Clustering of High Resolution Automotive Radar Detections and Subsequent Feature Extraction for Classification of Road Users

Eugen Schubert[1], Frank Meinl[1], Martin Kunert[1], Wolfgang Menzel[2]

[1]Robert Bosch GmbH, Advanced Engineering Sensor Systems
P.O. Box 16 61, 71226 Leonberg, GERMANY
Email: eugen.schubert@de.bosch.com

University Ulm, Institute of Microwave Techniques
Albert-Einstein-Allee, 41, 89081 Ulm, GERMANY

Abstract: *After successful entry into the automotive market some years ago radar sensor technology conquers more and more complex fields of automotive applications like Lane Change Assist or Automatic Emergency Braking systems. Future sensor generations will be challenged by constantly demanding resolution requirements not only requested by consumer rating organizations, but also by very demanding needs to realize automated driving tasks. Facing these novel challenges results in perpetual increasing of resolution requirements to guarantee highly detailed and accurate information of the surrounding area of the vehicle. Due to the high resolution operation mode a plenitude of scattering points reflected by each physical object will flood the detection list. This paper copes with the clustering of all these reflections into appropriate groups in order to exploit the advantages of multidimensional object size estimation and object classification.*

1. Introduction

With Adaptive Cruise Control as the first-time application radar sensor technology successfully entered the automotive market more than a decade ago. Since then radar sensors have been subjected to constantly increasing requirements due to the development of more complex applications like Lane Change Assist, Rear Cross Traffic Alert or Automatic Emergency Braking (AEB) systems. AEB systems for active pedestrian safety will be rated by EuroNCAP from 2016 on. Although many of the AEB systems already on the market are yet able to handle these tests, further developments aim to improve the mitigation performance for next generation vulnerable road user (VRU) protection systems. Hence sensor performance requirements will steadily increase over time. Furthermore, a bunch of novel active safety functions are waiting on the engineer's desks as next candidates for EuroNCAP protocol adoption beyond 2016 [1]. Some of them enlarge their AEB scope towards cyclists while other functions aim to support the driver when a high level of attention is required, like at turning manoeuvres with oncoming traffic, at dangerous intersections or during night. Beyond these, the challenge of automated driving will produce at least similar or even advanced resolution requirements to guarantee highly detailed and accurate information of the surrounding area of the vehicle. Estimation of object size, orientation or even object classification will soon become mandatory. Because radar performance is not significantly deteriorated by adverse environmental conditions like bad visibility at night, fog, rain or at distinctive backlight, future radar sensors must face the advanced resolution requirements claimed by novel applications to improve system robustness and overall traffic safety.

Pedestrian and cyclist radar characteristics were already derived from high resolution radar measurements, deeply analyzed and extensively discussed in previous work [2]. This paper is devoted to an important step in radar signal processing: the clustering of detections. This means grouping and assigning the bunch of detected scattering points within each physical object to adequate hypothesized objects. To form a realistic image of the surroundings, those groups of detections can be further used by multidimensional size estimation algorithms or extraction of features to enable object classification.

In the second section a brief insight in the used experimental sensor setup followed by the basic theory of the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) [3] is given. The third section presents a radar measurement of a scenario with multiple pedestrians and evaluates the capabilities of DBSCAN method. The last section deals with extraction of the multidimensional object size and the human step frequency as features for object classification.

2. Role of clustering in radar signal processing and brief insight in DBSCAN

The block diagram of an experimental sensor setup which is able to extract detection lists from high resolution radar measurements with adequate update times is depicted in Fig. 1. It consists of two main parts, the radar front-end and the digital signal processing part.

