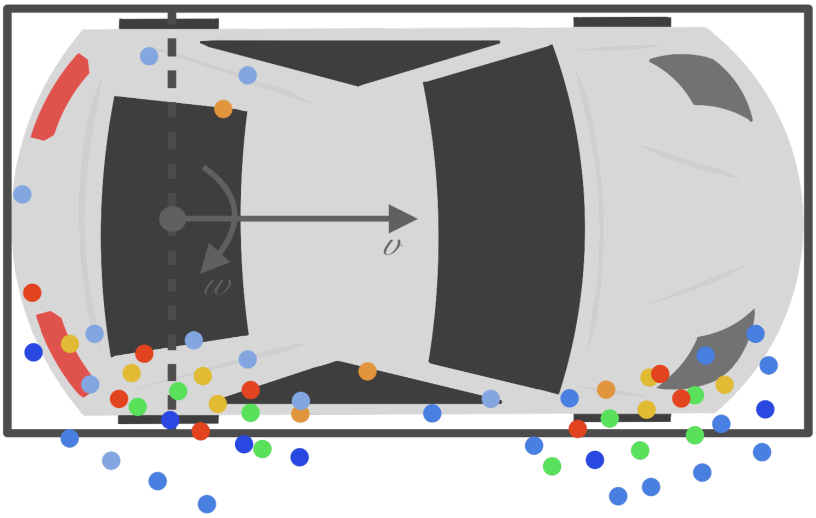


Instantaneous Full Motion Estimation Using Low Level Radar Measurements

Motion, Orientation and Size Estimation of an Extended Object Using One Automotive Radar in a Single Radar Frame

Master's Thesis in Systems, Control and Mechatronics

Karl-Oskar Gunnarsson
Jakob Nilsson

Department of Electrical Engineering
CHALMERS UNIVERSITY OF TECHNOLOGY
Gothenburg, Sweden 2020

Master's thesis 2020

Instantaneous Full Motion Estimation Using Low Level Radar Measurements

Motion, Orientation and Size Estimation of an Extended Object Using One Automotive Radar in a Single Radar Frame

KARL-OSKAR GUNNARSSON
JAKOB NILSSON

CHALMERS
UNIVERSITY OF TECHNOLOGY

Department of Electrical Engineering
Division of Signal Processing
Chalmers University of Technology
Gothenburg, Sweden 2020

Instantaneous Full Motion Estimation Using Low Level Radar Measurements
Motion, Orientation and Size Estimation of an Extended Object Using One Automotive Radar in a Single Radar Frame
KARL-OSKAR GUNNARSSON
JAKOB NILSSON

Supervisor: Niclas Carlström, Aptiv
Examiner: Tomas McKelvey, Department of Electrical Engineering

Master's Thesis 2020
Department of Electrical Engineering
Division of System and Control
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Cover: A drawing showing the beam vectors from one radar frame on a passenger car. A bounding box and the estimated rear axle is also shown with a velocity and a yaw rate.

Typeset in LaTeX
Gothenburg, Sweden 2020

Instantaneous Full Motion Estimation Using Low Level Radar Measurements
Motion, Orientation and Size Estimation of an Extended Object Using One Automotive Radar in a Single Radar Frame
KARL-OSKAR GUNNARSSON
JAKOB NILSSON
Department of Electrical Engineering
Chalmers University of Technology

Abstract

The objective of the thesis was to develop an algorithm that can estimate the size, orientation, velocity and yaw-rate of an extended object surrounding a host vehicle by using the available information in the Radar Data Cube (RDC) from a single radar in a single radar frame. The RDC was clustered using DBSCAN to sort out the scatter points belonging to the target. Then a RANSAC algorithm was used to estimate a bounding box of the target. Further it was assumed that the target vehicle could be modeled according to the Ackermann steering condition. The position of the rear axle was estimated using the assumption that the density of scatter points is higher around the wheels than on the rest of the side of a car.

The yaw rate was estimated by first calculating the Instantaneous Center of Rotation (ICR) line. The ICR line intersects the origin of the sensor and has a slope that can be calculated by solving a system of equations for two randomly sampled scatter points. This was done multiple times and the median of the results was used as an estimation of the slope. An ICR point was then found by computing the point where the ICR line intersects with a line coincident with the rear axle. The yaw rate could then be calculated using the slope of the ICR line and the position of the ICR point. Using the estimated yaw rate and the distance from the center of the rear axle to the ICR point, the instant velocity was estimated. To show that it would be possible to estimate the motion of a target instantaneously if there was less noise in the measurements, the results was modeled with a Coordinated Turn model and a Cubature Kalman filter was developed and applied.

It was shown that it is possible to compute the full motion, orientation and size instantaneously with the proposed model. The result was however considerably noisy. This noisy behavior was mitigated with the Kalman filter which resulted in an output with more stable estimates.

Keywords: Radar, Full Motion Estimation, Instantaneous Yaw Rate Estimation, Radar Data Cube

Acknowledgements

Throughout this project we have received help and support from many people, without whom this thesis would never have been written. First of all we would like to thank our supervisor Niclas Carlström for continuously guiding us. We would also like to thank our examiner Tomas McKelvey for valuable input, Alexander Lyckell for helpful advice and Zbigniew Pierzchala for providing logs from test drives. Lastly we would like to thank Aptiv for giving us the opportunity to do this project, and for providing office space and all materials needed.

Karl-Oskar Gunnarsson & Jakob Nilsson, Gothenburg, June 2020

Contents

List of Figures

List of Tables

List of Abbreviations

CFAR	Constant False Alarm Rate	xi, 10, 11, 17, 29
CKF	Cubature Kalman Filter	25, 26, 31, 32, 38, 39
CT	Coordinated Turn	27, 41
CUT	Cell Under Test	10, 11
DBSCAN	Density-Based Spatial Clustering of Applications with Noise	xi, 18, 19, 23, 38
FFT	Fast Fourier Transform	7, 8
ICR	Instantaneous Center of Rotation	xi, 22, 24, 25, 39
IF	Intermediate Frequency	5–8
LFMCW	Linear Frequency Modulated Continuous Wave	5, 6, 9
LRR	Long Range Radar	1
MRR	Mid Range Radar	1
MSE	Mean Square Error	xiii, 33–35, 38, 39
RANSAC	Random Sample Consensus	xi, 20–24, 37, 39
RDC	Radar Data Cube	xi–xiii, 1, 2, 5, 8–11, 17, 20, 29–32, 37–39
SRR	Short Range Radar	1, 6
SRR3	Short Range Radar Generation 3	1, 10, 11, 17, 27

Nomenclature

Aptiv A company that manufactures radars and develops radar trackers and feature functions for the automotive industry. . . . 1, 3

Azimuth The angle in the horizontal plane between the boresight of the radar and an object detected by a radar. xi, 1, 7–11, 13, 17–19, 24

Beam vector One bin in a range-Doppler map, which is described by a vector of complex numbers that is as long as the number of antennas on the radar . . . . xi, 8, 10, 11, 23, 37–39, 41

Bin One element in a range-Doppler map . 8, 10, 11, 18, 19, 29, 38

Boresight The axis in the center of a radar's field of view. . xi, xvii, 7, 9, 10, 12

Bounding box A rectangle that encapsulates an object. It has the length, width and pointing of the actual object. . . xi, 16, 17, 22–24, 38, 39

Chirp A frequency sweep from a radar. 5–7

Detection A peak in the range-Doppler map that is the usual output from a radar . . . . xi–xiii, 2, 5, 7, 10–16, 18, 29–32, 37–39, 41

Dwell time The time it takes to perform a chirp. 5

Elevation The angle in the vertical plane to a detected object. . . . 1, 9

Extended object An object that has more then one radar detection related to it. 2, 5, 11, 12, 41

Frame One radar measurement. One frame consist of several chirps. . xii, 2, 5, 6, 8, 17, 21, 23–25, 27, 31, 32, 37–39, 41

Heading The angle between the boresight of the sensor and the resultant of the velocity vectors of an object. . . . . . . xii, 5, 12, 32, 38

Kalman filter Algorithm used for estimating the states of a system given estimates of the states in previous time steps, measurements in the current time step, a measurement model, a motion model and noise or other uncertainties in the models. . 17, 26, 28, 33, 34, 39, 41

Pointing The angle between the boresight of the sensor and the direction of the bounding box. . . . xi–xiii, 11, 22–25, 27, 32, 33, 38, 39

Radar Detection system that uses radio waves to determine the range, angle, and velocity of objects . xi, xvii, 1, 2, 5, 6, 8–11, 14, 17, 23–25, 27, 29, 37, 39, 41

Range-Doppler map A two dimensional matrix showing the measured energy in each range and Doppler bin of a radar. . xi, xvii, 1, 2, 8, 10, 11, 18

Reference point A point on a vehicle where there is only linear speed and no speed caused by rotation of the vehicle. . 22, 24, 27, 28, 33, 38

Scatter point A point on a target where a radar signal is reflected. 17, 20, 24, 25, 38, 39

1
Introduction

The automotive industry is currently moving towards more autonomy and future cars might not require human drivers. There are still some crucial obstacles to overcome before vehicles become totally autonomous but many of the cars on the market today have some kind of autonomy or pilot assist systems.

1.1 Background

A crucial part of making cars more autonomous is to detect and classify surrounding objects and to decide their size, position, velocity and yaw rate. This is often done using radars, cameras and/or other sensors [1]. One of the companies developing radars and tracker schemes in the automotive industry is Aptiv. There are several different kinds of automotive radars that have various functions. Within Aptiv these include Short Range Radar (SRR), Mid Range Radar (MRR) and Long Range Radar (LRR). They have different fields of view and ranges. The type of radar used in this thesis is a Short Range Radar Generation 3 (SRR3). A radar operates by transmitting a radio signal and then receiving it again by antennas. When multiple antennas are used, both the range and the angles (azimuth and elevation) to what is measured can be computed by using the phase shift of the radio signal between the antennas. Usually an array of antennas are used to accurately compute these characteristics of the signal. The data that the antennas receive can be visualized in a range-Doppler map and by stacking the range-Doppler maps for all of the antennas in the radar one gets an RDC which contains all relevant low level information from the radar. An example of a range-Doppler map can be seen in Figure 1.1.

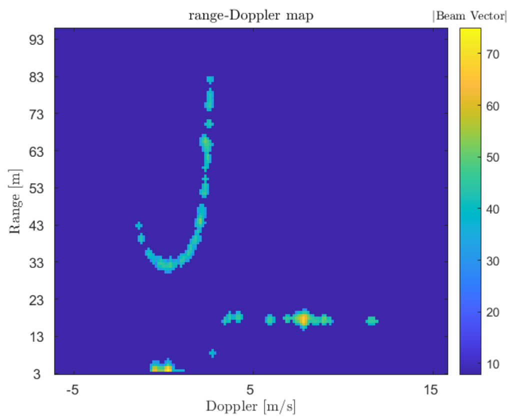


Figure 1.1: An example of a range-Doppler map with the amplitude of the radar signal.

Most automotive radars today use the information from the RDC to compute detections and only these detections are the output from the sensor. The reason for only giving the computed detections as output is that the RDC contains too much data for the on board processors to efficiently compute. The most common way of choosing detections is to take the peaks in the range-Doppler maps, that is the yellow areas in Figure 1.1. The rest of the RDC is filtered out by setting a threshold on what is to be categorized as noise. The detections are then grouped together (clustered) in some sort of object tracker to estimate the position, size and orientation of objects like cars, trucks and pedestrians. One distinction between modern automotive radars and classical airborne radars is that the targets are assumed to be extended objects in relation to point sources because the resolution of the radars are finer then the physical extent of the object. Hence, more than one detection per extended object are expected to be generated and the size and motion of objects can then be estimated.

In order to more accurately compute the size and orientation of the objects one could investigate if there is more useful information in the RDC than the detections that are the input to the common automotive radar trackers today [2]. In the case of the radar output in Figure 1.1, it is the information around the yellow areas that are categorized as noise in the radars used today.

1.2 Purpose

The purpose of the thesis is to investigate alternative tracking schemes than what is common today in an automotive setting by utilizing more of the data from the radar. This in order to develop better active safety functionalities which depend on an accurate description of the motion and size of surrounding extended objects.

1.3 Objective

The objective of the thesis is to develop an algorithm that can accurately estimate the size, orientation, velocity and yaw-rate of an extended object surrounding a host vehicle using more of the available information in the RDC than just the detections in one radar frame.

1.4 Scope

The data used in the project is the output from only one radar and hence no fusion between different radars or other sensors are considered. The object that is estimated is a passenger car which means that the result is not necessarily transferable to other kinds of objects, such as bicyclists and pedestrians, since the object is modelled as a rectangle with a motion as described by the Ackermann model. The thesis is also limited to post processing of collected data, meaning that the resulting algorithms are not intended to be able to function in real time in a vehicle.

When benchmarking the performance of the developed algorithms, the position, velocity and yaw rate of the target are compared to GPS data as ground truth. When evaluating the algorithms no new data is gathered. An existing log from a previous test drive at Aptiv is used since the data is gathered from a test facility where there are an accurate GPS system to track vehicles. The size estimates are compared to the actual size of the target given by the manufacturer of the car model used in the test drive.

2
Theory

In order to achieve the objective of the full motion estimation, the radar data had to be understood and extracted. Fundamental radar concepts are therefore described in Section 2.1 and the difference between radar detections and low level measurements in the RDC is detailed further in Section 2.2. From these radar measurements correlating to one extended object a crude method to compute the heading and velocity is described. Lastly, a model to simulate radar output in a single time instance from the motion of an extended object is proposed. Which was used to develop and verify the methods presented in this thesis. In the model, assumptions regarding the distribution of the radar measurements along the edges of an extended object are also explained.

