Reliable Orientation Estimation of Vehicles in High-Resolution Radar Images

Fabian Roos, Dominik Kellner, Jürgen Dickmann, and Christian Waldschmidt
Reliable Orientation Estimation of Vehicles in High-Resolution Radar Images

Fabian Roos, Student Member, IEEE, Dominik Kellner, Student Member, IEEE, Jürgen Dickmann, Member, IEEE, and Christian Waldschmidt, Senior Member, IEEE

Abstract—With new generations of high-resolution imaging radars the orientation of vehicles can be estimated without temporal filtering. This enables time critical systems to respond even faster. Based on a large data set this paper compares three generic algorithms for the orientation estimation of a vehicle. An experimental MIMO imaging radar is used to highlight the requirements of a robust algorithm. The well-known orientated bounding box and the so-called \textit{L-fit} are adapted for radar measurements and compared to a brute-force approach. A quality function selects the best fitted model and is a key factor to minimize alignment errors. Moreover, the reliability of the estimation is evaluated with respect to the aspect angle, the distance to the target, and the number of sensors. An approach to estimate the reliability of the current orientation estimation is introduced. It is shown that the root mean square error of the orientation estimation is 9.77° and 38% smaller compared to the common algorithm. In 50% of the evaluated measurements the orientation estimation error is smaller than 3.73°.

Index Terms—automotive radar, radar imaging, MIMO radar, orientation estimation, radar signal processing, Doppler radar, dimension extraction, bounding box estimation, reliability

I. INTRODUCTION

C AR manufacturers offer driver assistance and safety systems to prevent accidents and to enhance the comfort of driving. Especially radar sensors are commonly used for environment perception, because unlike camera, they can operate under severe weather conditions. To operate such systems in urban scenarios a more precise position, orientation, and dimension estimation is required compared to highways. This is due to a more dynamic environment with many different obstacles and fast changing situations. The driving path of approaching vehicles e.g., from side roads needs to be determined as fast as possible for real-time decision-making. With the availability of high-resolution image radars the possibility to estimate the orientation of vehicles arises as shown in [1]. Reasonable boundary conditions for automotive radar sensors to estimate the contour and thus the orientation of vehicles are investigated in [2]. A bandwidth larger than 1 GHz and an azimuth resolution better than 1° should be used. In this paper we show that a bandwidth of only 500 MHz is sufficient for orientation estimation.

To estimate the orientation of a target vehicle different approaches, like the Doppler distribution, box models, orthogonal line pairs, or radar response models, can be used and are explained in the following. Only algorithms using single measurements are considered, because the estimated orientation is used as input for a tracking not covered by this paper. This reduces latency in dynamic maneuvers.

In [3] the Doppler distribution of the target vehicle is used to set up the velocity vector and to determine the orientation if a linear motion is present, e.g., the vehicle travels along a straight line. But turning vehicles in urban scenarios are a key factor and are not covered by the approach.

The algorithms used in this paper are inspired from approaches utilized for laser scanners. The (orientated) bounding box approach is often selected which is based upon the determination of the convex hull. A cost-effective algorithm to find the minimal area box is the so-called \textit{rotating calipers} algorithm presented in [4]. A laser scanner can extract a detailed contour of vehicles, but if only the area of the enclosing rectangle is minimized, the resulting model can be misaligned. Therefore the authors of [5] introduce a symmetry assumption as an enhancement. It is assumed that the visible and the invisible part of the contour are symmetrical, which reduces the alignment error.

Often two visible sides of the vehicle are present, which can be modeled with two perpendicular lines. As the two sides are perpendicular and of different length, this fitted model is called \textit{L-fit}. This approach is applied to laser scanner data in [6]. The two orthogonal lines of the \textit{L-fit} can be found using the Hough transform, which is known from image analysis. The required geometric relations for detection are listed by [7]. [8] applies them to laser scanner data as well.

By using a radar response model as determined in [9] or [10] the different scattering centers of a vehicle are modeled and matched to the detected ones in the radar measurement. These models are determined from stationary targets and typically depend on the used radar sensors and type of vehicle. However, during motion the varying Doppler velocities of the targets can be exploited to resolve significant more reflections and to separate the target from close clutter.

