. 463–466 (cit. on p. 12).

24. Chen, V. C., F. Li, S. -. Ho, and H. Wechsler: 'Micro-Doppler effect in radar: phenomenon, model, and simulation study'. *IEEE Transactions on Aerospace and Electronic Systems* (2006), vol. 42(1): pp. 2–21 (cit. on pp. 12–14).

25. Chen, V.C.: *The Micro-Doppler Effect in Radar*. Norwood, MA 02062 USA: Artech House, 2011 (cit. on pp. 12, 13, 85, 87, 88).

26. Chojnacki, W., M. J. Brooks, A. van den Hengel, and D. Gawley: 'On the fitting of surfaces to data with covariances'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2000), vol. 22(11): pp. 1294–1303 (cit. on p. 39).

27. Costa, E.A. and R.B. Chadwick: *Signal Processing and Range Spreading in the FM-CW Radar*. U.S. Department of Commerce, National Oceanic and Atmospheric Administration Environmental Research Laboratories, 1984 (cit. on pp. 11, 16, 18–21, 26, 129).

28. Deep, Y., P. Held, S. S. Ram, D. Steinhauser, A. Gupta, F. Gruson, A. Koch, and A. Roy: 'Radar cross-sections of pedestrians at automotive radar frequencies using ray tracing and point scatterer modelling'. *IET Radar, Sonar Navigation* (6 June 2020), vol. 14: 833–844(11) (cit. on p. 49).

29. Diewald, F.: 'Objektklassifikation und Freiraumdetektion auf Basis bildgebender Radarsensorik für die Fahrzeugumfelderfassung'. PhD thesis. 2013 (cit. on pp. 10, 11, 17, 18, 91).

30. Doric, I., A. Reitberger, S. Wittmann, R. Harrison, and T. Brandmeier: 'A Novel Approach for the Test of Active Pedestrian Safety Systems'. *IEEE Transactions on Intelligent Transportation Systems* (2017), vol. 18(5): pp. 1299–1312 (cit. on p. 49).

31. Engels, F., P. Heidenreich, A. M. Zoubir, F. K. Jondral, and M. Wintermantel: 'Advances in Automotive Radar: A framework on computationally efficient high-resolution frequency estimation'. *IEEE Signal Processing Magazine* (2017), vol. 34(2): pp. 36–46 (cit. on pp. 3, 25, 26).

32. Engels, Florian, Philipp Heidenreich, Markus Wintermantel, Lukas Stacker, Muhammed Al Kadi, and Abdelhak M. Zoubir: 'Automotive Radars Signal Processing: Research Directions and Practical Challenges'. *IEEE Journal of Selected Topics in Signal Processing* (2021), vol.: pp. 1–1 (cit. on p. 16).

33. Ester, Martin, Hans-peter Kriegel, Jörg Sander, and Xiaowei Xu: 'A density-based algorithm for discovering clusters in large spatial databases with noise'. AAAI Press, 1996: pp. 226–231 (cit. on p. 31).

34. Flohr, F., M. Dumitru-Guzu, J. F. P. Kooij, and D. M. Gavrila: 'A Probabilistic Framework for Joint Pedestrian Head and Body Orientation Estimation'. *IEEE Transactions on Intelligent Transportation Systems* (2015), vol. 16(4): pp. 1872–1882 (cit. on p. 2).

35. Fogle, O. R. and B. D. Rigling: 'Micro-Range/Micro-Doppler Decomposition of Human Radar Signatures'. *IEEE Transactions on Aerospace and Electronic Systems* (2012), vol. 48(4): pp. 3058–3072 (cit. on p. 57).

36. FÖLSTER, F.: 'Erfassung ausgedehnter Objekte durch ein Automobil-Radar'. Technische Universität Hamburg-Harburg, 2006 (cit. on p. 27).

37. FORTMANN, T., Y. BAR-SHALOM, and M. SCHEFFE: 'Sonar tracking of multiple targets using joint probabilistic data association'. *IEEE Journal of Oceanic Engineering* (1983), vol. 8(3): pp. 173–184 (cit. on pp. 75, 131).

38. FRIIS, H. T.: 'A Note on a Simple Transmission Formula'. *Proceedings of the IRE* (1946), vol. 34(5): pp. 254–256 (cit. on p. 10).

39. GHALEB, A., L. VIGNAUD, and J. M. NICOLAS: 'Micro-Doppler analysis of wheels and pedestrians in ISAR imaging'. *IET Signal Processing* (2008), vol. 2(3): pp. 301–311 (cit. on p. 12).

40. GOODMAN, I.R., R.P. MAHLER, and H.T. NGUYEN: *Mathematics of Data Fusion*. Theory and Decision Library B. Springer Netherlands, 1997 (cit. on p. 81).

41. GRANSTROM, K., C. LUNDQUIST, and O. ORGUNER: 'Extended Target Tracking using a Gaussian-Mixture PHD Filter'. *IEEE Transactions on Aerospace and Electronic Systems* (2012), vol. 48(4): pp. 3268–3286 (cit. on p. 77).

42. GRANSTRÖM, K., C. LUNDQUIST, F. GUSTAFSSON, and U. ORGUNER: 'Random Set Methods: Estimation of Multiple Extended Objects'. *IEEE Robotics Automation Magazine* (2014), vol. 21(2): pp. 73–82 (cit. on pp. 66, 67, 71, 73).