2.1 Radars

A radar device fundamentally transmits electromagnetic waves that propagate through space, reflect on distant objects and are then received again by antennas on the radar device. Depending on the intended usage of the radar, the characteristics, such as frequency, bandwidth and dwell time, of the transmitted electromagnetic waves are different. The radar used to collect data in this thesis is a Linear Frequency Modulated Continuous Wave (LFMCW) radar. These sensors are capable of measuring the range, range rate and angles of a reflected wave. LFMCW radars transmit a linear frequency sweep called a chirp. The chirp is being transmitted on top of a carrier frequency, f_c. The difference between the lowest and highest frequency of a chirp is the bandwidth and the time it takes to perform a chirp is called dwell time. Several succeeding chirps are being transmitted during one radar frame [3].

2.1.1 Range Measurements

When the transmitted signal reflects on objects, a copy of the signal is being received by the antenna. Since it takes some time, τ_δ, for the signal to propagate from the radar to the object and back, the received signal will be shifted in time compared to the transmitted signal. By mixing the transmitted signal with the received signal the Intermediate Frequency (IF)-signal is obtained. The IF-signal is a new signal with a frequency equal to the difference between the transmitted and received frequencies in each time instance. The phase shift of the IF-signal will correspondingly be equal to the difference in phase shift between the transmitted and received signal. Since the chirps varies in frequency over time, the resulting IF-signals is different depending on the distance to the reflecting object. The distance can therefore be calculated from the IF-signal [3].

The carrier frequency can be different between radar frames to extract more information from the environment. An SRR can have both a frequency that works best at a short range and one that functions better on a mid range. These different frequencies generates different range-resolutions over different radar frames and are called look types.

2.1.2 Range Rate Measurements

If the object that the signal is reflected on has a velocity towards or away from the radar sensor, the IF-signal will change slightly between each chirp in a frame. For reasonable velocities the difference in frequency between chirps will be smaller than the range resolution. The phase, on the other hand, will change even at low velocities since the wavelength of the signal and the change in range between chirps is of the same order of magnitude. By using the phase difference between the IF-signals of each chirp, the time between the chirps and the wavelength of the radar wave, the radial velocity of the object relative to the radar can be calculated.

2.1.2.1 Ambiguity in Range Rate Measurements

Since sinusoidal waves are ambiguous, the phase shift between the chirps can be the same for different radial velocities. Hence there is a maximum radial speed that can be measured unambiguously by an LFMCW radar. If, for example, the radial displacement of an object that reflects a radar signal between two chirps is exactly half the wavelength of the radar wave, there is no way of knowing whether the object is moving towards the radar or away from it since the phase difference is the same in both cases. Since there is ambiguity in the range rate of objects it is assumed that the speed is low enough for the displacement between chirps to be less than half the wavelength of the radar wave. Hence the maximum radial velocity that an LFMCW radar can measure in a frame is a function of the wavelength and the time between the chirps, τ_s [3]. There are, however, methods for dealiasing the range rate, and in that way unambiguously measure radial velocities greater than this maximum. One such method is the one referenced in [4]. In this project, all radial velocities are assumed to be below the maximum radial velocity that the radar can measure unambiguously without dealiasing.

2.1.2.2 Range Rate Compensated for Velocity of the Sensor

Since the radar is mounted on a host vehicle that might be moving, the measurements will not reflect the actual velocities over the ground of detected targets, but the velocities relative to the host. Assuming that the velocity and curvature of the host vehicle, as well as the position of the sensor on the host, are known, the velocity of the sensor can be calculated and compensated for in order to get the over ground velocities of targets. When the over ground velocity is found the stationary objects can be sorted out. In Figure 2.1, an example of the position of a sensor on a host vehicle can be seen.

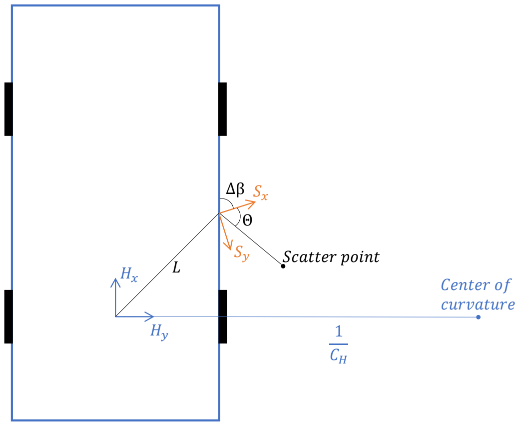


Figure 2.1: Position of a sensor on a host.

In Equation (2.1) the velocity of the sensor is calculated, given the position in Figure 2.1. The superscripts in brackets of the variables in Equation (2.1) and the following equations in this thesis indicates in which coordinate system the variable is defined. The subscripts indicate what object the variable belongs to. H is short for Host, S for Sensor, T for Target and D for Detection. Hence, $\mathbf{v}_S^{[S]}$ is the velocity of the sensor relative to the coordinate system of the sensor. $v_{S_x}^{[S]}$ and $v_{S_y}^{[S]}$ are the components in the direction of the x- and y-axis respectively. Further, $\Delta\beta$ is the boresight of the sensor. That is the angle between the x-axis in the host vehicle coordinate system and the x-axis in the sensor coordinate system. $\mathbf{v}_H^{[H]}$ is the velocity of the host vehicle in the host coordinate system, and $v_{H_x}^{[H]}$ and $v_{H_y}^{[H]}$ are the velocity components in the direction of the x- and y-axis respectively. C_H is the curvature of the target and $L_{S_x}^{[H]}$ and $L_{S_y}^{[H]}$ are the x- and y-coordinates of the origin of the sensor coordinate system in the host coordinate system.

$$\mathbf{v}_S^{[S]} = \begin{bmatrix} v_{S_x}^{[S]} \\ v_{S_y}^{[S]} \end{bmatrix} = \begin{bmatrix} \cos(\Delta\beta) & \sin(\Delta\beta) \\ -\sin(\Delta\beta) & \cos(\Delta\beta) \end{bmatrix} \begin{bmatrix} v_{H_x}^{[H]} \\ v_{H_y}^{[H]} \end{bmatrix} + \underbrace{C_H \left\| \mathbf{v}_H^{[H]} \right\|}_{\omega_H} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} L_{S_x}^{[H]} \\ L_{S_y}^{[H]} \end{bmatrix} \quad (2.1)$$

Using Equation (2.2), where θ is the azimuth, the compensated range rate, $\dot{r}_{comp}$, of a detection can then be calculated. In Equation (2.2), $\dot{r}$ is the range rate measured by the sensor.

$$\dot{r}_{comp} = \dot{r} + \left\| \mathbf{v}_S^{[S]} \right\| \cos\left(\theta - \tan^{-1}\left(\frac{v_{S_y}^{[S]}}{v_{S_x}^{[S]}} \right) \right) \quad (2.2)$$

2.1.3 Range-Doppler Map

When the IF-signals for each chirp is generated, they are low-pass filtered and then a Fast Fourier Transform (FFT) is performed over the IF-signals. This generates peaks in the IF-signal at frequencies corresponding to ranges of measured objects. A second FFT is then performed over the phases of the first FFT. This generates peaks at the phase shifts of the IF-signals corresponding to the range rates of the objects at the corresponding ranges. After performing these FFT:s, the results can be arranged in a two dimensional matrix called the range-Doppler map. An example of a range-Doppler map can be seen in Figure 2.2. Each element in the matrix is a bin pair that represent the energy in a Doppler bin and a range bin. Bin pairs without any objects are blue in the range-Doppler map and bin pairs where objects are measured are in other colors depending on how high the magnitude peak at that bin pair is. The range-Doppler map in Figure 2.2 shows the bin measurements multiplied with the corresponding range and Doppler resolutions. The resolution depends on the length of the FFT and the sampling frequency.

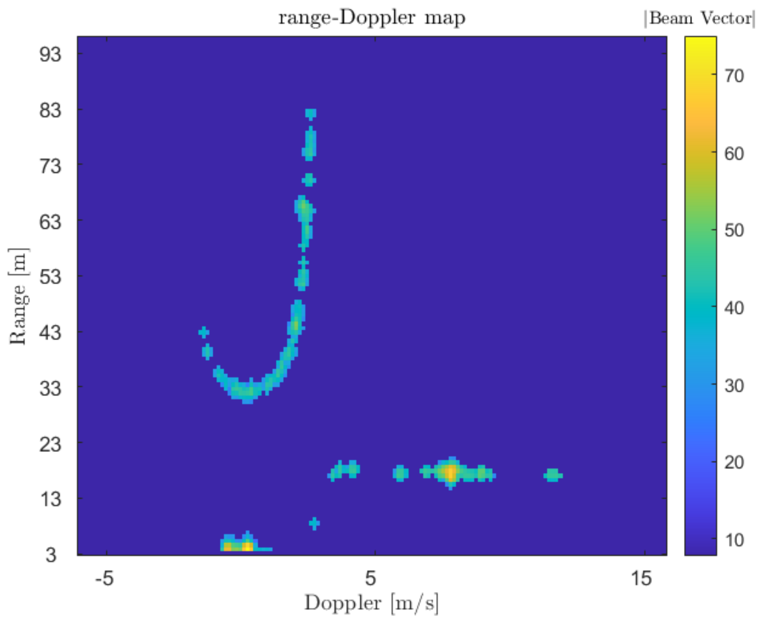


Figure 2.2: A range-Doppler map over moving target and host.

2.1.4 Radar Data Cube

In each radar frame one range-Doppler map, such as the one in Figure 2.2, for each antenna is generated. These can be represented as a three dimensional matrix, a data cube shown in Figure 2.3. Each element in the RDC is a complex valued number and the bin over all antennas in one frame is called a beam vector. The beam vector's complex valued numbers are used to compute the azimuth as explained in Section 2.1.5. Throughout this thesis the amplitude in each bin is computed as the norm of the beam vector.

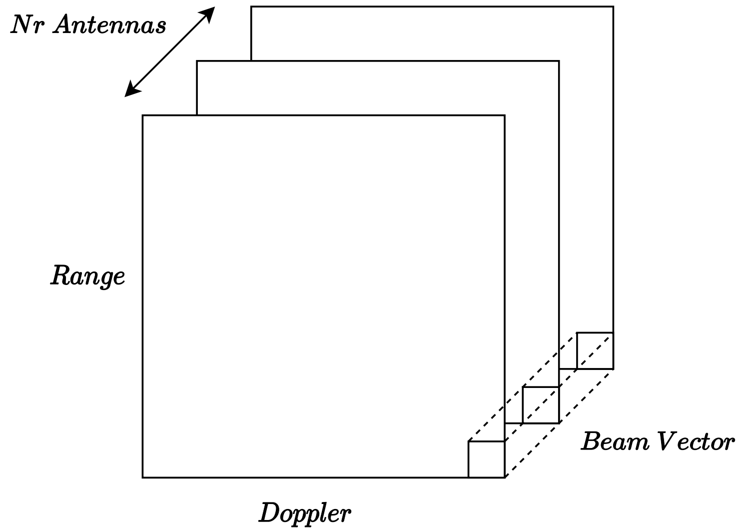


Figure 2.3: A visual representation of the RDC and a beam vector.

2.1.5 Angle Measurements

In order to measure the angles to objects, LFMCW radars have multiple antennas placed in rows and columns on the same plane. If an object that reflects a signal back to the antennas is not perpendicular to the radar plane, the different antennas do not receive the returning wave at exactly the same time. This leads to a phase shift between the different antennas' received signals. This phase shift can be used to calculate at what angle the object is relative to the radar. The phase shifts from antennas in vertical columns are used to calculate the elevation angle of the object relative to the radar. The phase shifts from antennas in horizontal rows are used to calculate the object's angle in the horizontal plane. A coordinate system is placed with its origin in the middle of the arrays of antennas. The x-axis is in what is called the boresight of the radar. The boresight is a line orthogonal to the plane of antennas. For most radars the boresight is in the the center of the radar's field of view which is also the case for the radar used in this thesis. When defining the angles to objects in the radar's field of view, the azimuth is the angle between the objects position and the boresight in the xy-plane and the elevation angle is the angle in the xz-plane [5]. The angles are visualized in Figure 2.4. In this project only the azimuth, and not the elevation angle, is considered.

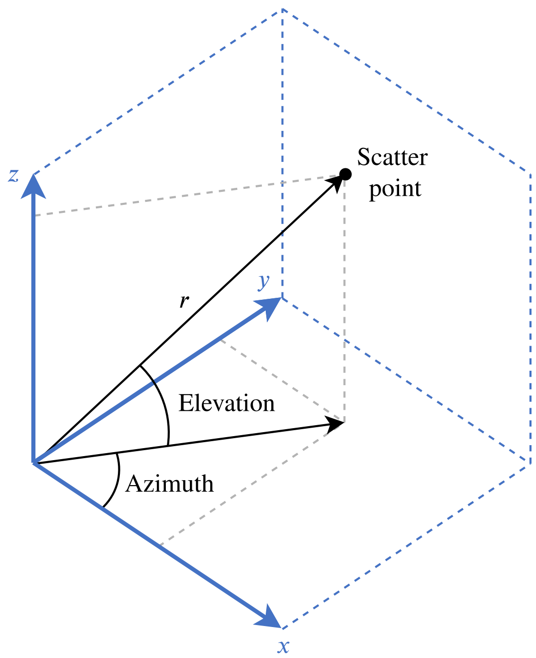


Figure 2.4: A visualization of the azimuth and elevation angles. The boresight of the radar is in the x-axis.

The SRR3 has an array of three antennas and hence the received signal per time instance is an array of three elements with complex numbers. In order to extract the angle from these complex numbers the phase difference between a narrow and wide antenna pair is computed and compared for each beam vector to remove any phase ambiguities [5].

2.2 Radar Signal Filtering

When the range-Doppler map is constructed the next challenge is to determine which information is the most relevant in order to find peaks in the data that could correlate to targets. This is normally done in a few different filtering steps that result in what is called detections. The first step is a type of CFAR, which is followed by subsequent thresholds on the received signal magnitude which result in first pass detections and then actual detections. All information above CFAR will however be considered in this thesis as useful.