In this paper, in addition to [11], a radar-specific orientation estimation algorithm is proposed, which is applied to measurement data gathered from an experimental MIMO radar sensor. Also the dimension estimation of the vehicle is presented together with a study on how the results depend on the number of used sensors and on the range to the target. Concepts to identify the quality of the current estimation are also presented.

F. Roos is with the Institute of Microwave Engineering, Ulm University, 89081 Ulm, Germany, Email: fabian.roos@uni-ulm.de
D. Kellner is with the Institute of Measurement, Control and Microtechnology, Ulm University, 89081 Ulm, Germany
J. Dickmann is with the Group Research and Advanced Engineering, Daimler AG, 89081 Ulm, Germany
C. Waldschmidt is with the Institute of Microwave Engineering, Ulm University, 89081 Ulm, Germany
This paper is organized as follows. In Section II the experimental setup and the preprocessing of radar data is explained. The requirements for the algorithms are discussed in Section III. This includes the description of how vehicles appear in the radar data. Three different orientation estimation algorithms are presented in detail in Section IV. Afterwards the scene is evaluated in Section V. The dependency on the introduced aspect angle, the range, and the number of sensors is shown. The possibility to extract the target vehicle’s dimensions is presented as well as its limitations. Also an approach to estimate the reliability of the estimation is introduced, followed by a conclusion in Section VI.

II. MEASUREMENT SETUP AND SIGNAL PREPROCESSING

At the front of the vehicle two experimental MIMO radar sensors using the 77 GHz frequency band are mounted at a height of 30 cm to 40 cm with an orientation directly towards the driving path as depicted in Fig. 2 (a). As stated in [12] a number of different antenna types have been used in the past: folded reflectarrays, lens antenna systems, mechanical rotating antennas, and planar patch configurations. The sensor for this paper uses a planar antenna system based on patches in a MIMO configuration with 10 Rx and 2 Tx channels. The placement of the transmit and receive antennas is depicted in Fig. 2 (b). While current available radar sensors mostly use the FMCW modulation technique as mentioned in [13], the transmit antennas emit alternatively a linear chirp-sequence modulation with 128 frequency ramps as shown in (c). This results in a time-division multiplexing and a nearly doubled aperture length using 256 ramps. The time delay between the two consecutive chirp blocks results in a Doppler dependent phase shift which must be known for the angle extraction. One antenna element in the virtual aperture is overlapping and the phase shift between the two chirp blocks can be determined. The receive antennas are equally spaced with a distance of 0.545λ and are evaluated in parallel with analog-to-digital converters with a resolution of 12 bit. The element distance is chosen as large as possible for a long aperture but so small that no ambiguities are present within the field of view. As a field...
of view $\pm 45^\circ$ can be evaluated. The transmit antennas consist of 12 patches in vertical direction and the receive antennas use 18. This leads to an elevation beamwidth of $\pm 3.5^\circ$. With a chirp duration $T_c$ of 20.48 $\mu$s and a chirp repetition time $T_r$ of 27.015 $\mu$s the pause of roughly 7 $\mu$s is required for the return and stabilization of the frequency ramp which is generated by a voltage controlled oscillator (VCO) and a fractional $N$ phase-locked loop ($N$-PLL) with a reference frequency of 50 MHz. Those timings yield an unambiguous maximal velocity of 36 m/s and a Doppler resolution of 0.28 m/s. Compared to radar sensors of series-production vehicles [13] the velocity resolution is enhanced by 50% due to the longer observation time. The adjustable bandwidth is set for this measurement to 500 MHz leading to a range resolution of 0.30 m. The internal low-pass filter limits the maximal distance to 53 m. The radar sensor outputs time-domain data which are stored for signal processing.