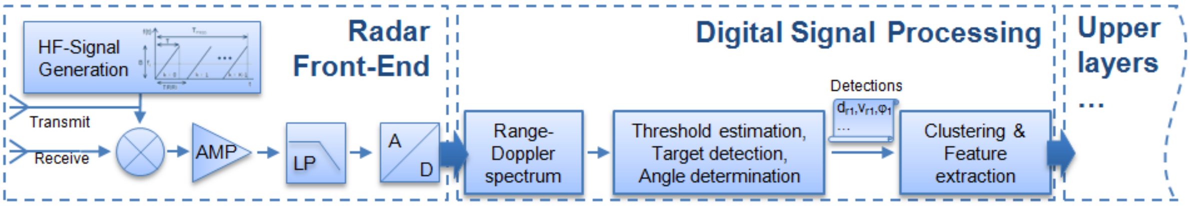


Fig. 1: Block diagram of the experimental radar sensor setup

The radar front-end is equipped with several antennas, off-the-shelf automotive radar Monolithic Microwave Integrated Circuits (MMICs) and high speed analogue digital converters capable for advanced FMCW chirp sequence modulation. With a typical parameter set a range resolution of approximately $\Delta d_r = 0.15$ m, a velocity resolution of $\Delta v_r = 0.1$ m/s and a maximum unambiguous velocity v_{max} of 100 km/h can be achieved, suitable for an application at urban speeds. The state of the art digital radar signal processing features a two dimensional Fast Fourier Transform (FFT) for range-Doppler spectrum calculation and detection threshold estimation based on ordered statistics constant false alarm rate (OS-CFAR). For angular position determination a uniform linear antenna array and a classical angle estimation algorithm is used. Thereby a vast list of detection-triples consisting of radial range d_r, radial velocity v_r and azimuth angle φ can be computed. Due to high resolution, it is possible to detect a plenitude of scattering points reflected by one physical object. In a next signal processing step, the major challenge is to cluster all these reflections to appropriate groups in order to exploit the advantages of multidimensional object size estimation or to extract features for classification purposes. The computational effort of object tracking, a further step in the signal processing chain, is thereby reduced because the number of tracks is decreased by previous clustering of the individual scattering points from the same object.

Many different clustering algorithms exist in literature. Because both hierarchical and partitional clustering algorithms suffer from robustness issues, are quite sensitive to noise or outliers, or require a-priori identification of the number of clusters, the density based cluster algorithm DBSCAN was chosen for implementation. At this point, only a short introduction into the algorithm is provided, while a detailed description is available in [3]. The principle of density based cluster algorithms is to identify groups of points by regions of different density according to the basic principle that the density outside of a cluster is lower than inside. Measured target properties like velocity, range, angle, or the radar cross section (RCS) are possible candidates to span a multi-dimensional feature space that can be processed by DBSCAN. Thereby two main operational parameters must be provided. A distance *Eps* that defines a neighbourhood applied to each point in which at least *MinPts* other points must appear to form a cluster. The shape of the neighbourhood depends on the distance function which can be chosen appropriate to the given geometrical and dynamic conditions. For simplicity the Euclidean distance was used in this work. Based on further conditions DBSCAN distinguishes between core and border points of a cluster and noise (points which could not be assigned to any cluster). Since the parameters *MinPts* and *Eps* are of global use they should be chosen according to the least dense cluster of the data base which is not considered to be pure noise.

3. Application of DBSCAN to high resolution radar measurements

To prove the suitability of DBSCAN it was applied to radar measurements of a scenario of three randomly walking pedestrians in front of the sensor. Fig. 2a shows the range-Doppler spectrum of the first measurement sequence at the start of the scenario. Beside some stationary, point-like targets the presence of three targets with kinematical spread can be separated. While the first person walks away radially it shows the largest kinematical spread. The third person shows the smallest spread caused by its nearly tangential motion direction with regard to the boresight of the sensor. The initial positions and walking directions can be easily derived from Fig. 2c. After application of OS-CFAR, peak detection and angle determination to the range-Doppler spectrum the scattering points can be transferred to a three dimensional space consisting of longitudinal distance x, lateral distance y and radial velocity v_r. Fig. 2b shows the detected scattering centres after clustering with DBSCAN to three separated clusters according to their physical origin. Thereby the parameter *MinPts* was chosen to a value of two, *Eps* in range domain to 1 m and *Eps* in velocity domain to 1 m/s. The centroid is marked with a larger green dot, stationary targets are coloured in black and noise detections are indicated with red crosses. Because the density in range domain of stationary detections differs from that of the moving detections, stationary and moving targets were separated by introducing an additional threshold of $v_{min} = 0.1$ m/s. As a result Fig. 2c shows trajectories of three clusters of the persons coloured according to their detection time from start of the scene (blue) to the end (red).