2.2.1 CFAR-filtered RDC

A CFAR is used to filter out the noise from the received radar signal in the RDC. The CFAR threshold of what is considered to be noise and what is relevant data is different in different scenarios and ranges and is therefore dynamically computed for each data set and range bin. One way to compute CFAR for a given range bin is to use the average of the signal power of $2N_{cfar}$ range bins around a Cell Under Test (CUT), as shown in Figure 2.5. This is called a cell averaging CFAR. The signal power for each cell is computed as the norm of the corresponding beam vector. In this way the general noise level is mitigated and only the relevant information in the data is left. If CFAR is larger than the signal power in the CUT, then it is considered to be noise [3].

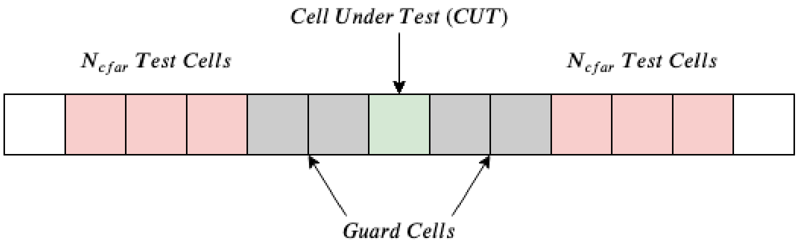


Figure 2.5: Computing CFAR with the sliding window approach over range bins.

2.2.2 Detections

From the CFAR-filtered range-Doppler map, detections can be computed. Detections can be chosen in various ways but the most common method is to take the peaks in the resulting CFAR-filtered range-Doppler map where a few adjacent bins around the peak are used to compute one detection. The detection does not necessarily represent one bin measurement since it is also common to interpolate between bins to increase the resolution of the detections. Hence, the number of detections is smaller than the number of bin measurements in the CFAR-filtered RDC. This result in detections with a range, azimuth, range rate and amplitude. Exactly how the detections are computed and which type of CFAR that is used in the SRR3 cannot be disclosed since it is a central part of Aptiv's software. In this thesis the detections used are the ones computed by the radar and the CFAR threshold used is also an output from the radar.

2.3 Velocity and Heading of Extended Objects

When a moving extended object is in the field of view of a radar, the radar signal reflects off of it and it is shown in the range-Doppler map as a cluster of points. This can be seen in Figure 2.2 as a collection of bins with high beam vector amplitude with a compensated range rate that is non-zero. If a moving extended object cause N bin measurements, the relations between these and the velocity of the vehicle can be expressed as in Equation (2.3). This relation will hold to compute the velocity and pointing of an object accurately if it is assumed that the extended object has a linear movement and the calculated velocity will thus be less accurate the more yaw rate an object has. The matrix is ill-conditioned if the azimuth spread is small as well and hence sensitive for noise in the compensated range rate [5]. In Equation (2.3) $\dot{r}_{comp}$ is the compensated range rate, θ is the azimuth angle and $v_{T_x}^{[S]}$ along with $v_{T_y}^{[S]}$ are the velocities of the target object in the sensor's coordinate system.

$$\begin{bmatrix} \dot{r}_{comp,1} \\ \vdots \\ \dot{r}_{comp,N} \end{bmatrix} = \begin{bmatrix} \cos(\theta_1) & \sin(\theta_1) \\ \vdots & \vdots \\ \cos(\theta_N) & \sin(\theta_N) \end{bmatrix} \begin{bmatrix} v_{T_x}^{[S]} \\ v_{T_y}^{[S]} \end{bmatrix} \tag{2.3}$$

The total velocity of the extended object is then given by Equation (2.4), and the heading of the target vehicle in the sensor's coordinate system, $\alpha_T^{[S]}$, is given by Equation (2.5) where $\Delta\beta$ is the boresight of the sensor.

$$\|\mathbf{v}_T\| = \sqrt{\left(v_{T_x}^{[S]}\right)^2 + \left(v_{T_y}^{[S]}\right)^2} \tag{2.4}$$

$$\alpha_T^{[S]} = \tan^{-1}\left(\frac{v_{T_y}^{[S]}}{v_{T_x}^{[S]}}\right) + \Delta\beta \tag{2.5}$$

2.4 Simulated Radar Measurements

To develop the algorithms used for estimating size, velocity and yaw rate, a theoretical model of a system containing a host vehicle with a sensor and one or more target vehicles was created. In the model, the host, sensor and targets are given velocities, curvatures and sizes when they are generated. Then detection points on the edge of the target can be generated. The velocities of these points depend on the velocity and the curvature of the target as well as on where the points are located relative to the center of rotation of the target vehicle. Equation (2.6) is used to calculate the velocity of the points. $\mathbf{v}_D^{[T]}$ is the velocity of a detection in the coordinate system of the target. $v_{D_x}^{[T]}$ and $v_{D_x}^{[T]}$ are the velocities in the x- and y-direction respectively. $v_{T_x}^{[T]}$ and $v_{T_x}^{[T]}$ are the x- and y-components of the velocity of the target vehicle in the coordinate system of the target. C_T is the curvature of the target. $L_{D_x}^{[T]}$ and $L_{D_y}^{[T]}$ are the x- and y-coordinates of the detection point in the target vehicle coordinate system.

$$\mathbf{v}_D^{[T]} = \begin{bmatrix} v_{D_x}^{[T]} \\ v_{D_y}^{[T]} \end{bmatrix} = \begin{bmatrix} v_{T_x}^{[T]} \\ v_{T_y}^{[T]} \end{bmatrix} + \underbrace{C_T \left\|\mathbf{v}_T^{[T]}\right\|}_{\omega_T} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} L_{D_x}^{[T]} \\ L_{D_y}^{[T]} \end{bmatrix} \tag{2.6}$$

To calculate what the sensor would measure if it detected the generated points, the velocities of the detections need to be calculated in the coordinate system of the sensor. This is done using Equation (2.7), where $\mathbf{v}_D^{[S]}$ is the velocity of the detection in the coordinate system of the sensor. $v_{D_x}^{[S]}$ and $v_{D_y}^{[S]}$ are the components in the x- and y-direction. $\Delta\gamma$ is the angle between the x-axis of the target coordinate system and the sensor coordinate system. $v_{D_x}^{[T]}$ and $v_{D_y}^{[T]}$ are the x- and y-components of the detection's velocity in the target coordinate system.

$$\mathbf{v}_D^{[S]} = \begin{bmatrix} v_{D_x}^{[S]} \\ v_{D_y}^{[S]} \end{bmatrix} = \begin{bmatrix} \cos(\Delta\gamma) & \sin(\Delta\gamma) \\ -\sin(\Delta\gamma) & \cos(\Delta\gamma) \end{bmatrix} \begin{bmatrix} v_{D_x}^{[T]} \\ v_{D_y}^{[T]} \end{bmatrix} \tag{2.7}$$

Equation (2.8) is then used to calculate the azimuth. θ is the azimuth of a detection in the coordinate system of the sensor. In this equation $p^{[S]}_{D_x}$ and $p^{[S]}_{D_x}$ are the x- and y-coordinates of a detection in the coordinate system of the sensor.

$$\theta = \tan^{-1}\left(\frac{p^{[S]}_{D_y}}{p^{[S]}_{D_x}}\right) \tag{2.8}$$

Equation (2.9) is used to calculate the range to a detection from the sensor. Just like in the previous equation, $p^{[S]}_{D_x}$ and $p^{[S]}_{D_x}$ are the x- and y-coordinates of a detection in the coordinate system of the sensor.

$$r = \sqrt{\left(p^{[S]}_{D_x}\right)^2 + \left(p^{[S]}_{D_y}\right)^2} \tag{2.9}$$

Equation (2.10) is used to calculate the range rate of a detection measured by the sensor. $v^{[S]}_{D_x}$ and $v^{[S]}_{D_x}$ are the x- and y-components of a detection's velocity in the coordinate system of the sensor. θ is the azimuth of the detection relative to the coordinate system of the sensor.

$$\dot{r} = v^{[S]}_{D_x} \cos(\theta) + v^{[S]}_{D_y} \sin(\theta) \tag{2.10}$$

A Matlab app was designed in order to visualize the generated scenario. An example of a visualization from the app can be seen in Figure 2.6, where a host, a sensor and a target have been generated. On the target 100 detections points have been added and the arrows represent the measured range rate from each detection.

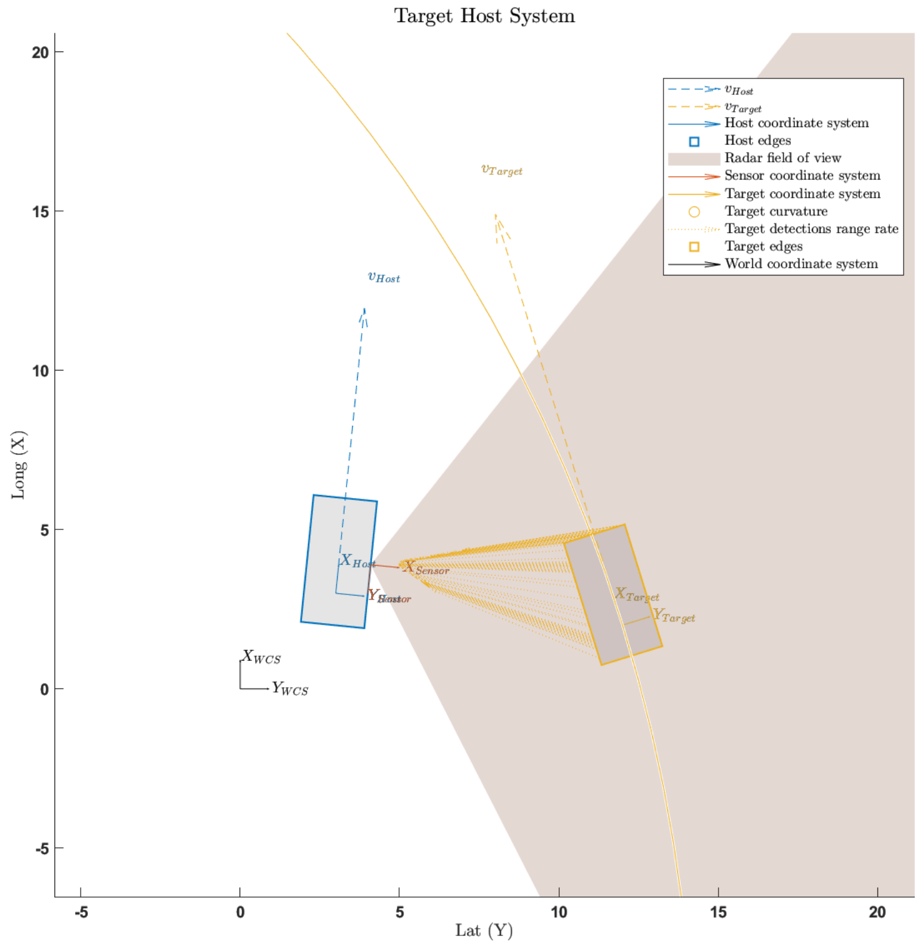


Figure 2.6: An example of generated data. The arrows from the detection points represent the range rates measured by the sensor on host.

2.4.1 Model for Generating Detections

The positions of the detections on the vehicle is placed where the highest probability of a detection in a given scenario is. The model used for this is the one in [6]. Here it is assumed that the majority of detections are generated on the corners of the vehicle and in the wheelhouses as shown in Figure 2.7, where the grey circle sectors represent the most plausible position of detection points and the direction in which the sensor has to be located in order for a detection to be generated. These places are the most plausible detection points because these are complex geometries with at least one area perpendicular to the radar which is a characteristic of a good reflector. There are more possible areas on a car that the propagating radar wave can reflect on, such as door frames and various edges. This is simulated by a small underlying uniform distribution of detections apart from the generated detections from the proposed model which can be seen in Figure 2.6.

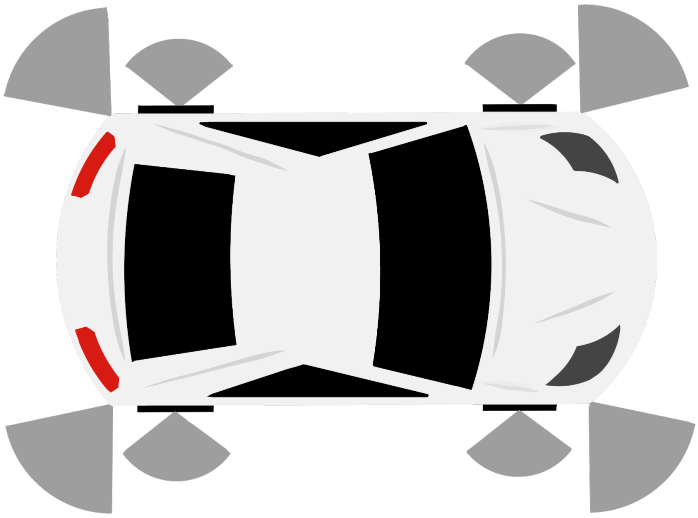


Figure 2.7: A plan view of the detection model for a passenger car. The gray circular sectors represent the most plausible detection points and in what directions the car has to be illuminated in order for a detection to be generated.