The typical angular resolution of state-of-the-art sensors is in between $3^\circ$ to $4^\circ$ using three to four channels as stated in [12]. With ten receive channels and the help of linear prediction as presented in [14] the angular resolution is enhanced to $1^\circ$. In contrast to state of the art automotive sensors that typically only measure a couple of scattering centers per target vehicle, the experimental MIMO radar sensor can usually register five to fifteen scattering centers. The number of resolvable centers is dependent on the orientation of the target car, the distance to it, and its velocity due to a better separability in range and velocity. For a peak detection the ordered statistic constant false-alarm rate (OS-CFAR, [15]) algorithm is used. After applying the OS-CFAR, there are multiple reflections of stationary and moving objects. To select only those reflection points of the vehicle, the clustering algorithm density based spatial clustering of applications with noise (DBSCAN, [16]) is used. The Doppler velocity is taken into account to identify moving targets.

The target vehicle as well as the vehicle with the radar sensors are both using an inertial measurement unit with differential GPS (DGPS) assistance to estimate the actual orientation and position to enable a precise error evaluation.

III. PROBLEM FORMULATION

Using radars, it is important to apply a robust approach which is motivated in Fig. 1. The target vehicle is driving in this part of the scene a figure eight, which is depicted with the trajectory (---). For several measurements the radar reflections (○) after applying the signal processing steps are shown together with a video image and the ground truth rectangle (----) representing the actual orientated vehicle. The east-north-up coordinate system is used, which means the driving direction of the car with the radar sensors is the $x$-axis and the $y$-axis is aligned to the left of it.

In contrast to laser scanners a radar is much more susceptible to detect multipath reflections. The transmitted electromagnetic wave can bounce off the ground and can get reflected from the underbody of a vehicle. This is clearly visible in Fig. 1 (a), where parts of the contour opposite to the viewing direction are detected. If the visible contour is orthogonal to the line of sight of the radar sensor, the reflections may outline and lead to detections in front of the actual target. This effect can be observed in (e) where the rear part is cluttered. The algorithm used should also cope with clustering faults which result in reflection points which do not support the vehicular model. Such measured points are called outliers and in (c) an example is shown. The algorithm should compensate such outliers so that the orientation is not flawed.

The gathered contour of the vehicle has, in contrast to the one obtained from laser scanners, stronger deviations and is dependent on the incident angle. As multipath reflections are present, the symmetry assumption of [5] cannot be assumed anymore, e.g., in Fig. 1 (d).

IV. ORIENTATION ESTIMATION ALGORITHM

After the clustering every target consists of several points which are used for the model fitting. The presented algorithms are all iterative ones, so that a quality function is required. Two algorithms are taken from literature and are enhanced to fulfill the requirements introduced in Section III. For comparison a brute-force algorithm is shown which also uses the quality function and should indicate the best possible outcome.

A. Two Perpendicular Lines (L-fit)

The idea to use two perpendicular lines, as presented in [6], must be adapted before it can be applied to radar data. The obtained reflections are cluttered and, as mentioned in Section III, measured points are also present inside the vehicle. In [6] the assignment of reflection points to one of the two lines is done using a corner candidate calculation. This approach cannot be used with radar data due to the fact that an assignment of points to one of the two lines is not trivial. There are points inside the vehicle, which originate from neither side. The proposed algorithm picks three points at random in each iteration. The first and the second point are assigned to the first line, the second and the third one belong to the perpendicular one. With this definition two perpendicular lines are calculated using linear regression as presented in [6]. For each iteration the current fit is evaluated using a quality function as presented in Section IV-D.
Fig. 3 shows two possible iterations, where (b) is the final accepted model. The three picked reflection points (●) are connected (-----) and are in most cases not orthogonal. Calculating the perpendicular fit results in the model (----).

B. Enhanced Orientated Bounding Box (EOBB)

The fundamental principle of this algorithm, which is introduced in [11], is the rotating calipers approach of [4]. After determining the convex hull the basic rectangle is created using the extreme values of the convex hull. In each iteration step the rectangle is rotated and aligned with one convex hull side (cf. iteration loop 1 in Fig. 4). The number of convex hull points is directly linked to the number of iterations.

