Fusing Laser and Radar Data for Enhanced Situation Awareness

Emanuel Eliasson

LIU-IEI-TEK-A--10/00850--SE
Linköping 2010

Linköping University
INSTITUTE OF TECHNOLOGY

Department of Management and Engineering
Division of Fluid and Mechatronic Systems
Linköpings universitet
SE-581 83 Linköping, Sweden

Fusing Laser and Radar Data for Enhanced Situation Awareness

Emanuel Eliasson

LIU-IEI-TEK-A--10/00850--SE
Linköping 2010

Supervisor: **Pär Degerman**
Scania CV AB

Examiner: **Lars Andersson**
IEI, Linköpings universitet

Department of Management and Engineering
Division of Fluid and Mechatronic Systems
Linköpings universitet
SE-581 83 Linköping, Sweden

Avdelning, Institution
Division, Department

Division of Fluid and Mechatronic Systems
Department of Management and Engineering
Linköpings universitet
SE-581 83 Linköping, Sweden

Datum
Date

2010-06-18

Språk
Language

☐ Svenska/Swedish
☒ Engelska/English
☐ _____

Rapporttyp
Report category

☐ Licentiatavhandling
☒ Examensarbete
☐ C-uppsats
☐ D-uppsats
☐ Övrig rapport
☐ _____

ISBN
__

ISRN
LIU-IEI-TEK-A--10/00850--SE

Serietitel och serienummer **ISSN**
Title of series, numbering __

URL för elektronisk version
http://www.iei.liu.se/flumes
http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-57928

Titel
Title

Fusion av laser- och radardata
för ökad omvärldsuppfattning

Fusing Laser and Radar Data
for Enhanced Situation Awareness

Författare Emanuel Eliasson
Author

Sammanfattning
Abstract

With an increasing traffic intensity the demands on vehicular safety is higher than ever before. Active safety systems that have been developed recent years are a response to that. In this master thesis Sensor Fusion is used to combine information from a laser scanner and a microwave radar in order to get more information about the surroundings in front of a vehicle. The Extended Kalman Filter method has been used to fuse the information from the sensors. The process model consists partly of a Constant Turn model to describe the motion of the ego vehicle as well as a tracked object. These individual motions are then put together in a framework for spatial relationships to describe the relationship between them. Two measurement models have been used to describe the two sensors. They have been derived from a general sensor model. This filter approach has been used to estimate the position and orientation of an object relative the ego vehicle. Also velocity, yaw rate and the width of the object have been estimated. The filter has been implemented and simulated in Matlab. The data that has been recorded and used in this work is coming from a scenario where the ego vehicle is following an object in a quite straight line. Where the ego vehicle is a truck and the object is a bus. One important conclusion from this work is that the filter is sensitive to the number of laser beams that hits the object of interest. No qualitative validation has been made though.

Nyckelord
Keywords Sensor Fusion, Kalman Filter, Laser Range Scanner, Radar

Abstract

With an increasing traffic intensity the demands on vehicular safety is higher than ever before. Active safety systems that have been developed recent years are a response to that. In this master thesis Sensor Fusion is used to combine information from a laser scanner and a microwave radar in order to get more information about the surroundings in front of a vehicle. The Extended Kalman Filter method has been used to fuse the information from the sensors. The process model consists partly of a Constant Turn model to describe the motion of the ego vehicle as well as a tracked object. These individual motions are then put together in a framework for spatial relationships to describe the relationship between them. Two measurement models have been used to describe the two sensors. They have been derived from a general sensor model. This filter approach has been used to estimate the position and orientation of an object relative the ego vehicle. Also velocity, yaw rate and the width of the object have been estimated. The filter has been implemented and simulated in Matlab. The data that has been recorded and used in this work is coming from a scenario where the ego vehicle is following an object in a quite straight line. Where the ego vehicle is a truck and the object is a bus. One important conclusion from this work is that the filter is sensitive to the number of laser beams that hits the object of interest. No qualitative validation has been made though.

Sammanfattning

Med en ökad intensitet av trafiken ställs det idag högre krav på fordonssäkerhet än det gjorts tidigare. Ett sätt att möta dessa krav är de aktiva säkerhetssystem som utvecklats de senaste åren. I det här examensarbetet har information från en scannande laser och en mikrovågsradar fusionerats ihop för att höja medvetenheten om vad som händer framför ett fordon. Extended Kalman Filter är metoden som har använts för att fusionera ihop informationen från sensorerna. Processmodellen består dels av en s.k. Constant Turn Model som används för att beskriva rörelser både för det egna fordonet och det objekt som följs. Och dels av ett ramverk där de båda modellerade rörelserna sätts i relation till varandra. Två stycken mätmodeller har använts för att beskriva sensorerna. Dessa har tagits fram från en generell sensormodell. På det här sättet har en skattning gjorts av position och orientering av det intressanta objektet relativt det egna fordonet. Även hastighet, vinkelhastighet och bredden på objektet har skattats. Filtret har implementerats och simulerats i Matlab. Det data som har spelats in och använts i det här arbetet kommer från ett scenario där det egna fordonet följer ett objekt med ganska rak

kurs där det egna fordonet är en lastbil och objektet är en buss. En viktig slutsats som kan dras efter det här arbetet. Är att filtret är känsligt för hur många laserträffar som fås på det intressanta objektet. Dock har ingen kvalitativ validering gjorts.