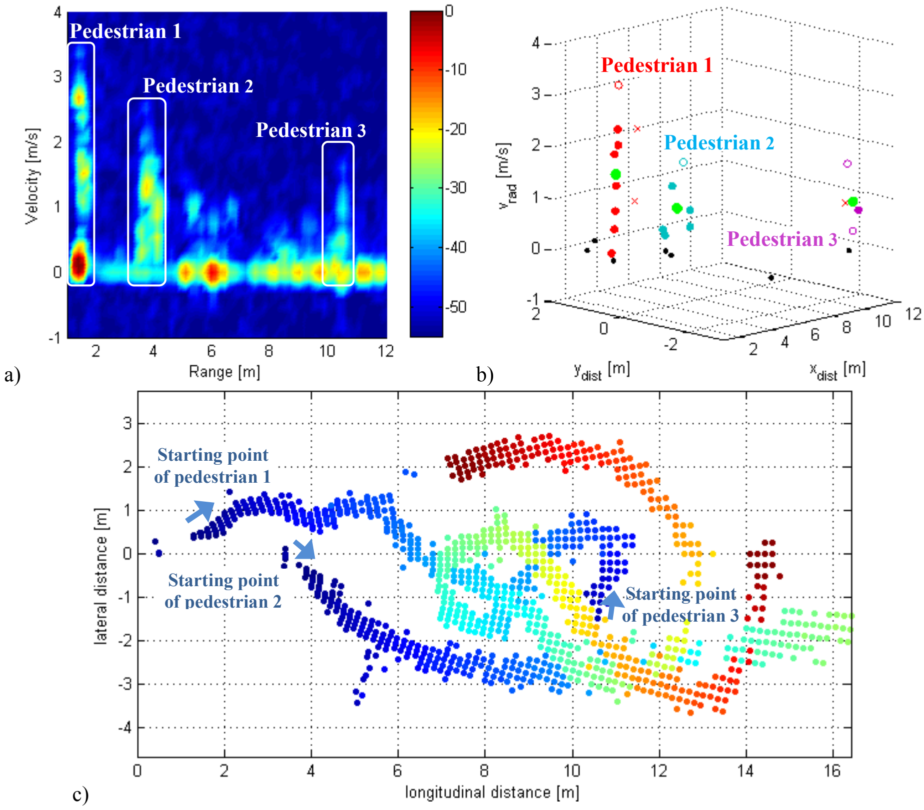
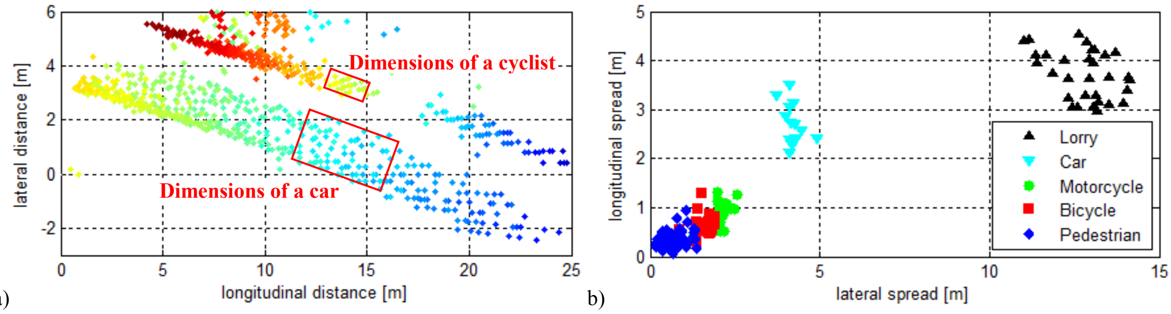