2.4.2 Measurement Noise and Outliers

In reality, the measured data is never as accurate as generated data, since there will always be some measurement noise. To make the generated data more realistic, measurement noise can be added to the calculated measurements. Equation (2.8) and (2.9) are then rewritten as (2.11) and (2.12).

$$\theta = \tan^{-1}\left(\frac{p_{D_y}^{[S]}}{p_{D_x}^{[S]}}\right) + n, \quad \text{where } n \sim \mathcal{N}\left(0,\, \sigma_\theta^2\right) \tag{2.11}$$

$$r = \sqrt{\left(p_{D_x}^{[S]}\right)^2 + \left(p_{D_y}^{[S]}\right)^2} + n, \quad \text{where } n \sim \mathcal{N}\left(0,\, \sigma_r^2\right) \tag{2.12}$$

There can also be correctly measured detections that have range rates that look completely wrong. For example if there is a detection on the wheel of a target vehicle, the range rate might differ greatly from the rest of the measured range rates on that target. These detections are called outliers. To make the generated data even more realistic an option to add outliers is available in the theoretical model. In Figure 2.8 the same scenario as in Figure 2.6 is visualized, but this time with noise and outliers added.

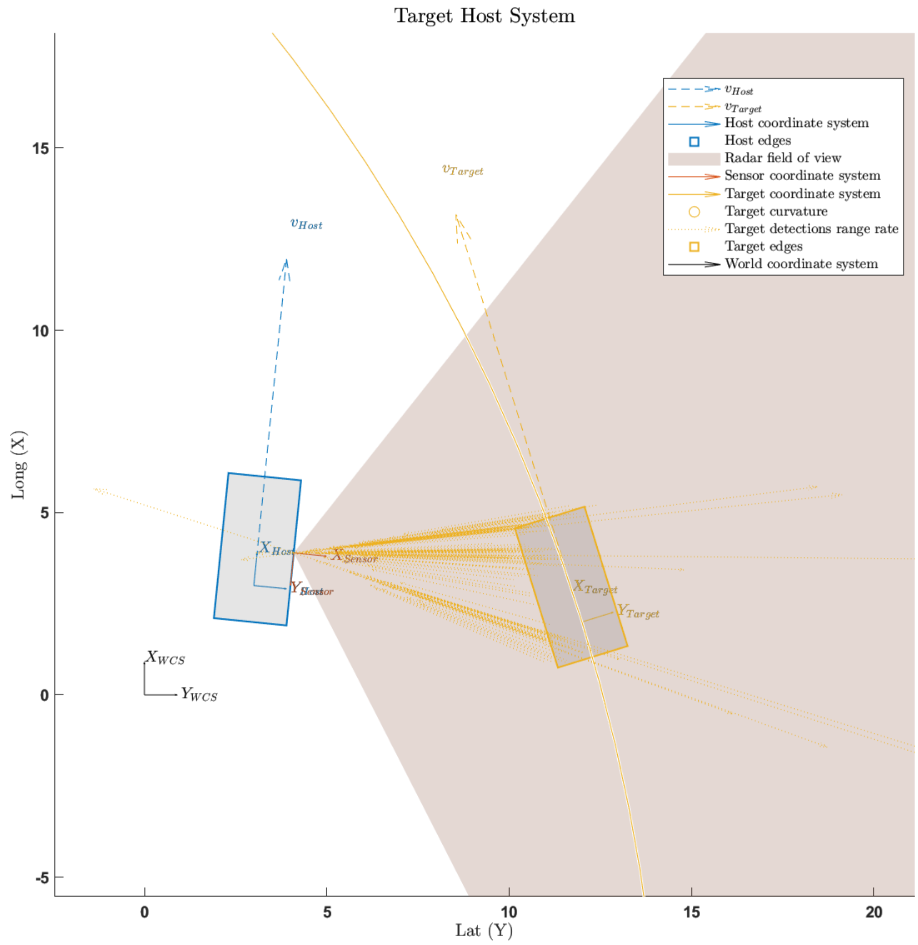


Figure 2.8: An example of generated data with noise and outliers. The arrows from the detection points represent the range rate measured by the sensor mounted on the host vehicle. Here it can be seen that the detection model in [6] is used since most detections are at the closest corner and at the positions of the wheels on the target bounding box.

3
Method

The objectives of this project was to estimate the size, orientation, velocity and yaw rate of a target vehicle using the RDC from a single radar frame. This was achieved in the steps that can be seen in the flowchart in Figure 3.1. The range, compensated range rate and azimuth of the scatter points were extracted from the RDC from an SRR3. Then all scatter points with an amplitude below CFAR were removed as described in Section 2.2.1. After all usable data was extracted, the first step was to develop a clustering algorithm to sort out any scatter points not belonging to the target vehicle. The clustering algorithm is described in detail in Section 3.1. In order to estimate the size and orientation of the target an algorithm for finding the bounding box was developed. This algorithm is described in Section 3.2. To be able to estimate around what point the target is yawing, the position of the target's rear axle was estimated using the assumption that the density and amplitude of the scatter points is greater around the wheels. This is described in Section 3.3. Using the rotation point and the measurements the yaw rate was then estimated. This is described in Section 3.4. Finally, in Section 3.5 it is described how the velocity was estimated. Since the noise level on estimations made in a single radar frame are high, a Kalman filter was developed to smooth out the yaw rate and velocity estimations over time. The Kalman filter is described in Section 3.6.

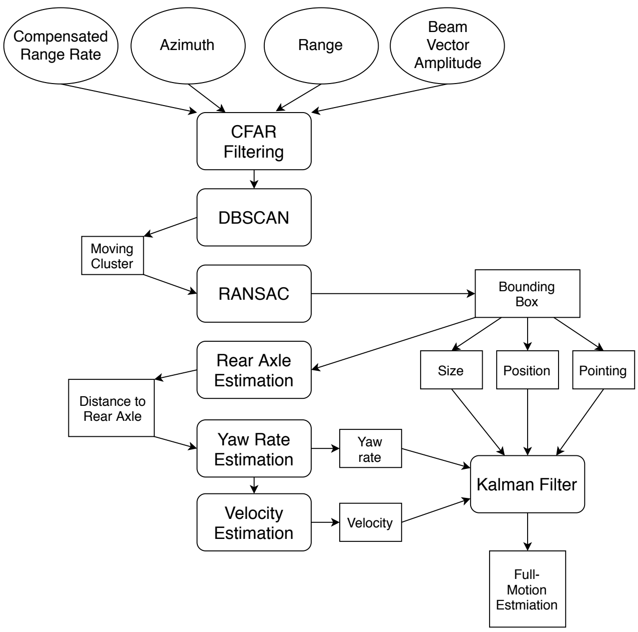


Figure 3.1: A flowchart describing what methods were used in this project.

3.1 Clustering Using DBSCAN

In order to estimate size, velocity and yaw rate of objects surrounding the sensor, the detections needed to be clustered into different targets. To do this, a version of DBSCAN was used. The classical DBSCAN algorithm only clusters data points based on their position [7]. In this case there can be detections in the same range and with almost the same azimuth that belongs to different objects, for example detections from different objects very close to each other. If the algorithm only clusters points based on their position these points might be classified as belonging to the same object. The outliers, for example detections on the wheels of a target vehicle, also need to be removed since their measured range rate does not reflect the motion of the target. Hence, only considering range and azimuth is not sufficient. By also taking range rate into account when doing the clustering, the risk of having outliers or detections from a different object in a cluster is decreased.

3.1.1 DBSCAN Algorithm

The DBSCAN algorithm clusters points that are close to each other. The classical DBSCAN algorithm only clusters spatial data and takes a euclidean distance, ϵ, and a minimum number of points, $minPts$, as input. The algorithm loops over all points and if a point p has more than $minPts$ in its neighbourhood, that is within the euclidean distance, ϵ, from p, the points belong to a cluster. Points with fewer than $minPts$ neighbours are considered as noise [7]. In the DBSCAN algorithm used in this project, however, the detections are clustered based on azimuth, range and range rate. Since there is not only spatial data as input, the euclidean distance can not be used as a threshold value. Instead the data is ordered in a range-Doppler map where the value of each bin is the azimuth instead of the magnitude of the signal. The threshold input, ϵ, is a 1-by-3 vector, where the first value describes how many range bins from a point its neighbourhood is. The second value describes how many Doppler bins from a point its neighbourhood is and the third element of ϵ is a threshold value for the azimuth. A flowchart of the DBSCAN algorithm can be seen i Figure 3.2.

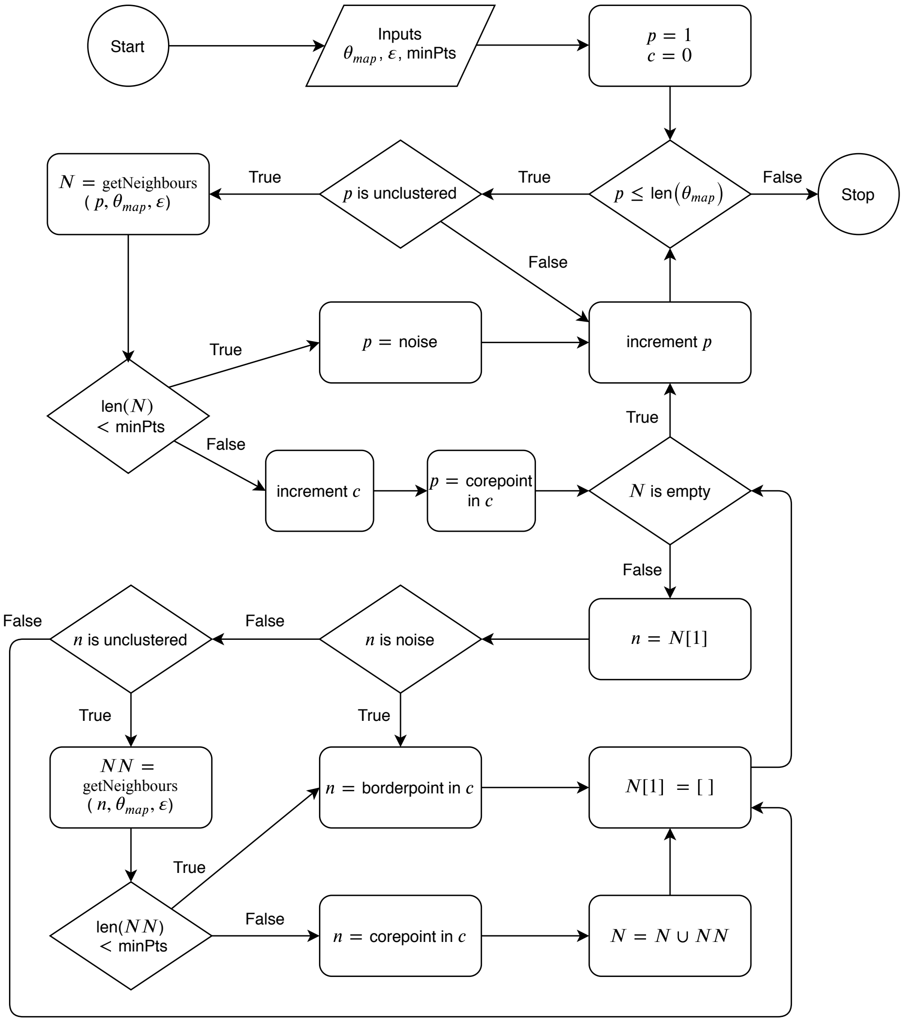


Figure 3.2: A flowchart describing the DBSCAN algorithm.

The DBSCAN algorithms uses another algorithm called getNeighbours, which takes an index of the azimuth map, the azimuth map and the threshold values ϵ as input and returns a list with the indices of the neighbours of the input point. Neighbours are in this case points that are in the range set by ϵ from the current point.

3.1.2 Cluster Properties

Besides labeling all bins with a cluster ID or as noise, the DBSCAN algorithm outputs information about each of the detected clusters. The properties provided are the size of the cluster, the number of core points, the number of edge points, an approximate location in the azimuth map and if the cluster is moving or not.

3.1.3 Moving and Stationary Cluster Classification

For each point in a cluster, the compensated range rate is calculated using Equation (2.2). If the average speed over ground is below 1 m/s the object is classified as stationary. Since only logs with one moving car was used, the largest moving cluster was assumed to be the target vehicle.

3.2 Bounding Box Estimation Using RANSAC

In order extract the length and width of a target vehicle, the edges of the target had to be determined from a moving cluster in the RDC. This was accomplished by using RANSAC. RANSAC is an algorithm for estimating the parameters of a mathematical model based on a set of data points with outliers. The method is non-deterministic since the result can vary from one time to another, even though the input data is the same. Most of the time the algorithm does, however, give a sufficiently good estimation. RANSAC is based on the method to take a minimum numbers of data points and fit a mathematical model to these points. The solution is then evaluated based on how many of the total number of points that are within a pre-set distance from the solution obtained by this minimal number of points. The points within that distance are called inliers. The algorithm is then iterated and if a solution is found that have more inliers than any of the previous solutions, that solution is saved. If the sought solution is a polynomial of grade N, the minimum number of points that RANSAC samples and fits a solution to is $N+1$ [8].

A modified version of the RANSAC algorithm was used to estimate the position of the sides of a target based on the measurements. Since two sides of the target (an L-shape) often is visible to the sensor, the algorithm was constructed to find L-shapes rather than finding a polynomial. The two sides visible was assumed to be perpendicular to each other. The lines aligning with the two sides can then be described using Equations (3.1) and (3.2).

$$y = kx + m_1 \tag{3.1}$$

$$y = -\frac{1}{k}x + m_2 \tag{3.2}$$

Since there are three parameters to be decided, k, m_1 and m_2, the RANSAC algorithm needs to take three random sample points to fit the model to. First k and m_1 are decided by fitting a polynomial of grade 1 to two of the points. Then m_2 is decided by finding a perpendicular line to the first polynomial that crosses the y-axis in m_2. For each of the two lines, all scatter points within the threshold distance from the line are added to a list of inliers. The threshold value used in this project was 0.5 m. Then the point where the two lines intersect is computed using Equations (3.3) and (3.4).

$$x_{\text{intersect}} = \frac{m_2 - m_1}{k + \frac{1}{k}} \tag{3.3}$$

$$y_{\text{intersect}} = k\, x_{\text{intersect}} + m_1 \tag{3.4}$$

The points that lies closer to the sensor than the intersection point does, are then removed from the lists of inliers before evaluating the solution in each iteration.