Acknowledgments

First I would like to thank my examiner and supervisor Lars Andersson for all guidance during this thesis. He has been having great patience with me and given me support throughout the work. I would also like to thank my supervisor Pär Degerman at Scania CV AB for the opportunity I got to do this thesis in cooperation with Scania CV AB.

I am also thankful for the help I got from the students Petter Källström and Mickael Karlsson as well as Henrik Schauman at Scania CV AB.

Contents

Chapter 1

Introduction

This thesis has been performed at the Division of Fluid and Mechatronic Systems at Linköping University in Linköping in cooperation with Scania CV AB in Södertälje. Throughout this work the vehicle in which the sensors are attached to, will be called ego vehicle. Whereas the vehicle that is observed will be called the tracked vehicle. The later one is also referred to as the object. Some state and input signal indexes are o for object concerning the tracked vehicle. Whereas the index for the ego vehicle is e.

1.1 Background

Sensor fusion simply means to fuse information from several sensors. The sensors should also measure different properties and/or operating range to give, in some way a result that is more valuable than the sensors themselves can produce. Otherwise the only benefit would be redundancy. In this work a laser range scanner measures the width of the tracked vehicle and a microwave radar measures the range rate for example. Most active safety systems in autmotive vehicles today are equipped with radars and cameras. Cameras however need a lot of image processing in order to give satisfying information. This thesis is investigating the concept of using the radar combined with a laser instead. In table 1.1 a rough overview of the properties of camera, laser scanner and radar approach is given. These are rough properties for the actual sensor plus some necessary pre-filtering.

	Camera	Laser Scanner	Radar
Azimuth angle	high resolution	medium resolution	low resolution
Range	low resolution	high resolution	medium resolution
Range rate	not	medium resolution	very high resolution
Weather conditions	sensitive	less sensitive	less sensitive
Work load	high	low	low

Table 1.1: Properties of different sensor approaches.

The pros of a camera are that it is relatively cheap compared to at least the laser scanner. And the camera measures the angles with high resolution because it is easy for it to see the contours of an object (in good conditions). There is a lot of information that can be extracted from the camera view that is not possible to get from the other sensors. To read off traffic signs and recognise different colors for example. Cons of the camera are much workload in general and that it is sensitive for weather conditions like rain, snow and fog. Range measurements are not very accurate and range rate is not supported for practical use i.e. without triangulating.

The benefits of using a laser scanner is that it is possible to get important information about the surrounding environment with less data than a camera for example. At least theoretically this small amount off data from the laser scanner should be enough to provide a system with information, that can match a camera when it comes to awareness of stationary and moving objects. It is possible to differentially get range rate from a laser scanner. It should be pointed out that the laser scanner used in this work is of direct detection type. However a laser working with coherent detection is able to measure the range rate with very high resolution. With laser the reflectivity of a surface can be distinguished. The laser scanner used here can detect the line markings on the road for example. Bad weather conditions are not as big problem for the laser as it is for the camera. The main drawbacks are less data compared to a camera meaning less information and that these types of sensors are relatively expensive today. A comparison of two vision sensors and a laser sensor is done in [6] and a comparison between two different laser scanners is made in [8].

The most important property of a radar sensor is the high resolution when detecting the speed of an object. This is of course a desirable property when the system is working in a dynamic environment like vehicular traffic. It is also less sensitive to bad weather conditions than the other two sensor types. On the other hand the radar is not very good in anything else than making range rate measurements. It is less good at measuring range and has poorer resolution when measuring azimuth angle than a laser scanner. Though the difference for resolution in measuring azimuth angle is small in this work. In [12] a radar is used to estimate the free space in front of a moving vehicle and in [11] the radar is used to track stationary objects for road mapping.

Why is it not enough to just have the information from the sensors? Why is a filter needed where the information is fused? To answer these questions lets look at an example with a collision avoidance system where there is a need of very accurate information very fast. The radar can give precise information about how fast a vehicle is approaching the ego vehicle. But it is also important to know if the target vehicle is in collision course with the ego vehicle. To get that information a camera or laser scanner is needed or some other sensor that can measure the relative orientation between the vehicles. Assuming that at least one of them

measures the relative position. It could be tempting to stop here and just use different measured quantities from the sensors and use the raw data with some conditions to trigger the breaks. There are however several problems with this approach. The sensors do have a certain limit when it comes to resolution. No matter how much improvements are done they will always give measurements with some uncertainty. Either because of physical limitations like not perfect lenses etc. Or in the process of translating the measured quantity into numerical information. Even if the sensors would be perfect the environment in which they are applied is not perfect as we know. Weather conditions, irregular geometry of vehicles, roads and other objects will make the data difficult to interpret. Another problem that arise if the data from sensors would be used separately is that extremely accurate calibration would be required. In order to know that they measure the same object or environment. Even for sensor fusion systems the calibration part is crucial. All these problems contributes to the total uncertainty that is not taken care of with this approach. For example in one instant the radar may give data about the situation being critical as well as the laser, saying that both relative speed and orientation combined with the position indicates an unpleasant immediate future. But when the next measurement comes maybe everything is fine from one sensor or both of them. Due to these problems mentioned above it will be impossible for an active breaking system to work for a realistic and dynamic environment. One way to overcome this particular problem could be to use more pre-filtering in conjunction with each sensor. But it would inevitably lead to loss of information.