43. GRANSTRÖM, KARL, MARCUS BAUM, and STEPHAN REUTER: 'Extended Object Tracking: Introduction, Overview and Applications'. *ISIF Journal of Advances in Information Fusion* (Dec. 2017), vol. 12(2) (cit. on pp. 73–77, 98).

44. GRESSMANN, M.: *Fahrradphysik und Biomechanik: Technik, Formeln, Gesetze*. Delius Klasing Verlag, 2017 (cit. on pp. 35, 36, 42).

45. HADERER, A., C. WAGNER, R. FEGER, and A. STELZER: 'Lateral velocity estimation using an FMCW radar'. *2009 European Radar Conference (EuRAD)*. 2009: pp. 129–132 (cit. on p. 19).

46. HARRIS, F. J.: 'On the use of windows for harmonic analysis with the discrete Fourier transform'. *Proceedings of the IEEE* (1978), vol. 66(1): pp. 51–83 (cit. on p. 26).

47. HARTENBERG, R. S. and J. DENAVIT: *Kinematic Synthesis of Linkages*. McGraw-Hill, 1964 (cit. on p. 85).

48. HE, Y., P. MOLCHANOV, T. SAKAMOTO, P. AUBRY, F. LE CHEVALIER, and A. YAROVOY: 'Range-Doppler surface: a tool to analyse human target in ultra-wideband radar'. *in IET Radar, Sonar Navigation* (2015), vol. 9(9): pp. 1240–1250 (cit. on p. 51).

49. HEIDENREICH, PHILIPP: 'Antenna Array Processing: Autocalibration and Fast High-Resolution Methods for Automotive Radar'. PhD thesis. Darmstadt: Technische Universität, Aug. 2012 (cit. on pp. 23, 24).

50. HELD, P., D. STEINHAUSER, A. KAMANN, T. HOLDGRÜN, I. DORIC, A. KOCH, and T. BRANDMEIER: 'Radar-Based Analysis of Pedestrian Micro-Doppler Signatures Using Motion Capture Sensors'. *2018 IEEE Intelligent Vehicles Symposium (IV)*. 2018: pp. 787–793 (cit. on pp. 3, 51–53, 55–57).

51. HELD, P., D. STEINHAUSER, A. KAMANN, A. KOCH, T. BRANDMEIER, and U. T. SCHWARZ: 'Micro-Doppler Extraction of Bicycle Pedaling Movements Using Automotive Radar'. *2019 IEEE Intelligent Vehicles Symposium (IV)*. 2019: pp. 744–749 (cit. on pp. 36, 38, 41, 43).

52. HELD, P., D. STEINHAUSER, A. KAMANN, A. KOCH, T. BRANDMEIER, and U. T. SCHWARZ: 'Normalization of Micro-Doppler Spectra for Cyclists Using High-Resolution Projection Technique'. *2019 IEEE International Conference on Vehicular Electronics and Safety (ICVES)*. 2019: pp. 1–6 (cit. on pp. 32, 36, 37, 44, 46).

53. HEUEL, S. and H. ROHLING: 'Pedestrian classification in automotive radar systems'. *2012 13th International Radar Symposium*. 2012: pp. 39–44 (cit. on pp. 3, 49).

54. HEUEL, S. and H. ROHLING: 'Pedestrian recognition based on 24 GHz radar sensors'. *11th International Radar Symposium*. June 2010: pp. 1–6 (cit. on p. 51).

55. HEUER, M., A. AL-HAMADI, and M. MEINECKE: 'An observation model for high resolution radar data in the context of an automotive pedestrian safety system'. *2013 14th International Radar Symposium (IRS)*. Vol. 2. 2013: pp. 714–719 (cit. on p. 64).

56. HEUER, M., A. AL-HAMADI, A. RAIN, M. MEINECKE, and H. ROHLING: 'Pedestrian tracking with occlusion using a 24 GHz automotive radar'. *2014 15th International Radar Symposium (IRS)*. 2014: pp. 1–4 (cit. on p. 64).

57. HO, Y. and R. LEE: 'A Bayesian approach to problems in stochastic estimation and control'. *IEEE Transactions on Automatic Control* (1964), vol. 9(4): pp. 333–339 (cit. on p. 68).

58. HÖGBOM, J. A.: 'Aperture Synthesis with a Non-Regular Distribution of Interferometer Baselines'. *Astronomy and Astrophysics Supplement Series*. Vol. 15. 1974: pp. 417–426 (cit. on p. 59).

59. INMAN, V.T., H.J. RALSTON, F. TODD, and J.C. LIEBERMAN: *Human walking*. Williams & Wilkins, 1981 (cit. on pp. 50, 84, 86).

60. JAY, F.: *IEEE Standard Dictionary of Electrical and Electronics Terms*. Institute of Electrical and Electronics Engineers, 1977 (cit. on p. 9).

61. KAMANN, A., P. HELD, F. PERRAS, P. ZAUMSEIL, T. BRANDMEIER, and U. T. SCHWARZ: 'Automotive Radar Multipath Propagation in Uncertain Environments'. *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. 2018: pp. 859–864 (cit. on p. 3).

62. KAMANN, A., D. STEINHAUSER, F. GRUSON, T. BRANDMEIER, and U. T. SCHWARZ: 'Extended Object Tracking using Spatially Resolved Micro-Doppler Signatures'. *IEEE Transactions on Intelligent Vehicles* (2020), vol.: pp. 1–1 (cit. on p. 12).