To cope with outlier points as depicted in Fig. 1 (c) in further iterations every hull point will be ignored and the rectangles are rotated and evaluated once again (cf. iteration loop 2 in Fig. 4).

Ignoring only one outlier point is not sufficient due to outlier points commonly appear in groups and ignoring a single point often does not alter the convex hull essentially.

C. Brute-force Approach as Best Case Scenario (MainDir)

The possible rotations of the rectangle of the EOBB algorithm is determined by the convex hull. Although one hull point is ignored the best orientation may not be found. To allow a fine angular rotation step size the following brute-force approach is applied, which obviously is computationally expensive. See [11] for a detailed description.

In the radar reflections a strong distinctive contour is visible which is the main direction and detected using the random sample consensus (RANSAC) algorithm of [17]. This main direction can be rotated in a small range to enhance the step size as indicated in Fig. 5 (a). For each main direction the maximal enclosing rectangle is set up and shrunk step by step. To select the best model the quality function of Section IV-D is applied.

D. Evaluating Each Fit: the Quality Function

By picking randomly three points using the proposed L-fit algorithm, a misaligned line pair as shown in Fig. 3 (a) is down-weighted using a quality function. The fitted $L$ should be shifted as far as possible towards the radar sensor position in order to ensure an alignment on the contour points. This is done using a corridor weighting function as depicted in Fig. 6. Points in the inner corridor support the current model, in contrast to points between the inner and outer corridor. Reflections outside of the outer corridor are probably from the other line or inside the vehicle and should only result in a constant error.

The EOBB and the brute-force algorithm use the same quality function to rate each iteratively calculated model. An error value needs to be minimized which is further called optimization variable. The optimization variable consists of several important factors instead of just using the area of the rectangle as in [5].

First of all the area of the box model should be as small as possible since a misaligned bounding box typically consumes more space. Clustered points which lie on the contour of the bounding box support the current chosen model as can be seen in Fig. 1. So each point is associated to a correspondent side as described in [11]. The sum of the distances from each point to the respective side is the second criteria. Points inside the rectangle are considered to support the current model due to possible multipath reflections and therefore called inliers.

Points outside the bounding box are not confirming the model and rated as outliers. The third important factor is the ratio between the number of outliers and inliers.

To summarize, the resulting area and the distances to the corresponding sides need to be minimized, and as many points as possible should be inside the box.
The error of the estimation is defined as

$$\varphi_{\text{err}} := \varphi_{\text{est}} - \varphi_{\text{GT}}. \quad (1)$$

The following two statistical parameters are applied to analyze the error: the root mean squarer (rms) error and the median absolute error

$$\text{mae} = \text{median} \{ \text{abs} (\varphi_{\text{err}}) \}. \quad (2)$$

The statistical parameters are listed in Table I. The well-known oriented bounding box (OBB) is sensitive to outliers since no compensation is available, which increases the rms error. Outlier points can be compensated by introducing a quality function \((OBB + QF)\) and both the mae and the rms error go down significantly by 47% and 30%, respectively. Applying the enhanced oriented bounding box (EOBB) approach, the implemented outlier treatment reduces the rms error again by 11% and the mae by 10%. The adapted L-fit is by 10% better in the rms error, but the mae is higher by 7% compared to the EOBB. The L-fit has less outliers as can be seen in the smaller rms error compared to the EOBB because it is like the MainDir not limited in the angular rotation steps. In exchange the mae is higher due the fact that the L-fit is challenged by front views of the vehicle. The brute-force algorithm (MainDir) is not restricted to the rotation steps by the convex hull and therefore leads to optimal results. Compared to the EOBB the rms error drops again by 19% and the mae by 12%.

A. Dependency of Aspect Angle

Accurate results can be achieved if two contour sides of the vehicle are visible as they increase the justification of the model. This dependency is shown with the introduced aspect angle as

$$\varphi_{\text{asp}} := \varphi_{\text{GT}} - \varphi_{\text{az}} \quad (3)$$

with the azimuth angle \(\varphi_{\text{az}}\) to the target vehicle.