We would like to have some method where the benefits of using several sensors are preserved. And does not lead to an increasing total uncertainty. We would also like this method to be able to forsee a small time step in a more stable way than the method above. All this points to a filter where it is possible to fuse the information from each sensor in a smart way. Meaning that each new measurement coming from a certain sensor is weighted dependent on how reliable it is. And where the physical quantities that we want to measure, evolves in accordance with a known model over time. There are a couple of methods to solve this problem. In this thesis the choice fell on a type of filter that is the origin of sensor fusion. The Kalman filter. This filter basically works in two phases after the initiation. One phase takes care of how the new measurements of the reality should be interpreted. I.e. how the new observation should affect the current view of the reality. The other phase is predicting what most likely would happen in the next moment according to the information we got this far and the model of the reality. Both these phases need to work if the filter overall is going to work properly. The first phase may be more obvious than the second. We need to get as much information as possible from each measurement. In this phase when a new measurement comes it also suppresses the uncertainty. But the filter is useless without the second phase. Everywhere sensors are used they measures their surroundings in order to see what happens. I.e. it is of no point to measure something that will never change. Now if something can change state it means it is dynamic and that time is always involved. The second phase takes care of this time issue but raises the uncertainty. It evolves the state of what is observed between the measurements

(phase one). The filter needs a couple of models to be able to fulfil its mission. The first phase uses a model of the sensor to interpret the impression to the state. In this work there is one model for the laser and another for the radar. The second phase needs a model of the dynamics i.e. the time dependent behaviour.

1.2 Purpose

The purpose has been to fuse the data from a microwave radar and one laser range scanner together, in order to investigate the performance when these sensors are used together. And also to present the difference between the decision basis given from the sensor fusion versus the one based on the observations. This configuration could be used in active safety systems were the extra information from a camera compared to a laser, is not required for the functionality of the filter. It is also interesting in systems were one can take advantage of the benefits of the laser, such as less computational workload and higher resolution for range measurements.

1.3 Goal

The goal is to give a vehicle better decision basis with the given extra sensors. Both better than without the extra sensors and also better than would be possible with only the information given by the sensors themselves. I.e. without the fusion. Better here means that the estimated quantities of the tracked object as well as the predictions of these quantities are more reliable than without the extra sensors and without the fusion.

1.4 Related Work

In [1] the author is investigating the spatial uncertainties in mobile robot teams. This work deals with the problem of decreasing the spatial uncertainty when several robots share information about their surroundings. It is a so called Simultaneous Localisation And Mapping (SLAM) problem. The author shows with simulated experiments how the spatial uncertainties for two robots affects the resulting uncertainty. This work is different than the objective here but the motion models as well as the sensor models have been taken from this work. Also how to deal with spatial uncertainties have been inspired by this work.

In [10] the author is dealing with the problem of estimating the motion of a vehicle and its surroundings to improve the drivers situation awareness. Also here sensor fusion is being used to get better situation awareness. However the models are different and a camera is used instead of a laser range scanner like in this thesis. In this work not only the motion of a tracked object is estimated but also properties of static objects such as road lines are estimated. A non-linear Kalman filter approach is used. The tracked vehicle is modelled with a similar model that has been used here namely a so called Coordinated (or Constant) Turn Model. whereas the

ego vehicle is modelled with a dynamic model called Single Track Model that involves geometric relationships as well as a tire model. The ego vehicle model is also extended with a road model. Three different setups of sensors are evaluated. One vehicle with a forward looking camera with rear, side and forward looking radars. Another setup with forward looking camera and radar. The third setup is equipped with extra internal (proprioceptive) sensors measuring axle height but no external (exteroceptive) sensors are used.

In [5] the problem of robot localization as well as the idea of using several robots for better localization is investigated. In this work a laser range scanner is used together with proprioceptive sensors to measure the surroundings. Also here a non-linear Kalman filter approach is used. Another similarity to the approach used here is the use of hough transform to estimate lines from the laser data.

1.5 Limitations

The main limitation in this work is that no data from a tracked vehicle is available. Making a qualitative verification impossible to preform. The environment is limited to two dimensions (x and y) in which the filter is working. The main reason for that is to keep it simple. However in this subject when it comes to active safety in a ground vehicle point of view, two dimensions is natural. In the way that all vehicles involved are moving in the same plane. Even though the road is hilly the vehicles are not flying around. No external information like maps and GPS have been used in the system. Only ego vehicle information from on-board sensors, and the laser and radar sensors via the CAN bus have been used. The radar is a major restriction on the functionality of the filter because it needs to lock on a target in order to get useful information from the laser. At least in the beginning of a tracking scenario. The radar can also be a limitation in the way that some pre-filtering is done before the information gets to the filter. Not necessarily when it comes to performance but because the author simply does not know what the pre-filter in the radar is doing. No multiple tracking has been considered in the work. Therefore no data association is needed when it comes to keep track of several objects. But it should not be difficult to extend the functionality to include this feature. The radar is the limit for multiple-tracking with the approach used in this work and it is able to keep track of four moving objects at the same time.