63. Kellner, D.: 'Verfahren zur Bestimmung von Objekt- und Eigenbewegung auf Basis der Dopplerinformation hochauflösender Radarsensoren'. PhD thesis. 2017 (cit. on pp. 27, 29).

64. Kellner, D., M. Barjenbruch, K. Dietmayer, J. Klappstein, and J. Dickmann: 'Instantaneous lateral velocity estimation of a vehicle using Doppler radar'. *Proceedings of the 16th International Conference on Information Fusion.* 2013: pp. 877–884 (cit. on p. 45).

65. Kellner, D., M. Barjenbruch, J. Klappstein, J. Dickmann, and K. Dietmayer: 'Wheel extraction based on micro doppler distribution using high-resolution radar'. *2015 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM).* 2015: pp. 1–4 (cit. on p. 12).

66. Kirubarajan, T. and Y. Bar-Shalom: 'Probabilistic data association techniques for target tracking in clutter'. *Proceedings of the IEEE* (2004), vol. 92(3): pp. 536–557 (cit. on p. 79).

67. Knoblauch, Richard L., Martin T. Pietrucha, and Marsha Nitzburg: 'Field Studies of Pedestrian Walking Speed and Start-Up Time'. *Transportation Research Record* (1996), vol. 1538(1): pp. 27–38 (cit. on p. 90).

68. Levinson, J. et al.: 'Towards fully autonomous driving: Systems and algorithms'. *2011 IEEE Intelligent Vehicles Symposium (IV).* 2011: pp. 163–168 (cit. on p. 1).

69. Ludloff, A.K.: *Praxiswissen Radar und Radarsignalverarbeitung.* Vieweg+Teubner Verlag, 2008 (cit. on p. 30).

70. Mahafza, B.R.: *Radar Systems Analysis and Design Using MATLAB.* Taylor & Francis, 2000 (cit. on pp. 10, 28).

71. Mahler, R.: 'PHD filters of higher order in target number'. *IEEE Transactions on Aerospace and Electronic Systems* (2007), vol. 43(4): pp. 1523–1543 (cit. on p. 77).

72. Mahler, R. P. S.: 'Multitarget Bayes filtering via first-order multitarget moments'. *IEEE Transactions on Aerospace and Electronic Systems* (2003), vol. 39(4): pp. 1152–1178 (cit. on p. 77).

73. Mahler, R.P.S.: *Statistical Multisource-multitarget Information Fusion.* Artech House information warfare library. Artech House, 2007 (cit. on p. 77).

74. Maybeck, P.S.: *Stochastic Models, Estimation, and Control.* ISSN. Elsevier Science, 1982 (cit. on p. 71).

75. Muntzinger, M. M., M. Aeberhard, S. Zuther, M. Mählisch, M. Schmid, J. Dickmann, and K. Dietmayer: 'Reliable automotive pre-crash system with out-of-sequence measurement processing'. *2010 IEEE Intelligent Vehicles Symposium.* 2010: pp. 1022–1027 (cit. on p. 1).

76. Padar, M. O., A. E. Ertan, and C. G. Candan: 'Classification of human motion using radar micro-Doppler signatures with hidden Markov models'. *2016 IEEE Radar Conference (RadarConf).* May 2016: pp. 1–6 (cit. on p. 51).

77. PATOLE, S. M., M. TORLAK, D. WANG, and M. ALI: 'Automotive radars: A review of signal processing techniques'. *IEEE Signal Processing Magazine* (2017), vol. 34(2): pp. 22–35 (cit. on pp. 3, 15, 49, 64).

78. PRIOLETTI, A., A. MØGELMOSE, P. GRISLERI, M. M. TRIVEDI, A. BROGGI, and T. B. MOESLUND: 'Part-Based Pedestrian Detection and Feature-Based Tracking for Driver Assistance: Real-Time, Robust Algorithms, and Evaluation'. *IEEE Transactions on Intelligent Transportation Systems* (2013), vol. 14(3): pp. 1346–1359 (cit. on p. 64).

79. PROTTER, M.H. and C.B.J. MORREY: *Intermediate Calculus.* Undergraduate Texts in Mathematics. Springer New York, 1996 (cit. on p. 38).

80. RAGESH, N. K. and R. RAJESH: 'Pedestrian Detection in Automotive Safety: Understanding State-of-the-Art'. *IEEE Access* (2019), vol. 7: pp. 47864–47890 (cit. on pp. 2, 64).

81. RAJ, R. G., V. C. CHEN, and R. LIPPS: 'Analysis of radar human gait signatures'. *IET Signal Processing* (2010), vol. 4(3): pp. 234–244 (cit. on p. 57).

82. REDMON, J., S. DIVVALA, R. GIRSHICK, and A. FARHADI: 'You Only Look Once: Unified, Real-Time Object Detection'. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).* 2016: pp. 779–788 (cit. on p. 2).

83. REID, D.: 'An algorithm for tracking multiple targets'. *IEEE Transactions on Automatic Control* (1979), vol. 24(6): pp. 843–854 (cit. on p. 75).

84. REN, S., K. HE, R. GIRSHICK, and J. SUN: 'Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2017), vol. 39(6): pp. 1137–1149 (cit. on p. 2).

85. RICHARDS, M.A.: *Fundamentals Of Radar Signal Processing.* McGraw-Hill Education (India) Pvt Limited, 2005 (cit. on pp. 28, 60).