Fig. 2: a) Range-Doppler spectrum of the first measurement cycle of a scenario with three pedestrians walking in different directions, b) clustered detections according their origin, c) trajectories of the detection clusters of the three persons coloured according their detection time from start of the scene (blue) to the end (red).

4. Identification and extraction of features for classification of road users

After grouping relevant detections to appropriate clusters, their individual consistencies can be used to gather additional information about the nature of the underlying physical object, which then can be used for classification purposes. Using the spatial extents of the objects is the most straight forward approach. Fig. 3a shows the measured trajectories of a scene in which a cyclist has been overtaken by a faster moving car while the radar sensor was placed next to the street. The colour represents the detection time stamp starting from blue and ending with red. The object dimensions in lateral and longitudinal direction can be clearly derived from the measurements. With additional information of objects orientation their physical widths and lengths can be calculated. According to Fig. 3a a cyclist shows a length of 2 m and a width of 0.8 m, while the car shows a length of 5 m and a width of 2 m. Fig. 3b compares the longitudinal and lateral extents derived from multiple measurements of different types of road traffic participants. Thereby a radially walking pedestrian shows the smallest lateral extent of about one meter in lateral and longitudinal direction, while the extent of other measured road traffic participants is always larger. It is obvious that the object specific difference in size could be used for classification purposes.



a) b)

Fig. 3: a) Birds eye view of the detections (coloured according their measurement time) of a scene in which a cyclist is being overtaking by a faster moving car. b) Comparison of longitudinal and lateral spatial spread extracted from radar measurements of different road users.

In contrast to other sensor technologies suitable for pedestrian detection like video, laser or ultrasonic, a significant advantage of radar is the ability to measure not only the position and the spatial extend of an object, but also the velocity of each scattering point. As already shown in [4] pedestrians and wheeled road traffic participants like cyclists show significant differences in their kinematical spreads and in their velocity signatures. Whereas the maximum kinematical spread of a wheeled object is equal to twice the mean velocity over the time the kinematical spread of a walking pedestrian changes periodically according to the gait cycle from nearly zero to approximately three times the mean walking velocity. Simple algorithms like comparison of the maximum velocity of a cluster with its kinematical centroid respectively its mean velocity may lead to an easy feature for object classification. Nevertheless the robustness of such a feature is questionable, because the illumination of the wheels by radar strongly depends on the orientation of the vehicle to the sensor. The evaluation of the periodically changing kinematical spread of a walking pedestrian depicts a more robust classification feature. Fig. 4a shows the velocity of the clustered scattering points of a radially walking pedestrian and in addition the velocity variance of the cluster over time. Obviously a periodicity of about two Hertz can be derived from the velocity variance over time. This frequency corresponds with the step frequency of the pedestrian derived from a cinematic analysis of the measured scene. This step frequency can be extracted as classification feature by application of a short time Fourier Transform of the velocity variance of several previous measurement cycles. As result of such an analysis the step frequency of approximately 2 Hz can be derived correctly from the spectrogram as provided at the bottom of Fig. 4a. Variance values of the 16 previous measurement sequences were used which gives a dead time for derivation of the feature of approximately half a second.