Instead of just counting the inliers when evaluation the solution in each iteration, the inliers are weighted based on their amplitudes. The sum of the weight of the inliers is the score of that solution. Hence, the best solution doesn't necessarily have the most inliers if it has many inliers of high amplitude.

The euclidean distance between the intersection point and the inlier farthest away from the intersection point on each line is the estimated length of the side of the target vehicle corresponding to that line. The length of the longest side is the length of the target vehicle and the length of the shortest side is the width of the target vehicle.

3.2.1 RANSAC with Only One Side of the Target Visible

In each iteration of the RANSAC algorithm, the score for each individual line is computed and the best one is saved. The best individual line after all iterations is used in case only one side of the target vehicle is visible. To decide whether one or two sides of the target are visible the ratio between the score of the two lines in the best solution is computed. If one side has less than 35% of the score of the other line, it is assumed that only one side is visible. Then the best saved individual line is used instead.

The euclidean distances between all inliers are calculated and the longest distance between two inliers is the estimated length of the side. If the length of the visible side is estimated as greater than 3 meters, the side is considered as the long side of the target. The width of the target vehicle is then estimated as the length multiplied with 0.4. If the length of the visible side is less than 3 meters it is instead considered as the short side. The length of target vehicle is then estimated as the width divided by 0.4.

3.2.2 Number of RANSAC Iterations

The number of RANSAC iterations N needed to obtain a probability p to sample only inliers when sampling n points and where the probability of each point being an inlier is κ, can be calculated using Equation (3.5) [8].

$$N = \frac{\log(1-p)}{\log(1-\kappa^n)} \tag{3.5}$$

To calculate a reasonable number of iterations the inlier rate had to be estimated. The inlier rate varied depending on which test log and which frame was considered. The proportion of inliers, κ in each frame of all the test logs was determined by checking the inlier rate after 1000 RANSAC iterations in each frame. The result can be seen in Figure 3.3. The mean value was 0.48 and the standard deviation was 0.116. A conservative inlier rate was then chosen as two standard deviations below the mean value, that is 0.248. A success rate p of 0.999 was chosen and the number of sampled points, n, was 3, since that is what is needed to form an L-shape. The minimum number of iterations needed was calculated as 450.

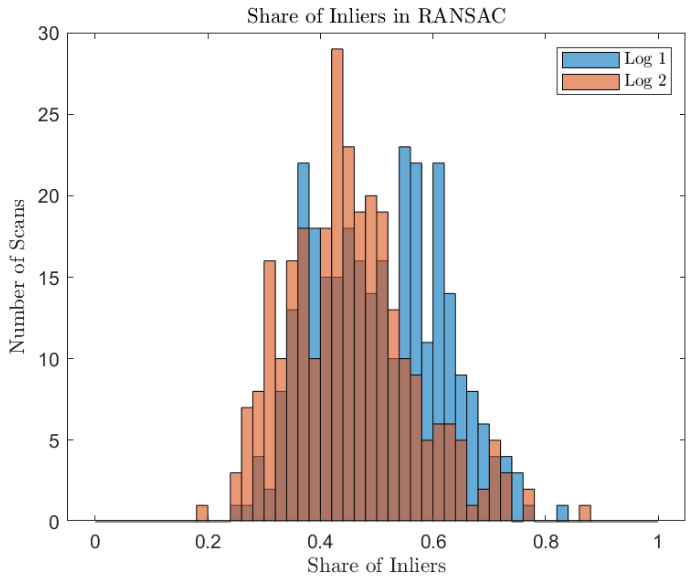


Figure 3.3: The share of inliers over two test logs where the target vehicle drives in a circle.

3.2.3 Unambiguous Pointing of Bounding Box

The RANSAC algorithm results in a bounding box that has a pointing which is ambiguous since it does not incorporate the direction of travel of the target, only the orientation. Assuming that the linear velocity of the target has a greater influence on the measured range rates than the yaw rate has, it is possible to adjust the pointing from RANSAC. If this assumption hold, the front and back side of the target vehicle can be determined. If the front or the back of the vehicle is visible, the mean of the range rates on these sides are used to determine whether the pointing is towards or away from the sensor. When only the side of the target vehicle is visible, the measured range rates along the length of the object should be lower in the front and larger in the back along the length of the target. This information is then used to adjust the pointing of the bounding box to be in the direction of travel.

3.3 Rear Axle Estimation

According to the Ackermann steering condition there is one point on the vehicle where there is only linear speed and no additional speed due to rotation of the vehicle. This point will throughout this report be called the reference point. The ICR line is perpendicular to the pointing of the target vehicle and intersects the reference point. The position of the reference point can be approximated to be in the middle of the rear axle. When assuming that the ICR lies in line with the rear wheels of the car, the assumption that there is no slip angle is also made, which means that the wheels are traveling in the direction that they are pointing [9]. This is a reasonable assumption for street cars that have a low lateral acceleration [9]. To define the reference point, the rear axle of the car has to be estimated. Since the pointing, and size of the object as well as the cluster corresponding to each bounding box are known, the rear axle position in the bounding box can be determined. For this to work the assumption that the area around the wheels have a higher density of beam vectors with high amplitudes reflected off of them than the rest of the side of the car, has to hold [6]. The amplitude of the beam vectors along the length of the bounding box can be computed [10]. This is done by adding the amplitudes of beam vectors along the length of the bounding box. 100 points along the edge of the bounding box is evaluated and the beam vectors with an Euclidean distance of 1 m is considered to be inliers. The aggregate contour plot over a whole log is shown in Figure 3.4. The peaks in the plot corresponds to the positions of the wheels and the first of these peaks is found in each radar frame.

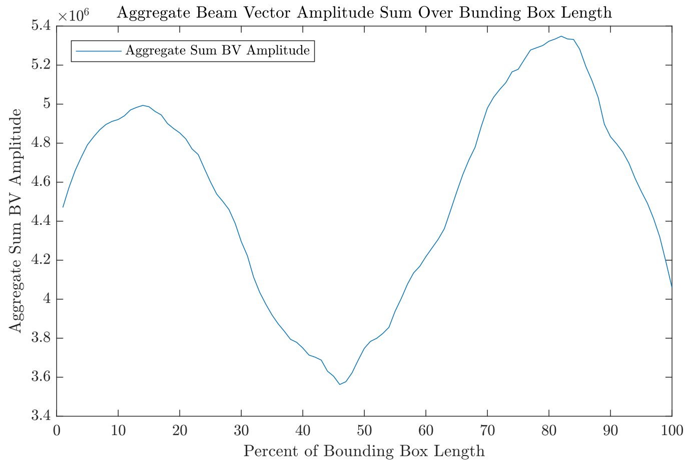


Figure 3.4: The sum of beam vectors near bounding box over its length.

This analysis result in a distance from the rear end of the bounding box to the two wheels on the side. Using the pointing of the bounding box, the distance to the rear axle from the rear end of the bounding box can be determined. The result for one scan is shown in Figure 3.5.

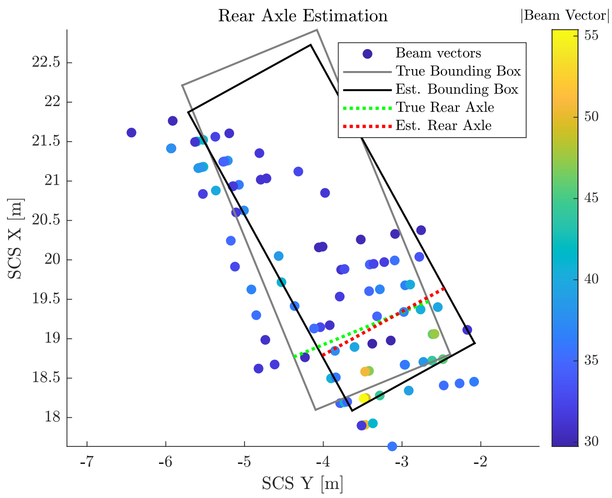


Figure 3.5: The rear axle extracted from the DBSCAN-clustered beam vectors, the bounding box from RANSAC and the true position of the bounding box with corresponding rear axle.

3.4 Yaw Rate Estimation

To estimate the yaw rate of a target vehicle in one radar frame with one sensor the ICR have to be known. The ICR is the center point of the curvature of the target. In order to estimate the ICR, the pointing, a reference point on the vehicle and the ICR line have to be known. The pointing is estimated with the modified RANSAC algorithm and the reference point is the middle of the rear axis that was estimated in Section 3.3. The ICR line describes all possible centers of rotation given the range rates and azimuths corresponding to the target vehicle. To compute the ICR line Equation (3.6), where ω is the yaw rate of the target, θ is the azimuth and $\dot{r}_{comp}$ is the compensated range rate, was used.

$$\dot{r}_{comp} = \underbrace{\omega y_{ICR}^{[S]}}_{\alpha} \cos(\theta) - \underbrace{\omega x_{ICR}^{[S]}}_{\beta} \sin(\theta) \tag{3.6}$$

Two scatter points from the target was sampled based on the norm of the corresponding beam vector and a linear system of two equations were solved to get α and β as in Equation (3.7). The subscripts i and j refers to the two sampled indices. The probability of sampling a certain index i is described in Equation (3.8) where BV is all the beam vectors correlating to one target. The center of rotation $(x_{ICR}^{[S]},\ y_{ICR}^{[S]})$ is in the sensor's coordinate system and the range rate $\dot{r}_{comp}$ is compensated for the velocity of the host vehicle [11].

$$\begin{bmatrix} \dot{r}_{comp,i}^{[S]} \\ \dot{r}_{comp,j}^{[S]} \end{bmatrix} = \begin{bmatrix} \cos(\theta_i) & \sin(\theta_i) \\ \cos(\theta_j) & \sin(\theta_j) \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix} \tag{3.7}$$

$$P(i) = \frac{||BV_i||}{\Sigma||BV||} \tag{3.8}$$

The system of equations in Equation (3.7) are solved 100 times using different sampled scatter points from the target. This results in 100 different estimates for α and β and the medians of the respective lists were assumed to be reasonable estimates. The estimates of α and β are then used to describe the ICR line as in Equation (3.9).

$$y^{[S]} = -\frac{\beta}{\alpha} x^{[S]} \tag{3.9}$$

According to the Ackermann steering condition the ICR lies perpendicular to the pointing of the bounding box of an object and hence the center of rotation, ICR, of an object is along the ICR line somewhere between ICR_{min} and ICR_{max} in Figure 3.6. Since the reference point P_{ref} is known the ICR can be determined as the intersection between the line that goes trough the rear axle of the bounding box and the ICR line as shown in Figure 3.6.

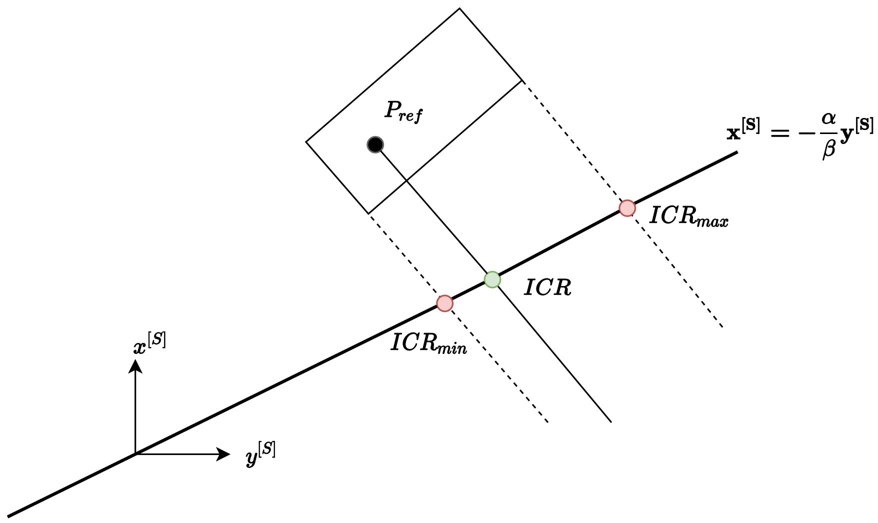


Figure 3.6: The ICR line and pointing of a target vehicle in the sensor coordinate system. The bounding box describes the target vehicle.