86. RISTIC, B., S. ARULAMPALAM, and N. GORDON: *Beyond the Kalman Filter: Particle Filters for Tracking Applications.* Artech House, 2003 (cit. on pp. 65, 66, 68–70).

87. ROHLING, H., S. HEUEL, and H. RITTER: 'Pedestrian detection procedure integrated into an 24 GHz automotive radar'. *2010 IEEE Radar Conference.* 2010: pp. 1229–1232 (cit. on p. 49).

88. ROHLING, H. and R. MENDE: 'OS CFAR performance in a 77 GHz radar sensor for car application'. *Proceedings of International Radar Conference.* 1996: pp. 109–114 (cit. on pp. 28, 30).

89. ROOS, F., D. ELLENRIEDER, N. APPENRODT, J. DICKMANN, and C. WALDSCHMIDT: 'Range migration compensation for chirp-sequence based radar'. *2016 German Microwave Conference (GeMiC).* 2016: pp. 317–320 (cit. on p. 21).

90. ROTA, G.-C.: 'The number of partitions of a set'. *The American Mathematical Monthly* (May 1964), vol. 71(5): pp. 498–504 (cit. on p. 76).

91.   RÜBSAMEN, MICHAEL: 'Advanced direction-of-arrival estimation and beamforming techniques for multiple antenna systems'. PhD thesis. Darmstadt: Technische Universität, June 2011 (cit. on p. 23).

92.   SAMPSON, P.: 'Fitting conic sections to "very scattered" data: An iterative refinement of the bookstein algorithm'. *Comput. Graph. Image Process.* (1982), vol. 18: pp. 97–108 (cit. on p. 39).

93.   SCHUBERT, E., M. KUNERT, A. FRISCHEN, and W. MENZEL: 'A multi-reflection-point target model for classification of pedestrians by automotive radar'. *2014 11th European Radar Conference.* 2014: pp. 181–184 (cit. on p. 51).

94.   SCHUBERT, E., F. MEINL, M. KUNERT, and W. MENZEL: 'High resolution automotive radar measurements of vulnerable road users – pedestrians amp; cyclists'. *2015 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM).* Apr. 2015: pp. 1–4 (cit. on pp. 35, 36).

95.   STEINHAUSER, D., P. HELD, A. KAMANN, A. KOCH, and T. BRANDMEIER: 'Micro-Doppler Extraction of Pedestrian Limbs for High Resolution Automotive Radar'. *2019 IEEE Intelligent Vehicles Symposium (IV).* June 2019: pp. 764–769 (cit. on pp. 58, 61, 63).

96.   STEWART, C. V.: 'Minpran: a new robust estimator for computer vision'. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (Oct. 1995), vol. 17(10): pp. 925–938 (cit. on p. 41).

97.   STOICA, P. and R.L. MOSES: *Spectral Analysis of Signals.* Pearson Prentice Hall, 2005 (cit. on pp. 23, 25, 26).

98.   STRUTZ, T.: *Data Fitting and Uncertainty: A practical introduction to weighted least squares and beyond.* Springer Fachmedien Wiesbaden, 2015 (cit. on pp. 39, 40).

99.   STURM, C., GANG LI, and U. LÜBBERT: '79 GHz automotive radar and its opportunities for frequency and bandwidth agile operation'. *2017 18th International Radar Symposium (IRS).* 2017: pp. 1–6 (cit. on p. 3).

100. SZPAK, Z. L., W. CHOJNACKI, and A. van den HENGEL: 'Guaranteed Ellipse Fitting with a Confidence Region and an Uncertainty Measure for Centre, Axes, and Orientation'. *Journal of Mathematical Imaging and Vision* (2014), vol. (cit. on pp. 38–40).

101. SZPAK, Z. L., W. CHOJNACKI, and A. van den HENGEL: 'Guaranteed Ellipse Fitting with the Sampson Distance'. *12th European Conference on Computer Vision* (Oct. 2012), vol. (cit. on pp. 38, 39).

102. THRUN, S., W. BURGARD, D. FOX, and R.C. ARKIN: *Probabilistic Robotics.* Intelligent Robotics and Autonomous Agents series. MIT Press, 2005 (cit. on pp. 65, 67–72).

103. TSAO, J. and B. D. STEINBERG: 'Reduction of sidelobe and speckle artifacts in microwave imaging: the CLEAN technique'. *IEEE Transactions on Antennas and Propagation* (1988), vol. 36(4): pp. 543–556 (cit. on p. 59).

104. VAN TREES, H.L.: *Optimum Array Processing: Part IV of Detection, Estimation, and Modulation Theory*. Detection, Estimation, and Modulation Theory. Wiley, 2004 (cit. on p. 23).

105. VITERBI, A.: 'Error bounds for convolutional codes and an asymptotically optimum decoding algorithm'. *IEEE Transactions on Information Theory* (1967), vol. 13(2): pp. 260–269 (cit. on p. 69).

106. WAGNER, T., R. FEGER, and A. STELZER: 'Cluster CLEAN: An application of CLEAN to LFMCW radar systems'. *2014 11th European Radar Conference*. 2014: pp. 173–176 (cit. on p. 59).

107. WAGNER, T., R. FEGER, and A. STELZER: 'Radar Signal Processing for Jointly Estimating Tracks and Micro-Doppler Signatures'. *IEEE Access* (2017), vol. 5: pp. 1220–1238 (cit. on pp. 64, 65).