For \(\varphi_{\text{asp}} = 0^\circ\) only the short contour of the rear view (cf. Fig. 1 (b)) is visible. The radar reflections from multipath propagation do not provide the whole vehicle’s dimensions, hence gaining a valid fit is challenging. This also holds for the front view with \(\varphi_{\text{asp}} = 180^\circ\) where the backscatter area is also smaller due to the front view of the vehicle. For the long side (cf. Fig. 1 (f)) with \(\varphi_{\text{asp}} = 90^\circ\) the errors are smaller compared to the short side. With one large side visible a stable contour alignment is provided. The algorithms yield optimal results if two contours are detected while the car is oncoming (cf. Fig. 1 (a)) with \(\varphi_{\text{asp}} = 135^\circ\) or departing (cf. Fig. 1 (c) and (d)) with \(\varphi_{\text{asp}} = 45^\circ\).

In Fig. 7 the dependency of the aspect angle for the different algorithms is shown. It can be clearly seen that for the front and rear view the errors are higher than for the long side and the L shape, respectively.

### Table I

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Two Sensors</th>
<th>One Sensor</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBB</td>
<td>rms in °</td>
<td>mae in °</td>
</tr>
<tr>
<td>OBB + QF</td>
<td>11.02</td>
<td>4.14</td>
</tr>
<tr>
<td>L-fit</td>
<td>8.83</td>
<td>3.99</td>
</tr>
<tr>
<td>EOBB</td>
<td>9.77</td>
<td>3.73</td>
</tr>
<tr>
<td>MainDir</td>
<td>7.88</td>
<td>3.30</td>
</tr>
</tbody>
</table>

---

Fig. 7. Function of the rms error (a) and the mae (b) from the aspect angle.
B. Dependency of Range

The radar sensor has an azimuth resolution of 1° after applying the linear prediction. This leads to a specific crossrange at a given distance. The farther away a target is, the larger the crossrange grows. A target at the edge of the field of view has therefore less reflections than a nearby target. This has an influence on the orientation estimation as depicted in Fig. 8. If less reflections are available, the estimation gets less reliable which can be seen in the increasing root mean square error.

C. Dependency of Number of Sensors

For target detection two radar sensors are used as mentioned in Section II. The lateral distance between the two sensors leads to different reflection centers due to the different aspect angles. This leads for some measurements to two separate strong reflection centers on the contour which enhances the model fitting as depicted in Fig. 9.

Using only one sensor, the rms error and the mae increase both by 17% for the EOB. Particular the L-fit requires a second strong reflection center, that is why the rms error increases by 32% and the mae by 21%.

D. Dimension Extraction

The presented algorithms allow the extraction of the dimensions of the target vehicle. The presented part of the scene from Fig. 1 is evaluated with respect to the box dimensions in Fig. 10. The target vehicle, a Mercedes-Benz E-class T model, has a length of 4.905 m and a width of 1.854 m. In the presented scene the mean values are 4.86 m and 2.06 m, respectively. This results in a root mean square error of 0.76 m and 0.48 m, respectively.

If both sides or the long side of the target is visible, the box dimensions agree with the actual values, e.g., (a), (c), (d), and (f). If the rear or front view is visible at least the length extraction is distorted, e.g., (b) and (e).

E. Reliability of the Current Fit

Once the model for the actual measurement is found, the question of the reliability of the actual fit arises. The approach to use the dimensions does not provide a valid rating due to the fact that, if only the rear or the front is visible, the error can be small as well, as can be seen in the example in Fig. 1 (b).
Dimensions are extracted with a root mean square error of 0.76 m and 0.48 m, respectively. The results deteriorate by 20% to 30% if only one sensor is used instead of two sensors. Evaluating the trend of the optimization variable resulting from the introduced quality function enables an evaluation of the reliability of the orientation estimation. Using this reliability value the orientation estimation can be enhanced by an integration in a temporal filter in future work.