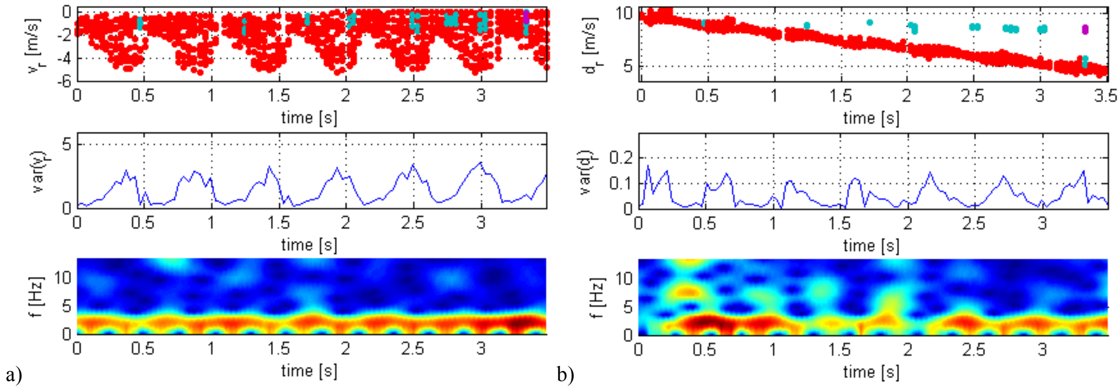


Fig. 4: Result of step frequency extraction based on radar detections in velocity domain (a) and range domain (b) of a radially walking pedestrian.

Because the spread in distance between the feet is also periodically changing, the same algorithm was applied to the detections in distance domain. The result is provided in Fig. 4b and shows less confidence compared to the step frequency extraction based on the kinematic spread. This is because the quotient of maximum step width and range resolution of about 7 is much smaller compared to the quotient of maximum kinematical spread and velocity resolution of about 50. One huge advantage of using the velocity variance instead of directly using the kinematic spread as classification feature is that the frequency of the variance is not significantly influenced by the walking direction of the pedestrian in relation to the sensor. With a high velocity resolution it is possible to extract the step frequency even from very small kinematical spread differences as they occur in the worst case of nearly tangential moving direction. Like already mentioned in [4] the extracted step frequency enables indication of the presence of a pedestrian walking along the road even under adverse visibility conditions. This indication can be used to prepare the driver to adopt his driving behavior. Using the step frequency as feature instead of just the kinematic spread can help to reduce unjustified warnings of such a safety function, because it helps to differentiate between walking pedestrians and other objects like turning cars, which also feature a kinematic spread.

Especially for AEB systems the avoidance of unjustified system interventions plays an important role. For such systems the mitigation capabilities are in strong relation to the earliest possible point in time where an activation of the brakes can be justified. Due to the high dynamics in pedestrian's motion this point lies very close to the collision time. Every additional indication about the actual intention of a detected pedestrian heading for a collision may be important for determination of the correct system reaction and consequently will lead to better mitigation capabilities while suppressing false system activations. To investigate the possibilities for contribution of high resolution radar to this topic, measurements of a laterally crossing pedestrian with and without stopping next to the road have been performed. Fig. 5a shows the lateral distance d_y, the radial velocity v_r and its variance extracted from radar measurements of a pedestrian crossing the road at 16 m distance from the sensor. Although this is a nearly tangential motion the periodic changes in velocity are clearly present and observable. Fig. 5b shows d_y, v_r and $var(v_r)$ derived from measurements where the pedestrian was told to stop next to the road in lateral distance of 2 m to the sensor. In Fig. 5b the pedestrian reaches the roadside at 1.4 s and starts to decelerate. A comparison of Fig. 5a and Fig. 5b shows differences in the kinematic spread and the velocity variance. While in case of continuously crossing in Fig. 5a the feet are accelerated one after each other to prepare the next step, Fig. 5b shows a nearly linear decreasing of the velocity after 1.4 s. Detailed evaluation of the radar measurements and the video captured in parallel allows the conclusion that this effect is caused by human motion characteristics. In case a walking pedestrian plans to stop next to the road the first foot reaches the stopping position similar to standard gait. After that and in contrast to the crossing pedestrian the second foot is placed slower next to the standing foot. This leads to less kinematic spread and less variance in velocity during the deceleration process, which may be used by AEB systems to indicate that the pedestrian will stop next to the road. Further a total time of approximately 600 ms can be derived from Fig. 5b for the complete deceleration process starting with placing the first foot next to the roadside and ending with the end of the upper body movement.