108. WINKLER, V.: 'Range Doppler detection for automotive FMCW radars'. *2007 European Microwave Conference*. 2007: pp. 1445–1448 (cit. on pp. 15, 17).

109. WINNER, H., S. HAKULI, F. LOTZ, and C. SINGER: *Handbook of Driver Assistance Systems: Basic Information, Components and Systems for Active Safety and Comfort*. Handbook of Driver Assistance Systems. Springer International Publishing, 2015 (cit. on pp. 1–3, 11, 12, 15, 22).

110. WORLD HEALTH ORGANIZATION: *Global status report on road safety 2018*. `https://www.who.int/publications-detail/global-status-report-on-road-safety-2018`. Online; accessed: 2020-12-21. Geneva, Switzerland, 2018 (cit. on p. 2).

111. XUEZHI WANG, S. CHALLA, and R. EVANS: 'Gating techniques for maneuvering target tracking in clutter'. *IEEE Transactions on Aerospace and Electronic Systems* (2002), vol. 38(3): pp. 1087–1097 (cit. on p. 79).

112. YAMADA, N., Y. TANAKA, and K. NISHIKAWA: 'Radar cross section for pedestrian in 76GHz band'. *2005 European Microwave Conference*. Vol. 2. 2005: 4 pp.–1018 (cit. on pp. 10, 49).

113. ZEDIKER, M. S., R. R. RICE, and J. H. HOLLISTER: 'Method for extending range and sensitivity of a fiber optic micro-Doppler ladar system and apparatus therefor'. *U.S. Patent 6,847,817*. 1998 (cit. on p. 12).

114. ZHAO, Z., P. ZHENG, S. XU, and X. WU: 'Object Detection With Deep Learning: A Review'. *IEEE Transactions on Neural Networks and Learning Systems* (2019), vol. 30(11): pp. 3212–3232 (cit. on p. 2).

List of Acronyms

Notation	Description
μD	Micro-Doppler
ADAS	Advanced driver assistance systems
AML	Approximated maximum likelihood
CA	Constant-acceleration
CFAR	Constant false alarm rate
CPHD	Cardinalized probability hypothesis density
CPI	Coherent processing interval
DBSCAN	Density-based spatial clustering of applications with noise
DFT	Discrete Fourier transform
DH	Denavit-Hartenberg
DOA	Direction-of-arrival
EKF	Extended Kalman filter
FFT	Fast Fourier transform
FMCW	Frequency-modulated continuous wave
GNN	Global nearest neighbor
HAD	Highly automated driving
IF	Intermediate frequency
IMM	Interacting multiple model
IR	Infrared
JPDA	Joint probabilistic data association
LM	Levenberg-Marquardt

Notation	Description
MEOT	Multiple-extended object tracking
MHT	Multiple hypothesis tracking
MoCap	Motion capture
MOT	Multiple-object tracking
MPOT	Multiple-point object tracking
OS-CFAR	Ordered-statistics constant false alarm rate
PDA	Probabilistic data association
PDF	Probability density function
PHD	Probability hypothesis density
PSF	Point spread function
RANSAC	Random sample consensus
RCS	Radar cross section
RF	Radio frequency
RFS	Random finite set
RMSE	Root mean square error
SNR	Signal-to-noise ratio
STFT	Short-time-Fourier-transform
UKF	Unscented Kalman filter
ULA	Uniform linear array
VRU	Vulnerable road user
VS	Visible spectrum

List of Figures

List of Tables

A Derivation of the rotation matrix 2.26

A cross product between two vectors can alternatively be expressed as the product of a skew-symmetric matrix and a vector

$$\vec{p} = \vec{u} \times \vec{r} \tag{A.1}$$

$$= \begin{pmatrix} u_y r_z - u_z r_y \\ u_z r_x - u_x r_z \\ u_x r_y - u_y r_x \end{pmatrix} \tag{A.2}$$

$$= \begin{pmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{pmatrix} \begin{pmatrix} r_x \\ r_y \\ r_z \end{pmatrix} = \hat{u}\vec{r}, \tag{A.3}$$

with

$$\hat{u} = \begin{pmatrix} 0 & -u_z & u_y \\ u_z & 0 & -u_x \\ -u_y & u_x & 0 \end{pmatrix}, \tag{A.4}$$

being a skew-symmetric matrix. This definition is often used for special orthogonal matrix groups (3-D rotation matrices SO(3)), which are defined by:

$$\text{SO}(3) = \{R \in \mathbf{R}^{3\times3} \mid R^T R = I, \det(R) = +1\}. \tag{A.5}$$

The time derivative of $R(t)R^T(t) = I$ yields

$$\dot{R}(t)R^T(t) + R(t)\dot{R}^T(t) = 0 \quad \Rightarrow \quad \dot{R}(t)R^T(t) = -\left(\dot{R}(t)R^T(t)\right)^T. \tag{A.6}$$

Note, that the matrix $\dot{R}(t)R^T(t) \in \mathbf{R}^{3\times3}$ is a skew-symmetric matrix. Consequently a vector $\vec{\omega} \in \mathbf{R}^3$ must exist such that:

$$\hat{\omega} = \dot{R}(t)R^T(t) \quad \Rightarrow \quad \dot{R}(t) = \hat{\omega}R(t). \tag{A.7}$$