Chapter 2

The Filter

The filter used in a work like this must be able to predict a new state from the motion model as long as no new information is available. And fuse the new information from the sensor models in the best way to update the state. This is a perfect job for the Kalman filter which can handle both state estimation and sensor fusion at the same time. Throughout this chapter the double time index will be used. $(n|m)$ should be read "at time n given all information up to time m". $\hat{\mathbf{x}}(k|k-1)$ for example means the predicted state at time k given all measurements until time $k-1$.

2.1 The Linear Kalman Filter

The Kalman filter approach (KF) is a common and important method in signal processing. In fact the KF is the optimal filter in the linear case if both process noises and measurement noises are Gaussian. Assume that the following linear model is used:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) + \mathbf{w}(k) \\ \mathbf{z}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{v}(k) \end{aligned} \tag{2.1}$$

Where the noises are gaussian:

$$\begin{aligned} \mathbf{E}[\mathbf{w}(k)] &= 0 \\ \mathbf{E}[\mathbf{v}(k)] &= 0 \\ \mathbf{E}[\mathbf{w}(k)\mathbf{w}^T(k)] &= \mathbf{Q}_w(k) \\ \mathbf{E}[\mathbf{v}(k)\mathbf{v}^T(k)] &= \mathbf{Q}_v(k) \end{aligned} \tag{2.2}$$

And the covariance is defined as:

$$\begin{aligned} \mathbf{E}[\tilde{\mathbf{x}}(k|k)\tilde{\mathbf{x}}^T(k|k)] &= \mathbf{P}(k|k) \\ \mathbf{E}[\tilde{\mathbf{x}}(k+1|k)\tilde{\mathbf{x}}^T(k+1|k)] &= \mathbf{P}(k+1|k) \end{aligned} \tag{2.3}$$

Where the state error is described by the first row in equation 2.4. The expectation mean is given by the second row in equation 2.4.

$$\begin{aligned} \tilde{\mathbf{x}}(k|k) &= \mathbf{x}(k) - \hat{\mathbf{x}}(k|k) \\ \hat{x} &= \mathbf{E}[x] \end{aligned} \tag{2.4}$$

If the noises in equation 2.2 are uncorrelated the KF will become according to equations 2.5 and 2.6. As mentioned in the introduction the KF can be divided into two parts. One for the time update (prediction step):

$$\begin{aligned} \hat{\mathbf{x}}(k+1|k) &= \mathbf{A}\hat{\mathbf{x}}(k|k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{P}(k+1|k) &= \mathbf{A}\mathbf{P}(k|k)\mathbf{A}^T + \mathbf{Q}_w(k) \end{aligned} \tag{2.5}$$

And one for the measurement update (estimation step):

$$\begin{aligned} \mathbf{K}(k) &= \mathbf{P}(k|k-1)\mathbf{C}^T\mathbf{S}(k)^{-1} \\ \mathbf{S}(k) &= \mathbf{C}\mathbf{P}(k|k-1)\mathbf{C}^T + \mathbf{Q}_v(k) \\ \nu(k) &= \mathbf{z}(k) - \mathbf{C}\hat{\mathbf{x}}(k|k-1) \\ \hat{\mathbf{x}}(k|k) &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)\nu(k) \\ \mathbf{P}(k|k) &= \mathbf{P}(k|k-1) - \mathbf{K}(k)\mathbf{C}\mathbf{P}(k|k-1) \end{aligned} \tag{2.6}$$

Where ν is the new information from the measurements called the innovation and $\mathbf{S}$ is the innovation covariance. The $\mathbf{A}$ matrix in the model affects how the states propagates over time and the $\mathbf{B}$ matrix decides how the input signal should affect the current state. The $\mathbf{C}$ matrix will determine how the measurements affect the states. These matrices can also be time dependent. The process noise, $\mathbf{w}$ determines how the uncertainty grows over time. And the covariance matrix, $\mathbf{Q}_w$ has an impact in terms of how much random walk that will occur in the states. The measurement noise, $\mathbf{v}$ decides how much the uncertainty should shrink after a new measurement. Where the corresponding covariance matrix $\mathbf{Q}_v$ affects how much the filter will rely on the measurements.

2.2 The Extended Kalman Filter

When dealing with realistic models however some of the models tend to be non-linear which implies a need of a filter that can handle non-linear models. Also here a number of methods are available like the Extended Kalman Filter (EKF) and the Uncented Kalman Filter (UKF) and the Particle Filter (PF). The most important distinction between these methods is that the KF based approaches assumes state uncertainties with Gaussian distributions whereas the PF filter can handle multi-modal distributions. These methods and their cousins can be found in [4]. In this thesis the EKF has been used where the non-linear models are linearised around a new working point for each time sample. In this way the KF is applied locally with an approximately linear model. This is done by the first order Taylor expansion of the model around the current estimate. It is however

not an optimal filter anymore as in the linear case. And the noises still needs to be Gaussian for satisfying results. To assume Gaussian noise is not a dangerous move in this case when it comes to measurement noise. This time a non-linear model is used:

$$\begin{aligned} \mathbf{x}(k+1) &= f(\mathbf{x}(k), \mathbf{u}(k)) \\ \mathbf{z}(k) &= h(\mathbf{x}(k)) + \mathbf{v}(k) \end{aligned} \tag{2.7}$$

Where f is a non-linear model that can be dynamic and h is a non-linear measurement model. And the same assumptions are made about the noises as in the linear case where the added process noise, $\mathbf{w}$ is replaced and modelled by a driving noise in the input signal. This noise and the corresponding input signal error covariance, $\mathbf{Q}_u$ is described in equation 2.8.