When the ICR line in Equation (3.9) and the center of rotation ICR, described by the coordinates $(x_{ICR}^{[S]},\ y_{ICR}^{[S]})$, is determined then the yaw rate, ω, can be computed. Using the expression in Equation (3.6) the yaw rate can be expressed as in Equation (3.10).

$$\omega = \frac{\alpha}{y_{ICR}^{[S]}} \tag{3.10}$$

3.5 Velocity Estimation

When the yaw rate and the center of rotation are known, the linear velocity of the target vehicle can be computed using Equation (3.11) where R is the radius of the turn which is the distance from ICR to P_{ref} in Figure 3.6. v is the velocity and ω is the yaw rate.

$$v = \omega \cdot R \tag{3.11}$$

3.6 Kalman Filter

Since the low level measurements from most radars are noisy, the estimated yaw rate and velocity are also noisy if scatter points from just one radar frame are used to make the estimation. To show that the method for estimating the yaw rate and velocity instantaneously would work if the noise level of the measured data was lower, a Cubature Kalman Filter (CKF) was used to get a better estimation of the position and orientation of the target in each frame. The general non-linear motion and measurement model used in the CKF are shown in Equation (3.12) and (3.13) [12]. $\hat{\mathbf{x}}_k$ is the state vector in a given time instance, $\mathbf{f}$ is the deterministic part of the motion model and $\mathbf{q}_{k-1}$ is added Gaussian noise sampled from a distribution with the variance $\mathbf{Q}_{k-1}$. The measurement model consists of a deterministic part $\mathbf{h}$ which describes the relation between the states $\hat{\mathbf{x}}_k$ and the measurements $\mathbf{y}_k$. The measurement model also consists of added Gaussian noise $\mathbf{r}_k$ that are sampled from a distribution with the variance $\mathbf{R}_k$.

$$\hat{\mathbf{x}}_k = \mathbf{f}\left(\hat{\mathbf{x}}_{k-1}\right) + \mathbf{q}_{k-1}, \quad \mathbf{q}_{k-1} \sim \mathcal{N}\left(\mathbf{0},\, \mathbf{Q}_{k-1}\right) \tag{3.12}$$

$$\mathbf{y}_k = \mathbf{h}(\hat{\mathbf{x}}_k) + \mathbf{r}_k, \quad \mathbf{r}_k \sim \mathcal{N}\left(\mathbf{0},\, \mathbf{R}_k\right) \tag{3.13}$$

The CKF is a non-linear Kalman filter that estimates the mean and variance of the states using sigma points $\mathcal{X}$. Sigma points are a set of points that represent the distribution of the states in each time instance. In every time instance the sigma points are constructed as in Equation (3.14). These are then weighted equally with the weight W. $(\mathbf{P}_{k-1|k-1})_i$ is the ith column of the variance of the states in the previous time step $k-1$.

$$\begin{aligned}
\mathcal{X}_{k-1}^{(i)} &= \hat{\mathbf{x}}_{k-1} + \sqrt{n}(\mathbf{P}_{k-1|k-1})_i^{1/2}, \qquad i = 1, 2, \ldots, n \\
\mathcal{X}_{k-1}^{(i+n)} &= \hat{\mathbf{x}}_{k-1} - \sqrt{n}(\mathbf{P}_{k-1|k-1})_i^{1/2}, \qquad i = 1, 2, \ldots, n \\
W &= \frac{1}{2n}
\end{aligned} \tag{3.14}$$

These points are then propagated through the prediction step and a prediction mean and variance are computed as shown in Equation (3.15) [12].

$$\begin{aligned}
\hat{\mathbf{x}}_{k|k-1} &\approx \sum_{i=1}^{2n} W \mathbf{f}(\mathcal{X}_{k-1}^{(i)}) \\
\mathbf{P}_{k|k-1} &\approx \sum_{i=1}^{2n} W(\mathbf{f}(\mathcal{X}_{k-1}^{(i)}) - \hat{\mathbf{x}}_{k|k-1})(\mathbf{f}(\mathcal{X}_{k-1}^{(i)}) - \hat{\mathbf{x}}_{k|k-1})^T + \mathbf{Q}_{k-1}
\end{aligned} \tag{3.15}$$

The update step uses the predicted mean $\hat{\mathbf{x}}_{k|k-1}$ and variance $\mathbf{P}_{k|k-1}$ and new sigma points are constructed as shown in Equation (3.16).

$$\begin{aligned}
\mathcal{X}_{k}^{(i)} &= \hat{\mathbf{x}}_{k|k-1} + \sqrt{n}(\mathbf{P}_{k|k-1})_i^{1/2}, \qquad i = 1, 2, \ldots, n \\
\mathcal{X}_{k}^{(i+n)} &= \hat{\mathbf{x}}_{k|k-1} - \sqrt{n}(\mathbf{P}_{k|k-1})_i^{1/2}, \qquad i = 1, 2, \ldots, n \\
W &= \frac{1}{2n}
\end{aligned} \tag{3.16}$$

Using the new sigma points from the prediction step, the mean and variance of the states can be updated with the measurements $\mathbf{y}_k$. The updated mean is then $\hat{\mathbf{x}}_{k|k}$ and the prediction is $\mathbf{P}_{k|k}$ as shown in Equations (3.17) [12].

$$\begin{aligned}
\hat{\mathbf{y}}_{k|k-1} &\approx \sum_{i=1}^{2n} \mathbf{h}(\mathcal{X}_k^{(i)})W \\
S_k &\approx \sum_{i=1}^{2n} (\mathbf{h}(\mathcal{X}_k^{(i)}) - \hat{\mathbf{y}}_{k|k-1})(\mathbf{h}(\mathcal{X}_k^{(i)}) - \hat{\mathbf{y}}_{k|k-1})^T W + \mathbf{R}_k \\
P_{xy} &\approx \sum_{i=1}^{2n} (\mathbf{h}(\mathcal{X}_k^{(i)}) - \hat{\mathbf{y}}_{k|k-1})(\mathbf{h}(\mathcal{X}_k^{(i)}) - \hat{\mathbf{x}}_{k|k-1})^T W \\
\hat{\mathbf{x}}_{k|k} &\approx \hat{\mathbf{x}}_{k|k-1} + P_{xy} S_k^{-1} (\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}) \\
\mathbf{P}_{k|k} &\approx \mathbf{P}_{k|k-1} - P_{xy} S_k^{-1} P_{xy}^T
\end{aligned} \tag{3.17}$$

The state vector used in the filter can be seen in Equation (3.18), where r_{c_x} is the x coordinate of the reference point, r_{c_y} is the y coordinate of the reference point on the target, v is the linear velocity of the target, φ is the pointing and ω is the yaw rate of the target.

$$\mathbf{x} = \begin{bmatrix} r_{c_x} \\ r_{c_y} \\ v \\ \varphi \\ \omega \end{bmatrix} \tag{3.18}$$

The motion model in the update step is described by a Coordinated Turn (CT) model, where the relations between the velocity, yaw rate, position and pointing are captured [13]. The motion model is shown in Equation (3.19) where T is the time between the radar frames which for the SRR3 is 50 ms. The velocity and yaw rate is modelled as constant with added Gaussian noise which has a variance of 0.001 on v and 0.001 on φ as in Equation (3.20).

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} x_1 + x_3 T \cos(x_4) \\ x_2 + x_3 T \sin(x_4) \\ x_3 \\ x_4 + T x_5 \\ x_5 \end{bmatrix} \tag{3.19}$$

$$\mathbf{Q} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.001 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.001 \end{bmatrix} \tag{3.20}$$

The measurement model, described in Equation (3.21), is constant but the velocity and pointing are excluded as measurements because these measurements have many outliers and are the most noisy which can make the Kalman filter less accurate. The measurement variance used is shown in Equation (3.22). The variance of the reference point is assumed to be larger then the variance of the yaw rate.

$$\mathbf{h}(\mathbf{x}) = \begin{bmatrix} x_1 \\ x_2 \\ x_5 \end{bmatrix} \tag{3.21}$$

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0.5 \end{bmatrix} \tag{3.22}$$

4
Results

To test the performance of the state estimation algorithms, estimates were done on a log from test drive. The log is from a case where the target vehicle drives in circles in front of the radar. The target vehicle in the log was a Peugeot Partner L1 2016. The size estimates were compared to the actual size of the target vehicle. The yaw rate and velocity estimates were compared to the yaw rates and velocities measured by GPS in the test drives. The performance of the algorithms was evaluated both for the case where all the data above the CFAR in the RDC was used and for the case where only detections were used. The detections in this case are the output from the radar where one detection is computed and interpolated from a few adjacent bin measurements, as described in Section 2.2.2. The number of RDC points and detections varied a lot depending on which side of the target vehicle was visible to the radar. As can be seen in Figure 4.1, the number of RDC points was significantly greater than the number of detections in all frames. The average number of RDC points was 123.42 and the average number of detections was 7.61.

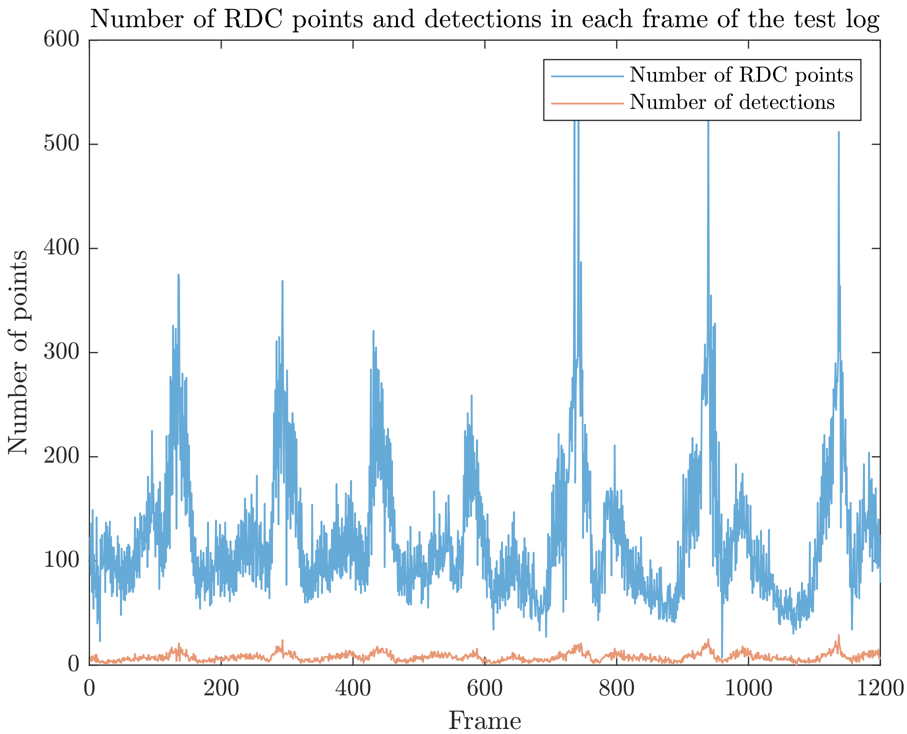


Figure 4.1: The number of RDC points and detections in each frame of the test log.

4.1 Size Estimation

The distributions of the size estimates from the test drive can be seen in Figure 4.2 where both the estimated size when using all the information in the RDC and only the detections are presented. The resulting mean values and standard deviations of the estimations over the whole log are presented in Table 4.1 and 4.2. The true length of the vehicle used in the test drives is 4.38 m and the width is 1.81 m which is presented as a vertical line in the histograms.

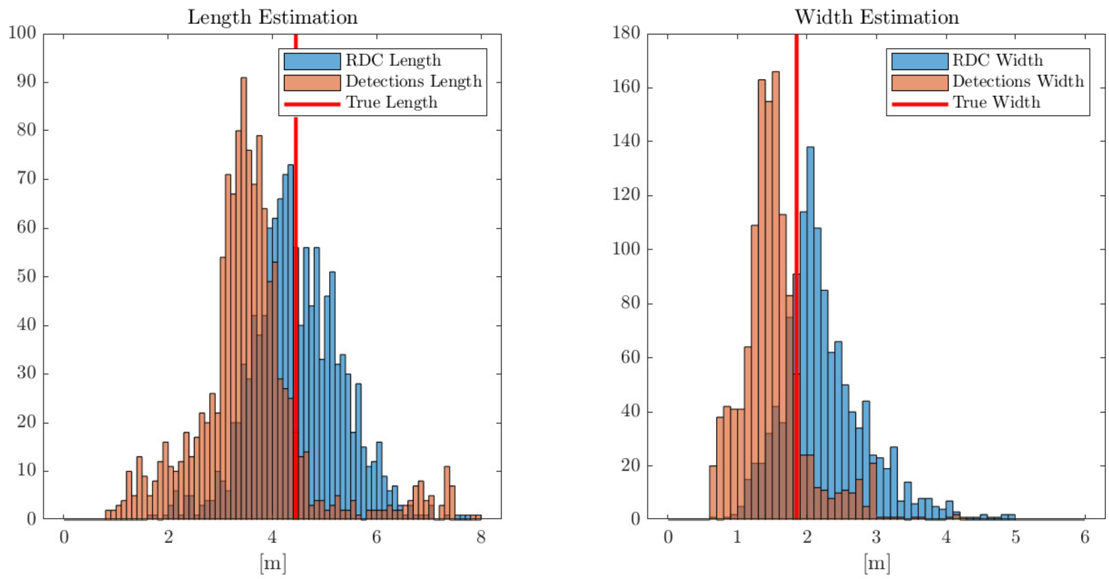


Figure 4.2: Histograms showing the estimation of the length and width of the target. The length is shown in the left histogram and width in the right histogram.

Table 4.1: Table over the estimated length mean, median and standard deviation for both when using the whole RDC and only the detections.

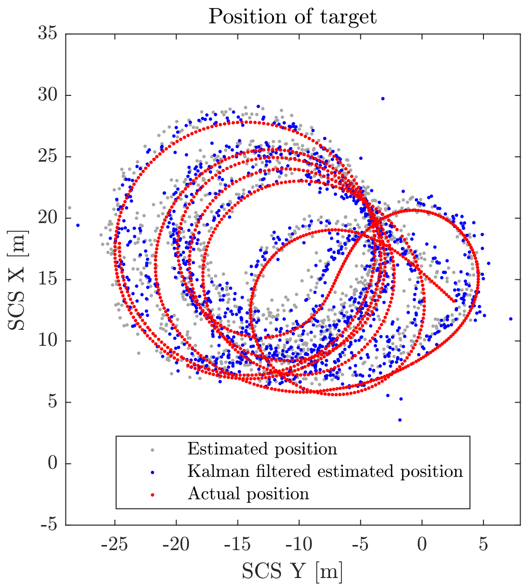
Length	**Mean**	**Median**	σ
RDC	4.58 m	4.42 m	1.21 m
Detections	3.62 m	3.50 m	1.24 m
True	4.38 m	4.38 m	

Table 4.2: Table over the estimated width mean, median and standard deviation for both when using the whole RDC and only the detections.

Width	**Mean**	**Median**	σ
RDC	2.25 m	2.13 m	0.64 m
Detections	1.51 m	1.47 m	0.48 m
True	1.81 m	1.81 m	

4.2 Position Estimation

The position of the center point of the target in each frame of the log is plotted in Figure 4.3. The figure shows one plot for the case where the whole RDC was used, and one plot for the case where only the detections were used. Both plots show the estimated position both before and after the CKF was applied. Figure 4.4 shows histograms of the position errors from the two cases.