Assuming $\vec{\omega}$ to be constant, we obtain

$$R(t) = e^{\hat{\omega}t}R(0), \tag{A.8}$$

where $e^{\hat{\omega}t}$ is the matrix exponential

$$e^{\hat{\omega}t} = I + \hat{\omega}t + \frac{(\hat{\omega}t)^2}{2!} + \cdots + \frac{(\hat{\omega}t)^n}{n!} + \cdots. \tag{A.9}$$

Setting $R(0) = I$ initially, we get

$$R(t) = e^{\hat{\omega}t}. \tag{A.10}$$

Since

$$\left(e^{\hat{\omega}t}\right)^{-1} = e^{-\hat{\omega}t} = e^{\hat{\omega}^T t} = \left(e^{-\hat{\omega}t}\right)^T, \tag{A.11}$$

then

$$\left(e^{\hat{\omega}t}\right)^T \left(e^{\hat{\omega}t}\right) = I, \tag{A.12}$$

yielding the determinant $\det\left(e^{\hat{\omega}t}\right) = \pm 1$. Moreover,

$$\det\left(e^{\hat{\omega}t}\right) = \det\left(e^{\frac{\hat{\omega}t}{2}} e^{\frac{\hat{\omega}t}{2}}\right) \tag{A.13}$$

$$= \left(\det\left(e^{\frac{\hat{\omega}t}{2}}\right)\right)^2 \geq 0, \tag{A.14}$$

showing that $\det\left(e^{\hat{\omega}t}\right) = 1$. As a consequence, the matrix $R(t) = e^{\hat{\omega}t}$ can be considered to be the 3-D rotation matrix, that rotates around the axis $\vec{\omega} \in \mathbf{R}^3$ with scalar angular velocity Ωt rad, where $\Omega = \|\vec{\omega}\|$. If $\vec{\omega}$ provides the rotation axis and the angular velocity, $R(t) = e^{\hat{\omega}t}$ at time t. The rotation matrix can be computed using Rodrigues's formula. Let $\vec{\omega}' \in \mathbf{R}^3$ with $\|\vec{\omega}'\| = 1$ and $\vec{\omega} = \Omega\vec{\omega}'$, one can show that the power of $\hat{\omega}'$ can be reduced by

$$\hat{\omega}'^3 = -\hat{\omega}'. \tag{A.15}$$

Simplifying the exponential series

$$e^{\hat{\omega}t} = I + \hat{\omega}t + \frac{(\hat{\omega}t)^2}{2!} + \cdots + \frac{(\hat{\omega}t)^n}{n!} + \cdots, \tag{A.16}$$

with

$$e^{\hat{\omega}t} = I + \left(\Omega t - \frac{(\Omega t)^3}{3!} + \frac{(\Omega t)^5}{5!} - \cdots\right)\hat{\omega}' \tag{A.17}$$

$$+ \left(\frac{(\Omega t)^2}{2!} - \frac{(\Omega t)^4}{4!} + \frac{(\Omega t)^6}{6!} - \cdots\right)\hat{\omega}'^2 \tag{A.18}$$

$$= I + \hat{\omega}'\sin(\Omega t) + \hat{\omega}'^2(1 - \cos(\Omega t)). \tag{A.19}$$

Therefore

$$\boldsymbol{R}_t = e^{\hat{\omega}t} = I + \hat{\omega}'\sin(\Omega t) + \hat{\omega}'^2(1 - \cos(\Omega t)). \tag{A.20}$$

B Derivation of the mixed radar signal 2.52

The transmitted signal is [27]

$$S_{\mathrm{T}}(t_m, m_s) = A \cos\left(2\pi \left(f_c + \dot{f}\frac{1}{2}t_m\right)t_m + \varphi_0\right). \tag{B.1}$$

The received signal is

$$S_{\mathrm{R}}(t_m, m_s) = \tilde{A} \cos\left(2\pi \left(f_c + \dot{f}\frac{1}{2}(t_m - \tau)\right)(t_m - \tau) + \varphi_0\right). \tag{B.2}$$

Demodulation is achieved by mixing the two signals according to

$$\cos(\alpha) \cdot \cos(\beta) = \frac{1}{2}\cos(\alpha - \beta) + \frac{1}{2}\cos(\alpha + \beta). \tag{B.3}$$

Applying (B.3) to (B.1) and (B.2) with subsequent low-pass filtering yields the difference between the original signals

$$S_{\mathrm{IF}}(t_m, m_s) = \tilde{A} \cos\left(2\pi\left(\Delta\varphi(t_m, m_s)\right)\right), \tag{B.4}$$

where

$$\Delta\varphi(t_m, m_s) = f_c\tau - \dot{f}\frac{1}{2}\left(\tau^2 - 2t_m\tau\right). \tag{B.5}$$

Substituting

$$\tau = \frac{2r(0)}{c} + \frac{2v_{\mathrm{r,rel}}}{c}m_sG + \frac{2v_{\mathrm{r,rel}}}{c}t_m, \tag{B.6}$$

in (B.5) and dropping all terms containing $1/c^2$ yields

$$\Delta\varphi(t_m, m_s) = f_c\frac{2r(0)}{c} + m_sf_c\frac{2v_{\mathrm{r,rel}}}{c}G \tag{B.7}$$

$$+ f_c\frac{2v_{\mathrm{r,rel}}}{c} + \dot{f}\frac{2r(0)}{c} + \underbrace{\dot{f}\frac{2v_{\mathrm{r,rel}}}{c}m_sGt_m}_{a} + \underbrace{\dot{f}\frac{2v_{\mathrm{r,rel}}}{c}t_m^2}_{b}. \tag{B.8}$$