$$\begin{aligned} \mathbf{E}[\tilde{\mathbf{u}}(k)] &= 0 \\ \mathbf{E}[\tilde{\mathbf{u}}(k)\tilde{\mathbf{u}}^T(k)] &= \mathbf{Q}_u(k) \\ \tilde{\mathbf{u}}(k) &= \mathbf{u}(k) - \hat{\mathbf{u}}(k|k) \end{aligned} \tag{2.8}$$

The same distinction with time and measurement update as in the linear case will be done here. The time update becomes:

$$\begin{aligned} \hat{\mathbf{x}}(k+1|k) &= f(\hat{\mathbf{x}}(k|k), \hat{\mathbf{u}}(k)) \\ \mathbf{P}(k+1|k) &= \mathbf{F}_x(k)\mathbf{P}(k|k)\mathbf{F}_x^T(k) + \mathbf{G}_u(k)\mathbf{Q}_u(k)\mathbf{G}_u^T(k) \end{aligned} \tag{2.9}$$

And the measurement update becomes:

$$\begin{aligned} \mathbf{K}(k) &= \mathbf{P}(k|k-1)\mathbf{H}_x^T(k)\mathbf{S}(k)^{-1} \\ \mathbf{S}(k) &= \mathbf{H}_x(k)\mathbf{P}(k|k-1)\mathbf{H}_x^T(k) + \mathbf{Q}_v(k) \\ \nu(k) &= \mathbf{z}(k) - h(\mathbf{x}(k|k-1)) \\ \hat{\mathbf{x}}(k|k) &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k)\nu(k) \\ \mathbf{P}(k|k) &= \mathbf{P}(k|k-1) - \mathbf{K}(k)\mathbf{H}_x(k)\mathbf{P}(k|k-1) \end{aligned} \tag{2.10}$$

Where $\mathbf{F}_x(k) = \nabla f_x$ and $\mathbf{G}_u(k) = \nabla f_u$ are Jacobians evaluated at $\mathbf{x}(k) = \hat{\mathbf{x}}(k|k)$ and $\mathbf{u}(k) = \hat{\mathbf{u}}(k)$ respectively. Whereas the Jacobian $\mathbf{H}_x(k) = \nabla h_x$ is evaluated at $\mathbf{x}(k) = \hat{\mathbf{x}}(k|k-1)$.

2.3 Application Specific Filter

In this thesis an EKF with the same structure as described in the foregoing section has been used. Here some additional explanation will be done in order to clarify how the filter was organised for the task in this work. The filter used in this thesis is described with the model in equation 2.11.

$$\begin{aligned} \mathbf{x}(k+1) &= f(\mathbf{x}(k), \mathbf{u}(k)) \\ \mathbf{z}_l(k) &= h_l(\mathbf{x}(k)) + \mathbf{v}(k) \\ \mathbf{z}_r(k) &= h_r(\mathbf{x}(k), \mathbf{u}(k)) + \mathbf{v}(k) \end{aligned} \tag{2.11}$$

However the innovation covariance, $\mathbf{S}$ in equation 2.10 is only valid for the laser model, h_l in equation 2.11. For the radar model, h_r the input vector dependency ($\mathbf{u}$) has to be taken care of. It is solved with an extra expression that consists of a Jacobian, $\mathbf{J}_u = \nabla h_u$ and the input signal covariance matrix, $\mathbf{Q}_u$. This innovation covariance will have the appearance in equation 2.12 and is further described in appendix A.

$$\mathbf{S}(k) = \mathbf{H}_x(k)\mathbf{P}(k|k-1)\mathbf{H}_x^T(k) + \mathbf{J}_u(k)\mathbf{Q}_u(k)\mathbf{J}_u^T(k) + \mathbf{Q}_v(k) \tag{2.12}$$

An overview of the structure of the filter is presented in figure 2.1. Where the block $f(\mathbf{x}, \mathbf{u})$ is the motion model (process model) described in section 3.1. $g_l(\mathbf{x}, \mathbf{Z}_l)$ is a filter that converts the raw measurements from the laser, $\mathbf{Z}_l$ to observations, $\mathbf{z}_l$ described in section 3.2.3. The blocks $h_l(\mathbf{x})$ and $h_r(\mathbf{x}, \mathbf{u})$ are the sensor models (observation models) for the laser and the radar respectively described in sections 3.2.4 and 3.2.7. The states and input signals that have been used throughout this work looks as in equation 2.13.