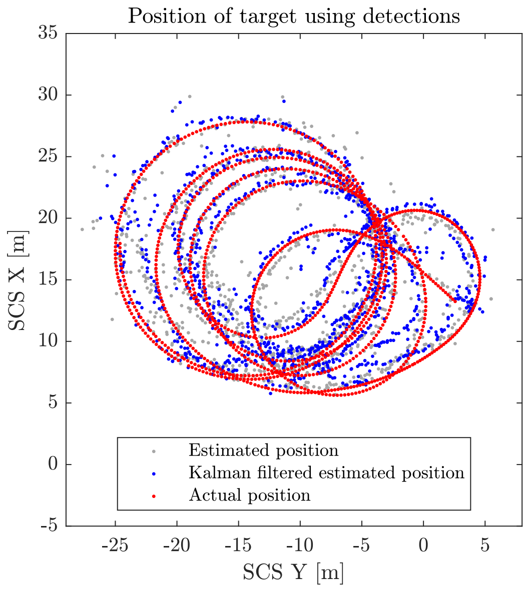
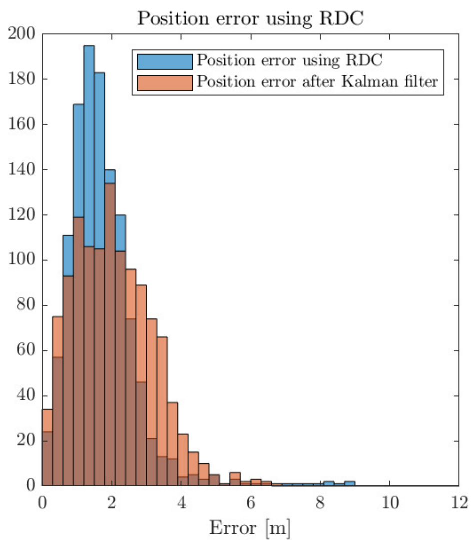


Figure 4.3: The leftmost plot shows the estimated position using the RDC and the true position in each frame. To the right is the estimated position using the detections and the true position.



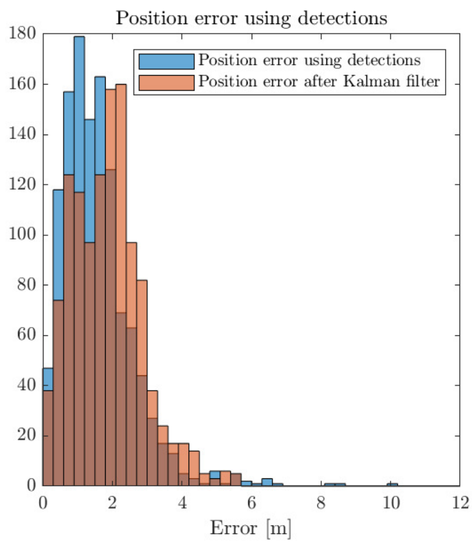


Figure 4.4: The leftmost plot shows error of the estimated position using the RDC compared to the true position. To the right is the position error from the case where only the detections were used.

4.3 Pointing Estimation

The estimated pointing in each frame in the log, before and after the CKF, from the case where the whole RDC was used is plotted in the upper half of Figure 4.5.

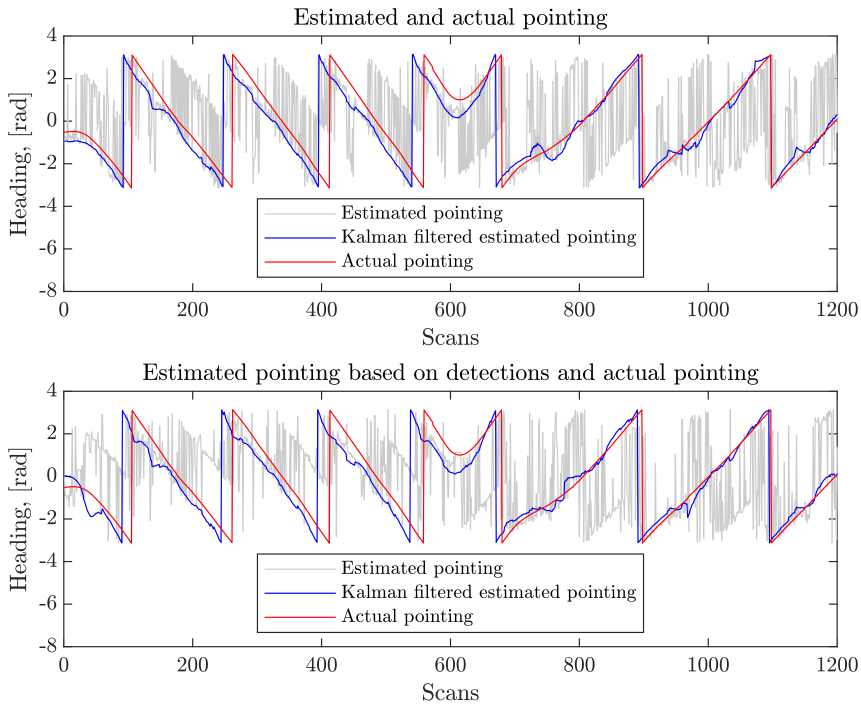


Figure 4.5: The upper plot shows pointing estimations from the RDC and true pointing. Below is the pointing estimations from the detections and true pointing.

The lower half the figure shows the corresponding plot for the case where only the detections were used to estimate the states. The mean square error of the estimations compared to the ground truth GPS data is shown in Table 4.3.

Using the least squares approach described in Section 2.3 the result for the heading in Figure 4.6 is obtained. If correct, the pointing and heading should have a strong correlation, however it differs considerably from the pointing as shown.

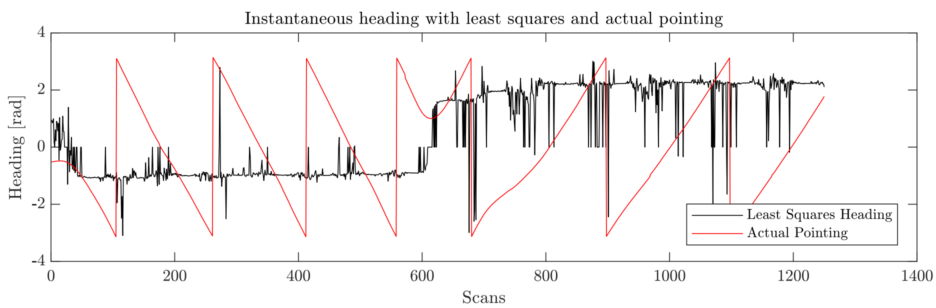


Figure 4.6: Instantaneous heading estimation using a least squares approach and the actual pointing.

Table 4.3: Table over the MSE for the different pointing estimates.

	RDC	RDC Kalman	Dets	Dets Kalman	Least Squares
MSE	4.6	2.0	5.2	2.5	5.9

4.4 Rear Axle Estimation

The result presented in Figure 4.7 is the euclidean distance between the estimated reference point and the actual reference point from GPS data which lies in the center of the vehicle 0.73 m from the back side.

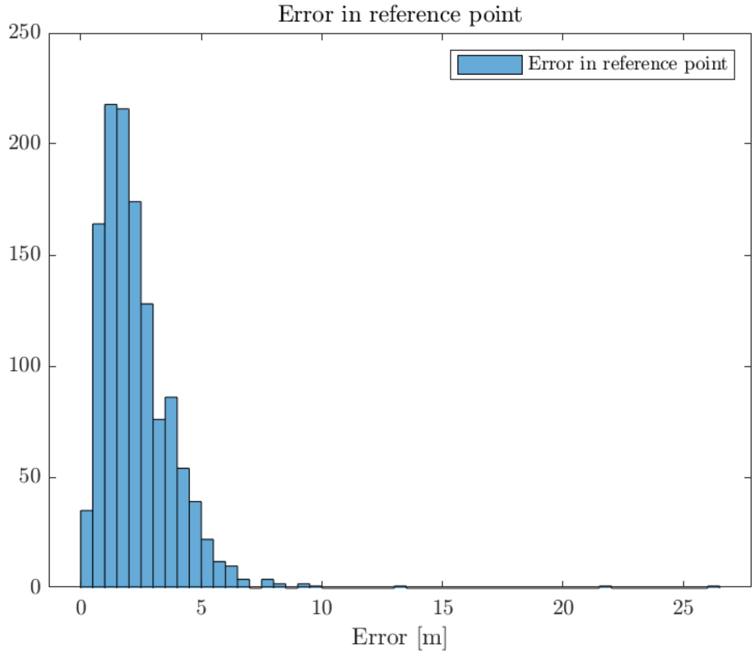


Figure 4.7: The error of the estimated reference point compared to the actual reference point.

4.5 Yaw Rate Estimation

The yaw rate was estimated using only measurements from a single time instance and with the reference point on the target estimated as the center of the rear axle. Both the instantaneous yaw rate and the Kalman filtered yaw rate compared to the actual yaw rate are presented in Figure 4.8. The MSE of the estimates compared to the GPS data is shown in Table 4.4.

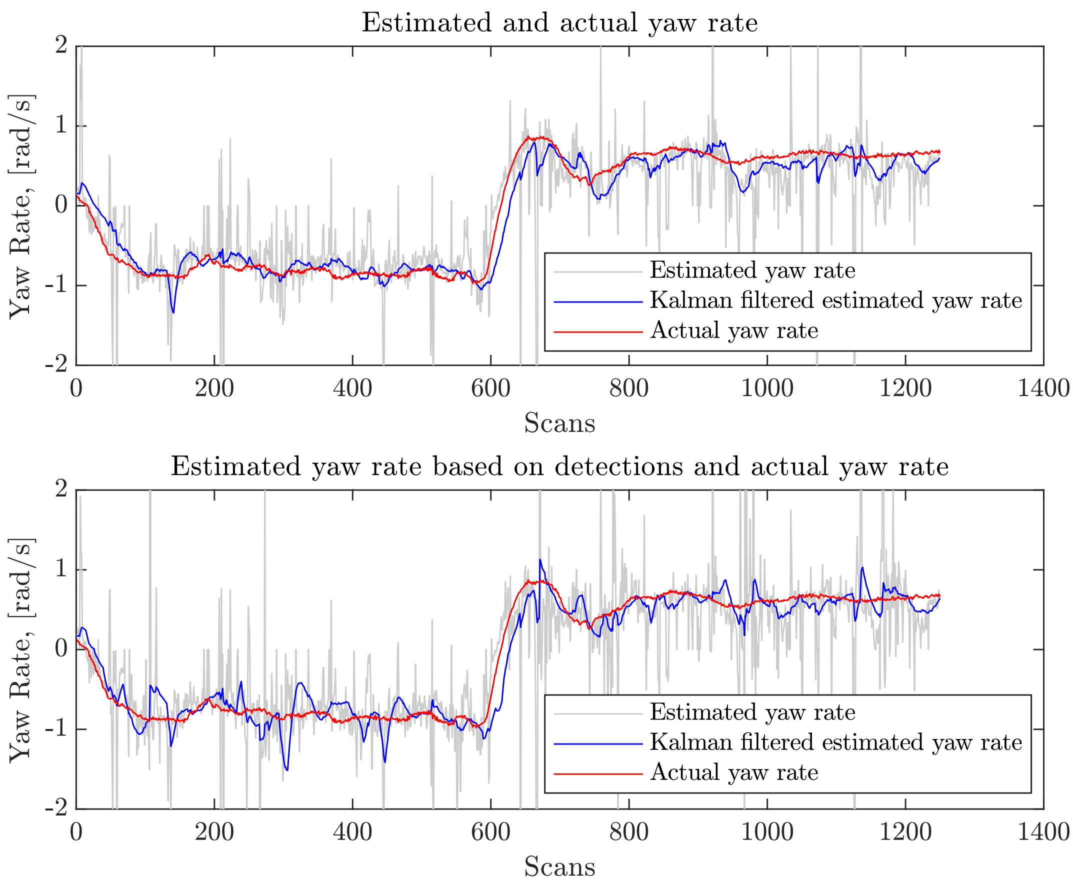


Figure 4.8: Yaw rate estimation results for the whole RDC and detections.

Table 4.4: Table over the MSE for the different yaw rate estimates.

	RDC	RDC Kalman	Dets	Dets Kalman
MSE	0.47	0.032	9.1	0.16

4.6 Velocity Estimation

The velocity estimates from the log are presented in Figure 4.9 together with the actual velocity. This velocity estimation is based on the computations described in Section 3.5 and then Kalman filtered. The mean square error of the estimates compared to the GPS data is shown in Table 4.5.