The terms a and b in (B.7) appear to be maximal when $t_m = T/2$. Assuming a typical automotive radar parametrization, a and b are relatively small and can be dropped. Consequently, the final intermediate frequency signal is

$$S_{\mathrm{IF}}(t_m, m_s) = \cos\left(2\pi\left(\Delta\varphi(t_m, m_s)\right)\right) \tag{B.9}$$

$$\Delta\varphi(t_m, m_s) = f_c\frac{2r(0)}{c} + m_sf_c\frac{2v_{\mathrm{r,rel}}}{c}G + \left(f_c\frac{2v_{\mathrm{r,rel}}}{c} + \dot{f}\frac{2r(0)}{c}\right)t_m. \tag{B.10}$$

C Calculation of the marginal association probabilities 4.51

The JPDA algorithm evaluates the conditional probabilities of the following joint events at time k: [9], [13], [37]

$$\theta = \bigcap_{j=1}^{m} \theta^{i,j}, \tag{C.1}$$

where

$$\theta^{i,j} \hat{=} \{\text{measurement } j \text{ originated from target } i\}, \quad j = 1, \ldots, m \quad i = 0, \ldots, n. \tag{C.2}$$

A binary validation matrix is defined according to

$$\Omega = \left[\omega^{i,j} \right], \quad j = 1, \ldots, m \quad i = 0, \ldots, n, \tag{C.3}$$

where $\omega^{i,j}$ indicates whether measurement j is in the validation gate of object i. The index $i = 0$ stands for "no object" resulting in a corresponding column of Ω having only units. An exemplary validation matrix for one specific event including two objects and four measurements is of the form

$$\Omega = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix}, \tag{C.4}$$

where the rows refer to the measurements $j = 1, \ldots, m$ and the columns refer to the objects $i = 0, \ldots, n$ including the "no object" association as the first column. Each event θ can be represented by an event matrix

$$\hat{\Omega}(\theta) = \left[\hat{\omega}^{i,j}(\theta) \right] \tag{C.5}$$

where the binary elements in $\hat{\Omega}$ correspond to the associations in event θ. Thus

$$\hat{\omega}^{i,j}(\theta) = \begin{cases} 1, & \text{if } \theta^{i,j} \in \theta \\ 0, & \text{if otherwise.} \end{cases} \tag{C.6}$$

The feasible events are those that comply to the point-object assumption in (4.39) with

1) A measurement can originate only from one source, i.e.,

$$\sum_{i=0}^{n} \hat{\omega}^{i,j}(\theta) = 1, \quad j = 1, \ldots, m \tag{C.7}$$

2) At most one measurement can originate from a target, i.e.,

$$\delta_i(\theta) = \sum_{j=1}^{m} \hat{\omega}^{i,j}(\theta) \leq 1, \quad i = 1, \ldots, n \tag{C.8}$$

The determination of feasible events is performed by scanning Ω and assign the rows and columns (except for column $i = 0$) for every event according to the given point-object constraints.

The binary target detection indicator $\delta_i(\theta)$ reveals whether any measurement is associated to object i in event θ. The measurement association indicator

$$\tau^j(\theta) = \sum_{i=1}^{n} \hat{\omega}^{i,j}(\theta) = 1, \quad j = 1, \ldots, m \tag{C.9}$$

denotes whether measurement j is associated to an object in event θ. Consequently, the number of clutter (unassociated) measurements in event θ is

$$\phi(\theta) = \sum_{j=1}^{m} \left[1 - \tau_j(\theta)\right]. \tag{C.10}$$

The determination of the joint event probabilities is done using Bayes' rule. The conditional probability of a joint event that is conditioned on all meassurements up to and including time k can be expressed as

$$\begin{aligned}
\Pr\left(\theta_k | \mathbf{Z}_{1:k}\right) &= \Pr\{(\theta_k | \mathbf{Z}_k, \mathbf{Z}_{1:k-1}\} \\
&= \frac{1}{c} p\left(\mathbf{Z}_k | \theta_k, \mathbf{Z}_{1:k-1}\right) \Pr\{\theta_k | \mathbf{Z}_{1:k-1}\} \\
&= \frac{1}{c} p\left(\mathbf{Z}_k | \theta_k, \mathbf{Z}_{1:k-1}\right) \Pr\{\theta_k\},
\end{aligned} \tag{C.11}$$

where c denote a normalization constant. The measurements' joint probability density conditioned on the joint event θ reads

$$p\left(\mathbf{Z}_k | \theta_k, \mathbf{Z}_{1:k-1}\right) = \prod_{j=1}^{m_k} p\left(\mathbf{z}_k^j | \theta_k^{i,j}, \mathbf{Z}_{1:k-1}\right), \tag{C.12}$$

where

$$p\left(\mathbf{z}_k^j | \theta_k^{i,j}, \mathbf{Z}_{1:k-1}\right) = \begin{cases} f^i\left[\mathbf{z}_k^j\right] & \text{if } \tau^j(\theta) = 1 \\ V^{-1} & \text{if } \tau^j(\theta) = 0. \end{cases} \tag{C.13}$$