VI. CONCLUSION

For the orientation estimation of target vehicles in high-resolution radar images three different algorithms are presented using single measurements. Two state of the art approaches of the orientated bounding box (OBB) and the L-fit from literature are adapted. The orientation error is less than 3.73° in 50% of the evaluated measurements applying the enhanced orientated bounding box (EOBB). A root mean square error of 9.77° is achieved, which is a reduction by 38% compared to the well-known orientated bounding box algorithm. The improved outcome is due to the consideration of possible outliers in radar data. Precise target vehicle's dimensions are extracted with a root mean square error of 0.76 m and 0.48 m, respectively. The results deteriorate by 20% to 30% if only one sensor is used instead of two sensors. Evaluating the trend of the optimization variable resulting from the introduced quality function enables an evaluation of the reliability of the orientation estimation. Using this reliability value the orientation estimation can be enhanced by an integration in a temporal filter in future work.

REFERENCES

Fabian Roos (S’15) received the M.Sc. degree from the Karlsruhe Institute of Technology (KIT), formerly Universität Karlsruhe (TH), in 2014. Since September 2014 he is working towards the Ph.D. degree as a Research Assistant at the Institute of Microwave Engineering at Ulm University, Germany. He is interested in automotive radar signal processing, especially for chirp sequence radar. Currently he is working on adaptivity for chirp sequence radar.

Dominik Kellner (S’15) was born in Munich, Germany in 1985. He received the Dipl.-Ing. degree (equivalent to M.Sc. degree) in mechatronics and information technology from TU Munich, Germany, in 2010. He is currently working towards the Ph.D. degree at Ulm University, Germany. His research interests include automotive radar and environmental perception with a focus on motion state estimation of extended objects. Further research topics are radar alignment and ego motion state estimation using high resolution Doppler radar.

Jürgen Dickmann is Manager Active Sensors at DAIMLER Research and Advanced Engineering. He is responsible for the development of active Sensors (Radar and Scanner) and environmental perception in driver assistance- and active safety systems. His present focus is in sensor-technology like Radar, Lidar and Laser Scanner technology for highly automated functions/self-driving cars. He received his Diploma degree in electrical engineering from University Duisburg in 1984 and did his Dr.-Ing. degree in 1991 from the Rheinisch Westfälische Technische Hochschule Aachen (RWTH), Germany. In 1992 he was awardee from VDE/ITG. In 1986 he started his career at AEG Research Center Ulm, where he conducted research on III/V processing techniques, mm-Wave devices and mm-MMICs. Between 2005 and 2009 he was in charge of teams developing sensor technologies, sensor fusion- and situation analysis concepts. In 2015 he was also in charge of the transfer into series for the complete environmental perception including all sensors and navigation/localization components for the next S-Class.

Christian Waldschmidt received the Dipl.-Ing. (M.S.E.E.) and the Dr.-Ing. (Ph.D.E.E.) degrees from the University Karlsruhe (TH), Karlsruhe, Germany, in 2001 and 2004, respectively. From 2001 to 2004 he was a Research Assistant at the Institut für Höchstfrequenztechnik und Elektronik (IHE), Universität Karlsruhe (TH), Germany. Since 2004 he has been with Robert Bosch GmbH, in the business units Corporate Research and Chassis Systems. He was heading different research and development teams in microwave engineering, rf-sensing, and automotive radar. In 2013 Christian Waldschmidt returned to academia. He was appointed as Director of the Institute of Microwave Engineering at Ulm University, Germany, as full professor. The research topics focus on radar and rf-sensing, mm-wave and submillimeter-wave engineering, antennas and antenna arrays, MIMO, rf and array signal processing.

He authored or coauthored over 100 scientific publications and more than 20 patents. Additionally, he is vice chair of IEEE MTT-27 Technical Committee (wireless enabled automotive and vehicular applications), executive committee board member of the German MTT/AP joint chapter, and member of the ITG committee Microwave Engineering (VDE). He is a reviewer for multiple IEEE transactions and letters. In 2015, he served as TPC chair for IEEE MTT-S International Conference on Microwaves for Intelligent Mobility and as associate guest editor for IEEE Transactions on Microwave Theory and Techniques.