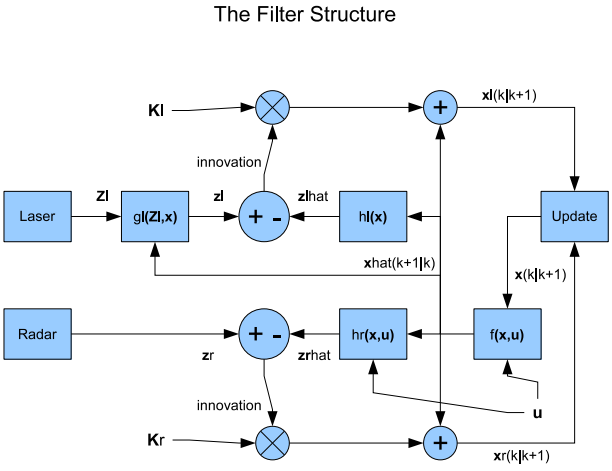


Figure 2.1: Overview of the structure of the filter.

$$\begin{aligned} \mathbf{x} &= (x, y, \phi, v_o, \omega_o, W) \\ \mathbf{u} &= (v_e, \omega_e, dv_o, d\omega_o) \end{aligned} \tag{2.13}$$

Where x is the relative distance in the longitudinal direction seen from the ego vehicle. Whereas y is the relative distance in lateral direction seen from the ego vehicle. A positive y means that the tracked vehicle is to the left of the ego vehicle. The relative orientation, ϕ is the angle between the vehicles driving directions. v_o is the velocity of the tracked vehicle in its direction of travel. The yaw rate of the

tracked vehicle, ω_o is the angular velocity about the vertical axis. And W is the width of the tracked vehicle. v_e and ω_e is the velocity and yaw rate for the ego vehicle. dv_o and $d\omega_o$ should be seen as change of velocity and angular velocity over one discrete time instant. They are however always zero in this work. The reason for having these input signals though they are zero is to affect the motion model of the tracked vehicle with the corresponding noises. In figure 2.2 an example of a situation is presented to illustrate how the state and input signals are defined.

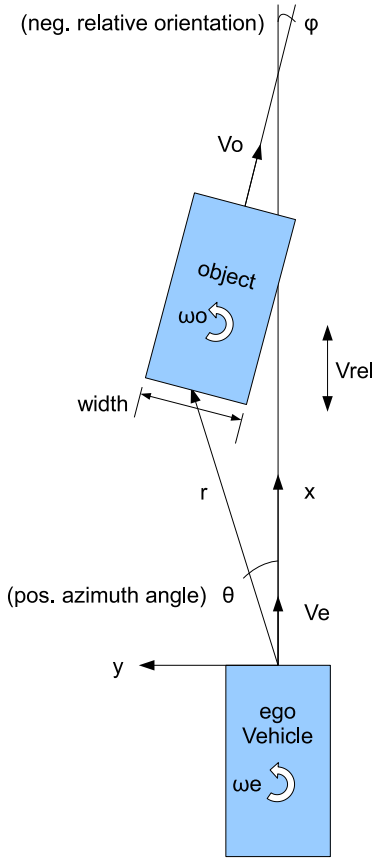


Figure 2.2: Overview of how the states and some of the input signals are defined.

Chapter 3

The Models

There are two types of models that are needed in a work of this type. Motion models are used for prediction. I.e. to predict how the states propagates over time. The sensor models are used to determine how new measurements should affect the states.

3.1 Motion Models

There are two motions that needs to be modelled, the ego vehicle motion and the motion of the tracked vehicle. In this work the same model has been used to describe both motions with different parameter settings. The dynamics of the vehicles are assumed to be low and therefore a kinematic model is used. The ego vehicle motion is driven by measured speed and yaw rate whereas the object motion is driven by constant acceleration to make it more robust. These individual motions are then put in a framework to describe the relationship between them.

The first section, 3.1.1 explains the difference between holonoimic and nonholonomic systems. In section 3.1.2 the Constant Turn model (CT model) is presented that has been used to describe the individual motions of the vehicles. This model is derived in appendix A. And in the next section, 3.1.4 a framework for the two vehicular models is presented. These relationships are called spatial relationships or stochastic relationships and are more described in appendix B.

3.1.1 Holonomic and Nonholonomic Systems

The motion models in this thesis will describe a so called nonholonomic system. Which basically means that the states are historically dependant of the control input, apart from a holonomic system that is only dependant on the current input signal.

Constraint equations that only involve relative positions of points in a mechanical system is said to be holonomic, and the associated mechanical system is said to

be a holonomic system. On the other hand if a system has constraint equations including velocities, accelerations, or derivatives of system coordinates. The constraint equations are said to be nonholonomic, and the mechanical system is said to be a nonholonomic system. Suppose that the constraint equation defining a mechanical system has the form as in equation. 3.1.

$$f(c_r, t) = 0 \tag{3.1}$$

where c_r $(r = 1, ..., n)$ are coordinates of the system (n is the number of degrees of freedom of the unrestrained system). Such a constraint is said to be holonomic (or geometric). Whereas if the constraint equation has the form as in equation 3.2 the constraint is said to be nonholonomic (or kinematic).

$$f(c_r, \dot{c}_r, \ddot{c}_r, ..., t) = 0 \tag{3.2}$$

I.e. a mechanical system with only holonomic constraints gives a holonomic system but if at least one constraint is nonholonomic it gives a nonholonomic system. The fact that holonomic systems produce algebraic constraint equations, whereas nonholonomic systems produce differential constraint equations. Give rise to a number of implications for a nonholonomic system such as error propagation for example. For a filter using the CT model presented in the next section or any other model describing a nonholonomic system. This means that the uncertainty of the current state will grow as long as no new measurements are available. More information about mechanical systems can be found in [7].