with

$$f^i\left[\mathbf{z}_k^j\right] \triangleq \mathcal{N}\left[\mathbf{z}_k^j; \hat{\mathbf{z}}_{k|k-1}^i, \mathbf{S}_k^i\right], \tag{C.14}$$

and $\hat{\mathbf{z}}^i_{k|k-1}$ denotes the predicted object detection for object i and $\mathbf{S}^i_k$ is the innovation covariance. V is the volume of the surveillance region where clutter associations are assumed to be uniformly distributed. Plugging (C.13) into (C.12) leads to

$$p\left(\mathbf{Z}_k|\theta_k, \mathbf{Z}_{1:k-1}\right) = V^{-\phi(\theta)} \prod_{j=1}^{m_k} \left[f^i \left[\mathbf{z}^j_k \right] \right]^{\tau^j(\theta)}, \tag{C.15}$$

where $\phi(\theta)$ indicates the total number of clutter measurements in event θ_k, and $\tau^j(\theta)$ selects the individual measurement densities in accordance to their associations.
The prior probability $\Pr(\theta_k)$ in (C.11) can be expressed by

$$\Pr\{\theta_k\} = \Pr\{\theta_k|\delta(\theta), \phi(\theta)\}\Pr\{\delta(\theta), \phi(\theta)\}, \tag{C.16}$$

where the first factor in (C.16) is

$$\Pr\{\theta_k|\delta(\theta), \phi(\theta)\} = \left(A^{m_k}_{m_k-\phi} \right)^{-1} = \left(\frac{m_k!}{\phi!} \right)^{-1} = \frac{\phi!}{m_k!}, \tag{C.17}$$

where A^m_l denotes arrangements of l out of m distinct objects. The second factor in (C.16) is

$$\Pr\{\delta(\theta), \phi(\theta)\} = \prod_{i=1}^{n} \left(P^{\mathrm{D},i} \right)^{\delta_i} \left(1 - P^{\mathrm{D},i} \right)^{1-\delta_i} \mu_F(\phi), \tag{C.18}$$

where $\mu_F(\phi)$ is the probability mass function of the clutter model. The indicators $\delta_i(\theta)$ select the probabilities of detection and no detection, respectively, considering the event θ_k. Plugging (C.17) and (C.18) into (C.16) gives the prior probability of a joint association event θ_k as

$$\Pr\{\theta_k\} = \frac{\phi!}{m_k!} \mu_F(\phi) \prod_{i=1}^{n} \left(P^{\mathrm{D},i} \right)^{\delta_i} \left(1 - P^{\mathrm{D},i} \right)^{1-\delta_i}. \tag{C.19}$$

The posterior probability of θ_k is achieved by plugging (C.15) and (C.19) into (C.11):

$$\Pr\left(\theta_k|\mathbf{Z}_{1:k}\right) = \frac{1}{c}\frac{\phi!}{m_k!} \mu_F(\phi) V^{-\phi} \prod_{j=1}^{m_k} \left[f^i \left[\mathbf{z}^j_k \right] \right]^{\tau^j} \prod_{i=1}^{n} \left(P^{\mathrm{D},i} \right)^{\delta_i} \left(1 - P^{\mathrm{D},i} \right)^{1-\delta_i}. \tag{C.20}$$

By summing over all joint events where the marginal event occurs, the marginal association probabilities are obtained:

$$\beta^{i,j} = \Pr\{\theta^i_k|\mathbf{Z}_{1:k}\} = \sum_{\theta} \Pr\left(\theta|\mathbf{Z}_{1:k}\right) \hat{\omega}^{i,j}(\theta), \quad j = 1, \ldots, m, \quad i = 0, \ldots, n \tag{C.21}$$

$$\beta^{i,0} = 1 - \sum_{j=1}^{m} \beta^{i,j}, \quad i = 0, \ldots, n. \tag{C.22}$$

Curriculum Vitae

Personal data

Name	Patrick Held
	born on 10.07.1990 in Ulm, Germany

Education

2017-now	**Chemnitz Technical University**, Chemnitz, Germany
	Institute of Physics
	External PhD student; Supervisor: Prof. Ulrich T. Schwarz
	Degree: Doctor rerum naturalium (Dr. rer. nat.)
2015-2017	**University of Applied Sciences Munich**, Munich, Germany
	Mechatronics (Sensors & actuators)
	Thesis: *Model-based object detection using automotive radar in the near vehicle environment*
	Degree: Master of Engineering (M.Eng.)
2011-2015	**University of Applied Sciences Ulm**, Ulm, Germany
	Mechatronics (Mechatronic vehicle systems)
	Thesis: *Radar-based investigation of electromagnetic waves' backscattering behavior on flowing waters*
	Degree: Bachelor of Engineering (B.Eng.)

Work Experience

2017-now	**University of Applied Sciences Ingolstadt, CARISSMA**, Ingolstadt, Germany
	Institute of Safety in Future Mobility (C-ISAFE)
	Scientific assistant, PhD candidate program
2016-2017	**Continental - ADC GmbH**, Lindau, Germany
	Divison Chassis & Safety, Advanced Engineering Radar
	Master's candidate

Erklärung

Ich erkläre, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs- oder Beratungsdiensten (Promotionsberaterinnen oder Promotionsberater oder anderer Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen. Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt. Ich erkläre hiermit, dass ich mich noch nie an einer in- oder ausländischen wissenschaftlichen Hochschule um die Promotion beworben habe oder gleichzeitig bewerbe.

06.05.2021   Patrick Held