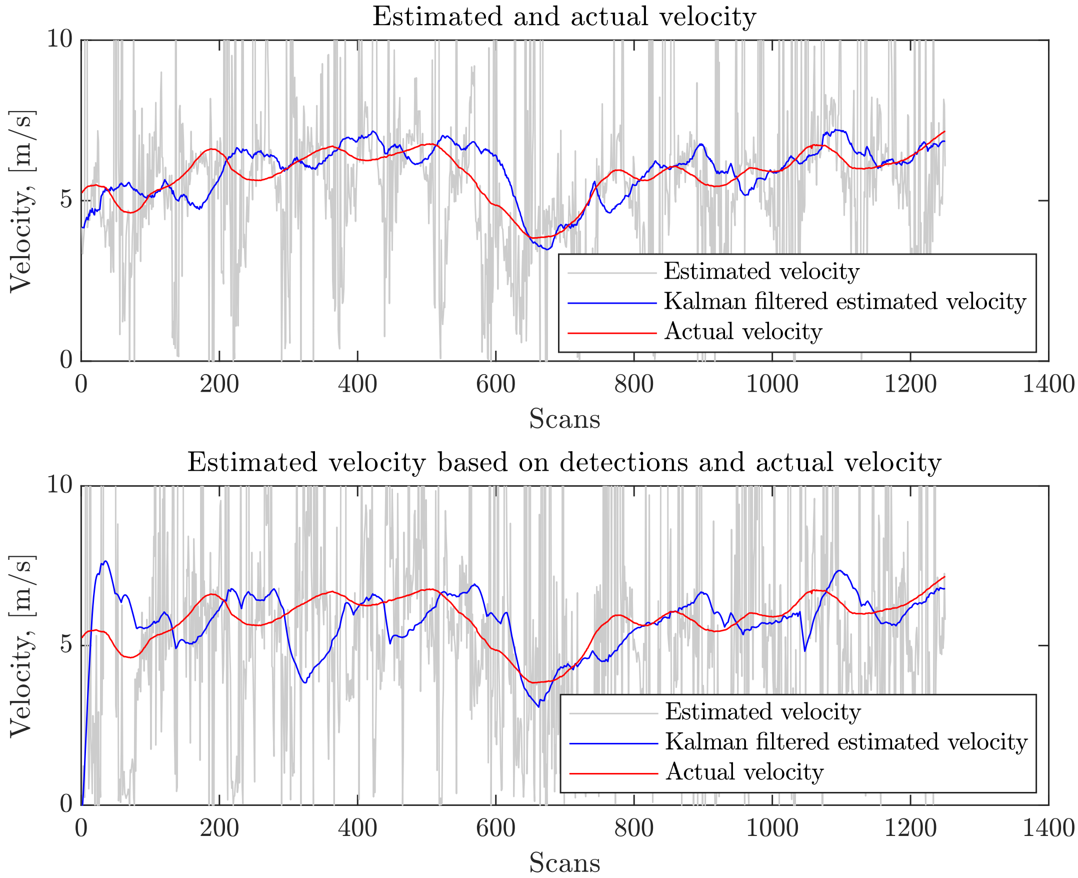


Figure 4.9: Velocity estimation results for the whole RDC and detections.

Using the method of computing the velocity instantaneously presented in Section 2.3 yielded the result in Figure 4.10. Here, the velocity estimates are far larger than the actual velocity.

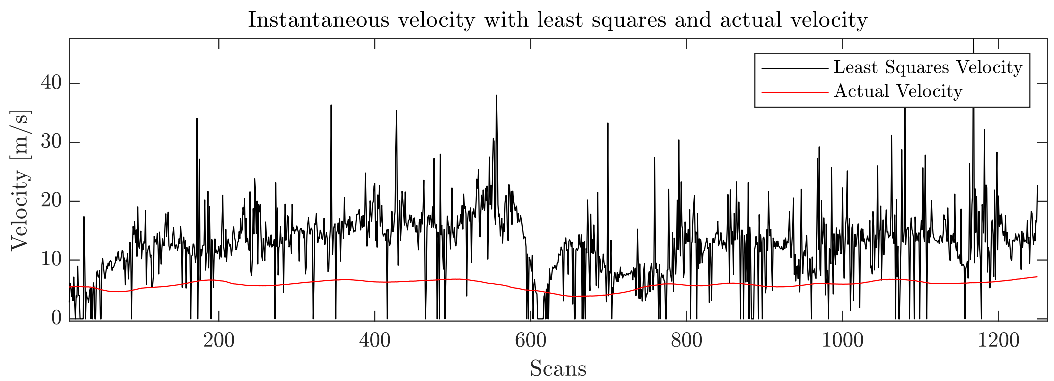


Figure 4.10: Instantaneous velocity estimation using a least squares approach.

Table 4.5: Table over the MSE for the different velocity estimates.

	RDC	RDC Kalman	Dets	Dets Kalman	Least Squares
MSE	120	0.48	4700	0.95	80

5
Discussion

Throughout this thesis a number of assumptions have been made to constrain the project and to develop a solution. It was assumed that the examined radar log only contained one moving object and that the object was a passenger car that behaved according to the Ackermann steering model. For this assumption to hold the car has to experience no slip, the car has to travel in the direction in which the wheels are pointing. This is true in most cases where the speed of a vehicle is low but could be an issue when the yaw rate is high, which could have affected the results of the velocity and yaw rate estimations. The assumption that it was known that the object in question was a passenger car was deemed reasonable since it is reasonable to assume that the radar trackers of the future will have different models for tracking different objects [14]. The setting where the developed model fits in is where the radar outputs a number of detections or beam vectors with a label indicating the type of object, then the model presented in this thesis can be applied.

Since the results presented are from only one log where the host car is stationary on a test track with a target car circulating in front of the radar the results are not necessarily transferable to all other scenarios and use cases. In a cluttered traffic scenario with a lot of cars travelling at the same speed as the host for example, the clustering will be much more difficult which could impact the estimation of the object's states. In a scenario where the host vehicle is moving, the uncertainty in the host's velocity will be added to the uncertainty in the estimation of the target's motion.

5.1 Size Estimation

The estimated size was better when using all the beam vectors in the cluster compared to the estimations when using only the detections. This is because the detections usually don't arise along the whole edges of the vehicle but only on the nearest corner of the target and around the visible wheels. In the RANSAC algorithm the size will then be underestimated since it will see the distance from the rear corner to the front wheel as the whole length. Using all the beam vectors however, the likelihood of getting data from a greater part of the edge increases since there might be low level energy coming from the whole edge of the vehicle that can be captured and the size estimate will be better. As shown in 4.1 the estimates using the whole RDC was generally better compared to using only the detections. There are also more outliers when using the detections which is probably because in some frames there are only a few detections, three or less, which will make the RANSAC solution poor since it samples three points in each iteration.

5.2 Position Estimation

The position is the easiest state to estimate, since it only depends on the position of the scatter points and not on the range rates. As can be seen in Section 4.2, the position estimates are close to the ground truth. The estimations are not improved by the CKF. This is due to the fact that the center point of the target is not a state in the CKF. The center point after the CKF is instead derived from the estimation of the reference point, the rear axle estimation and the estimated pointing. The uncertainties in all these estimates are propagated and hence the uncertainty of the the position of the center point after the CKF is larger than before the CKF.

5.3 Pointing Estimation

The estimated pointing is in most frames fairly good. The MSE was, as expected, smaller when using the whole RDC than when only using detections, both before and after the CKF had been applied. In the estimated pointing before the CKF there are, however, some frames where the error in the pointing is close to π radians. In those cases the position and size of the bounding box oftentimes are somewhat correct, but the rear end of the car is estimated as the front end and vice versa. A general approach to solve this ambiguous pointing is to fit the cluster to a velocity profile and thus compute the heading which should be close to the pointing of the vehicle. This was done as shown in Figure 4.6, but it did not yield any satisfactory result. This is because when the yaw rate is considerably high, as in the log used, it is not possible to fit the correct velocity profile to the bin measurements of the cluster as described in Section 2.3. The proposed method to solve the issue presented in Section 3.2.3 does not seem to work in every scan either which is probably due to the yaw rate contributing more to the range rate then the linear speed of the object.

5.4 Rear Axle Estimation

Evaluating the rear axle estimation is difficult since it is highly dependent on the estimation of the, size, pointing and position of the bounding box. This makes any comparisons relative to the bounding box, for example the distance from the rear side of the bounding box to the estimated position of the rear axle, irrelevant. As discussed in Section 5.3 the pointing estimate is sometimes about π radians off from the true pointing. In those cases, the rear axle estimation algorithm will estimate the front axle as the rear axle. The solution presented is also dependent on that the position and size of the bounding box is estimated correctly since it aims to find the peaks of beam vector amplitudes along the nearest side of the bounding box. Another approach to the problem is to use another iteration of DBSCAN on the clustered beam vectors in order to find the peaks as in [10]. This will however have the same difficulty with the pointing since an accurate instantaneous velocity profile can not be established when the vehicle's yaw rate is high. The behaviour of faulty rear axle estimation dependent on ambiguous pointing is shown in Figure 4.7 where most errors are around 1 m but there is a tendency for a second peak in the histogram around an error of 4 m. In those cases the position of the front axle has been estimated and not the rear axle.

5.5 Yaw Rate Estimation

Estimating the yaw rate for a general cluster in a single radar frame is not possible since there are too many unknown variables. If the assumptions made in this report are valid for a specific object, the yaw rate can be extracted in a single radar frame from one cluster as has been shown. However, the method for estimating yaw rate in a single radar frame has some flaws. As mentioned in Section 3.4 the method is dependent on that the reference point on the rear axle is known. The biggest disadvantage of this method is that its accuracy is very dependent on how accurate the bounding box around the target can be estimated. Especially the accuracy of the estimation of the pointing of the target is crucial for obtaining good results from this method. As can be seen in Section 4.5 the yaw rate estimates are considerably noisy. This is most likely due to the fact that in some instances the estimation of the bounding box is of poor quality.

The yaw rate had a smaller MSE when using the whole RDC compared to detections both before and after the CKF had been applied. This is due to a better estimate of the bounding box and pointing in each frame. More scatter points, as is the case when using the whole RDC, makes the estimation of α and β more accurate and hence the ICR more accurate as well.

5.6 Velocity Estimation

The velocity estimation in a single time instance is very noisy and can not be utilized in a tracker as is. This high level of noise is probably due to the fact that the turn radius is poorly estimated when the line perpendicular to the pointing of the bounding box and the ICR line are close to parallel. Then small errors in the pointing or the slope of the ICR line yield large errors in the turn radius which propagates to the velocity estimation. The Kalman filter where the velocity was excluded as an input resulted in better results however.

5.7 Future Work

There is potential to further improve upon the presented solution and method. If applied to a radar with a higher resolution than the one in used in this project, the results are expected to be better since it would yield more beam vectors per object and it would be easier to extract the position of the rear axle of the vehicle.

Furthermore the solution would have to be verified on more data. During this project there was only data from one test drive available with ground truth. To verify that the methods could work in a more general setting other logs, with target vehicles driving in different speeds, with different yaw rates and in different parts of the radar's field of view would be needed. It is likely that the tuning of the RANSAC algorithm and the Kalman filter would need to be adjusted to work in other settings.

In order to incorporate some, or all the methods used in this thesis in a radar system in a commercial car, all algorithms would need to be more efficient. To make the code efficient enough to run on a car's on board processors was, however, not a part of the scope for this project.

Assuming that a classification of the cluster is given as an input to the developed algorithm it could be extended to cope with other kinds of objects than passenger cars. In order to achieve this the assumption of Ackermann steering has to be modified to match the movement of for example trucks and bicycles. Pedestrians can, however, not be modelled with a motion model like the ones used when modelling vehicles and should be modelled as for example a random walk.

6 Conclusion

The objective of this thesis was to estimate the full motion, size and orientation of an extended object in a single radar frame, which was achieved. The estimates are dependent on that the assumption that the extended object is a passenger car with a steering model described with Ackermann steering holds. However, the result is very noisy and is improved considerably if a Kalman filter with a CT model is applied to the computed state estimates of the object. The result is considerably better when using all beam vectors in a cluster as compared to only detections.

Bibliography

[1] T. Luettel, M. Himmelsbach, and H. Wuensche, "Autonomous ground vehicles—concepts and a path to the future," *Proceedings of the IEEE*, vol. 100, no. Special Centennial Issue, pp. 1831–1839, 2012.

[2] K. Granstrom, M. Baum, and S. Reuter, "Extended object tracking: Introduction, overview and applications," *Journal of Advances in Information Fusion*, vol. 12, pp. 139–174, 12 2016.

[3] T. Jeffrey, *Phased-Array Radar Design - Application of Radar Fundamentals*. Raleigh, NC: SciTech Publishing, 2009. [Online]. Available: https://app.knovel.com/hotlink/toc/id:kpPARDARF5/phased-array-radar-design/phased-array-radar-design

[4] M. Stachnik, D. Cieslar, and K. Kogut, "Method of determining the de-aliased range rate of a target," U.S. Patent 20 180 356 498, 12, 2018. [Online]. Available: http://www.freepatentsonline.com/y2018/0356498.html

[5] N. Carlström and A. Öqvist, "Target identification using low level radar measurements," Master's thesis, Chalmers University of Technology, Chalmersplatsen 4, 412 96 Gothenburg, 06 2018.

[6] M. Bühren and B. Yang, "Simulation of automotive radar target lists using a novel approach of object representation," in *2006 IEEE Intelligent Vehicles Symposium*, 01 2006, pp. 314–319.

[7] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, "A density based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, ser. KDD'96. AAAI Press, 08 1996, p. 226–231.

[8] M. A. Fisher and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, p. 381–395, 06 1981. [Online]. Available: https://doi.org/10.1145/358669.358692

[9] W. F. Milliken and D. L. Milliken, *Race Car Vehicle Dynamics*. Warrendale, PA: SAE International, 1995.

[10] D. Kellner, M. Barjenbruch, J. Klappstein, J. Dickmann, and K. Dietmayer, "Wheel extraction based on micro doppler distribution using high-resolution

radar," in *2015 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*, 06 2015, pp. 1–4.

[11] J. Schlichenmaier, L. Yan, M. Stolz, and C. Waldschmidt, "Instantaneous actual motion estimation with a single high-resolution radar sensor," in *2018 IEEE MTT-S International Conference on Microwaves for Intelligent Mobility (ICMIM)*, 04 2018, pp. 1–4.

[12] H. Arasaratnam and S. Haykin, "Cubature kalman filters," *IEEE Transactions on Automatic Control*, vol. 54, no. 6, pp. 1254–1269, 07 2009.

[13] M. Roth, G. Hendeby, and F. Gustafsson, "Ekf/ukf maneuvering target tracking using coordinated turn models with polar/cartesian velocity," in *17th International Conference on Information Fusion (FUSION)*, 07 2014, pp. 1–8.

[14] R. O. Chavez-Garcia and O. Aycard, "Multiple sensor fusion and classification for moving object detection and tracking," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, no. 2, pp. 525–534, 2016.