3.1.2 Constant Turn model with Polar Velocity

As mentioned earlier the same kinematic motion model has been used to describe the motion of the ego vehicle as well as the motion of the tracked vehicle. The model has been derived in [1] and has originally been used to describe motions of moving robots. It is based on the assumption that the vehicle is moving along a fixed curvature. I.e. it can move straight or turn with a certain constant angular velocity. Though this curvature is recalculated for each time step in the filter, the point is that the modelled vehicle is prevented from making moves in lateral direction. Following state and input vectors are used in the CT model.

$$\mathbf{x}(k) = [x(k), y(k), \phi(k)]^T \tag{3.3}$$

$$\mathbf{u}(k) = [v(k), \omega(k)]^T \tag{3.4}$$

The following equations describe an exact discrete constant turn model not dependent of the sampling rate. The state transition is linear in $\mathbf{U}(k)$ but not in the state $\mathbf{x}(k)$.

$$\begin{aligned} \mathbf{x}(k+1) &= f_{CT}(\mathbf{x}(k), \mathbf{u}(k)) \\ f_{CT}(\mathbf{x}(k), \mathbf{u}(k)) &= \mathbf{x}(k) + \mathbf{R}(k)\mathbf{U}(k) \end{aligned} \tag{3.5}$$

$$\mathbf{R}(k) = \begin{bmatrix} \cos(\phi(k)) & -\sin(\phi(k)) & 0 \\ \sin(\phi(k)) & \cos(\phi(k)) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{3.6}$$

For a given sample interval T the driving velocities, v and ω will correspond to a travelled distance, x and y described by the two first rows in equation 3.7. Where the rotated angle corresponds to the last row in equation 3.7.

$$\mathbf{U}(k) = \begin{bmatrix} \frac{v(k)}{\omega(k)} \sin(T\omega(k)) \\ \frac{v(k)}{\omega(k)}(1 - \cos(T\omega(k)) \\ T\omega(k) \end{bmatrix} \tag{3.7}$$

With this model the covariance matrix describing the uncertainty of the states will become according to equation 3.8.

$$\mathbf{P}(k+1|k) = \mathbf{F}_x(k)\mathbf{P}(k|k)\mathbf{F}_x^T(k) + \mathbf{G}_u(k)\mathbf{Q}_u(k)\mathbf{G}_u^T(k) \tag{3.8}$$

Where the Jacobians with respect to state and input vectors becomes according to equation 3.9 and 3.10 respectively.

$$\mathbf{F}_x(k) = \begin{bmatrix} 1 & 0 & \frac{\hat{v}(k)(\cos(\hat{\phi}(k|k)+T\hat{\omega}(k))-\cos(\hat{\phi}(k|k))}{\hat{\omega}(k)} \\ 0 & 1 & \frac{\hat{v}(k)(\sin(\hat{\phi}(k|k)+T\hat{\omega}(k))-\sin(\hat{\phi}(k|k))}{\hat{\omega}(k)} \\ 0 & 0 & 1 \end{bmatrix} \tag{3.9}$$

$$\mathbf{G}_u(k) = \mathbf{R}(k) \begin{bmatrix} \frac{\sin(T\hat{\omega}(k))}{\hat{\omega}(k)} & \frac{\hat{v}(k)(T\hat{\omega}(k)\cos(T\hat{\omega}(k))-\sin(T\hat{\omega}(k))}{\hat{\omega}^2(k)} \\ \frac{1-\cos(T\hat{\omega}(k))}{\hat{\omega}(k)} & \frac{\hat{v}(k)(\cos(T\hat{\omega}(k))-1)+\hat{\omega}(k)T\sin(T\hat{\omega}(k))}{\hat{\omega}^2(k)} \\ 0 & T \end{bmatrix} \tag{3.10}$$

The covariance matrix for the driving input signals becomes according to equation 3.11. Where the input signals are modelled as zero mean, uncorrelated, white sequences with standard deviations σ_v and σ_ω.

$$\mathbf{Q}_u = \begin{bmatrix} \sigma_v^2 & 0 \\ 0 & \sigma_\omega^2 \end{bmatrix} \tag{3.11}$$

The model is very attractive when dealing with irregular sampling due to the exact discretisation. This has also been the case in this work were all available data has been sampled with irregular time steps. Similar and other models can be found in [9].

3.1.3 Constant Turn model with lateral slip included

If there is information about the sideways motion of the vehicle, v_{lat} then lateral slip can be added in the CT model. However lateral slip has not been used in the model for this work. With lateral slip included the new input signal vector will become according to equation 3.12.

$$\mathbf{u}_{lat}(k) = [v(k), v_{lat}(k), \omega(k)]^T \tag{3.12}$$

Where the corresponding pose change, $\mathbf{U}_{lat}$ still is kinematically correct in equation 3.13.

$$\mathbf{U}_{lat}(k) = \begin{bmatrix} \frac{v(k)}{\omega(k)}\sin(T\omega(k)) - \frac{v_{lat}(k)}{\omega(k)}(1-\cos(T\omega(k)) \\ \frac{v(k)}{\omega(k)}(1-\cos(T\omega(k)) + \frac{v_{lat}(k)}{\omega(k)}\sin(T\omega(k)) \\ T\omega(k) \end{bmatrix} \tag{3.13}$$

With the new input added a new Jacobian, $\mathbf{G}_{u_{lat}}$ is needed as well as a new input noise covariance matrix, $\mathbf{Q}_{u_{lat}}$. They will become as in equation 3.14 and 3.15 respectively.