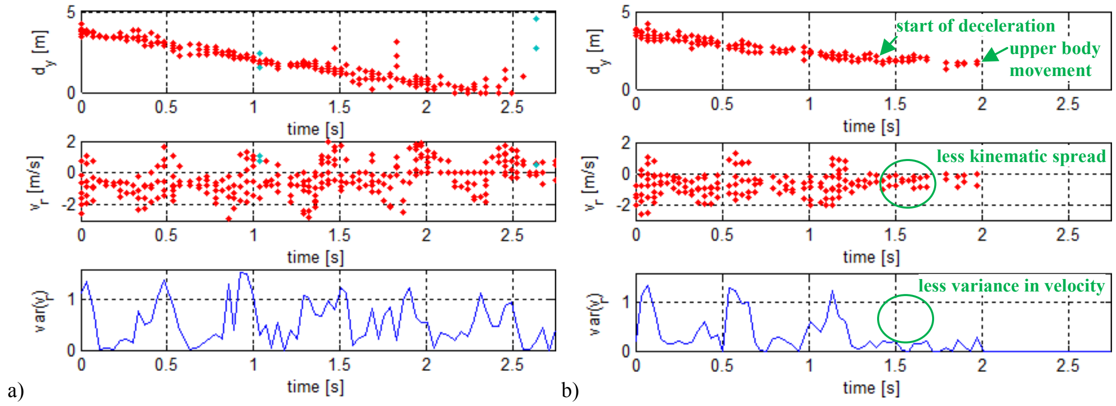


Fig. 5: Detected lateral distances d_y, radial velocities v_r and variance of the velocities of a pedestrian crossing the road in 16 m longitudinal distance (a) in comparison to a pedestrian stopping next to the road (b).

4. Conclusion

Facing the challenges of future driver assistance functions like left turning assist or highly automatic driving the suitability of the DBSCAN algorithm to the detections of high resolution radar measurements has been investigated. Beside advantages regarding computational effort for tracking detections, clustering of detections additionally enables the extraction of classification features. Object size, object orientation and even more object specific features like the human step frequency have been derived from clustered reflections provided by high resolution radar measurements. Furthermore investigations on the extraction of the actual intention of a detected pedestrian heading for a collision have been carried out. It was shown that the change in velocity variance can be used to indicate the pedestrian's intention to stop before the actual stopping process has been completed by the pedestrian.

5. Acknowledgements

The research leading to these results has received funding from the European Union's Seventh Framework Programme under grant agreement No.295 364 (www.deserve-project.eu) and the German Federal Ministry of Education and Research (BMBF). The consortium wants to express its gratitude to the European Commission and the BMBF for supporting the project.

References:

[1] EuroNCAP, „Roadmap 2020“, http://www.euroncap.com/files/Euro-NCAP-2020-Roadmap---June-2014-2---0-e11c0984-af94-420e-9d63-63edc8538745.pdf, June 2014

[2] E. Schubert, F. Meinl, M. Kunert, W. Menzel, “High Resolution Automotive Radar Measurements of Vulnerable Road Users–Pedestrians & Cyclists”, International Conference on Microwaves for Intelligent Mobility (ICMIM), April 2015, Heidelberg, Germany

[3] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, „A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise“, Proc. 2nd Int. Conf. on Knowledge Discovery and Data Mining, Portland, OR, AAAI Press, 1996, pp. 226-231

[4] E. Schubert, M. Kunert, W. Menzel, A. Frischen, “A Multi-Reflection-Point Target Model for Classification of Pedestrians by Automotive Radar”, 11th European Radar Conference (EuRAD), 2014, Rome, Italy