$$\mathbf{G}_{u_{lat}}(k) = \mathbf{R}(k) \begin{bmatrix} \frac{\sin(T\hat{\omega}(k))}{\hat{\omega}(k)} & \frac{\cos(T\hat{\omega}(k))-1}{\hat{\omega}(k)} & \frac{\hat{v}(k)(T\hat{\omega}(k)\cos(T\hat{\omega}(k))-\sin(T\hat{\omega}(k))}{\hat{\omega}^2(k)} \\ \frac{1-\cos(T\hat{\omega}(k))}{\hat{\omega}(k)} & \frac{\sin(T\hat{\omega}(k))}{\hat{\omega}(k)} & \frac{\hat{v}(k)(\cos(T\hat{\omega}(k))-1)+\hat{\omega}(k)T\sin(T\hat{\omega}(k))}{\hat{\omega}^2(k)} \\ 0 & 0 & T \end{bmatrix} \tag{3.14}$$

$$\mathbf{Q}_{u_{lat}}(k) = \begin{bmatrix} \sigma_v^2 & 0 & 0 \\ 0 & \sigma_{v_{lat}}^2 & 0 \\ 0 & 0 & \sigma_\omega^2 \end{bmatrix} \tag{3.15}$$

The new covariance update for the added lateral slip will become according to equation 3.16. Where the second part is changed but the first part is the same as for the case without lateral slip.

$$\mathbf{P}(k+1|k) = \mathbf{F}_x(k)\mathbf{P}(k|k)\mathbf{F}_x^T(k) + \mathbf{G}_{u_{lat}}(k)\mathbf{Q}_{u_{lat}}(k)\mathbf{G}_{u_{lat}}^T(k) \tag{3.16}$$

3.1.4 Motion Model with Spatial Relationships

With the CT model mentioned earlier a model that describes the relationship between the ego vehicle and tracked vehicle can be derived. The ego motion is feed with state and input signals in equation 3.17. The input signal comes directly from the vehicles built in sensors via the CAN network in the ego vehicle. And the state vector is always reset to zero making the coordinate system to be attached to the ego vehicle.

$$\begin{aligned} \mathbf{x}_e(k) &= [0,0,0]^T \\ \mathbf{u}_e(k) &= [v_e(k), \omega_e(k)]^T \end{aligned} \tag{3.17}$$

Note that though $\mathbf{x}_e(k)$ always starts with zero it will not necessary produce a prediction that is zero. I.e. $\mathbf{x}_e(k+1) \neq [0,0,0]^T$ when $\mathbf{u}_e(k) \neq [0,0]^T$. The motion of the tracked vehicle is described with measured relative position and orientation. The state vector and input signals that feeds the model looks as in equation 3.18.

$$\begin{aligned} \mathbf{x}_o(k) &= \mathbf{x}(k) = [x(k), y(k), \phi(k)]^T \\ \mathbf{u}_o(k) &= [v_o(k) + Tdv_o(k), \omega_o(k) + Td\omega_o(k)]^T \end{aligned} \tag{3.18}$$

The state vector $\mathbf{x}_e(k+1)$ describing ego vehicle motion and $\mathbf{x}_o(k+1)$ describing tracked vehicle, can be seen as inner states of a model that handles the spatial relationships between them. But we need another notation to fully understand how the inner states affect the spatial relationships. Let $X_{ab} = [x_{ab}, y_{ab}, \phi_{ab}]^T$ denote a spatial relationship between point a and point b. With the nomenclature below it would mean a move of the ego vehicle.

a:	ego vehicle start position at time k
b:	ego vehicle end position at time k+1
c:	tracked vehicle start position at time k
d:	tracked vehicle end position at time k+1

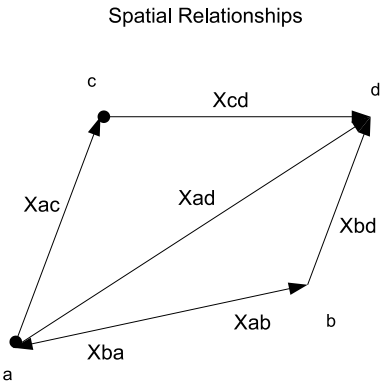


Figure 3.1: The spatial relationships in the example.

Figure 3.1 illustrates the spatial relationships in the example. The relative state at time k, $\mathbf{x}(k)$ describing the relationship between the motion of the ego vehicle and tracked vehicle corresponds to X_{ac}. And the corresponding predicted state for time k+1, $\mathbf{x}(k+1)$ that we wish to derive will be X_{bd}. The relationship X_{ab} is the predicted move of the ego vehicle with the CT model when expressed in coordinates of point a. And X_{ad} is the move of the tracked vehicle from the CT model when expressed in coordinates of point a. Note that X_{ab} corresponds to $\mathbf{x}_e(k+1)$ and X_{ad} corresponds to $\mathbf{x}_o(k+1)$ (seen from point a). X_{cd} would be the predicted motion of the tracked vehicle from its point of view. To get the relationship X_{bd} we need two spatial operations described in appendix B. The uncertainty in point b needs to be moved and added with the uncertainty in point d. This is done by first using the inverse relationship operator ($\ominus$) on X_{ab}. And then using the result with the compounding operator ($\oplus$) together with X_{ad}. The individual motions described by the CT model and the relationship between them

is described by equations 3.19 and 3.20.

$$\begin{aligned} X_{ab} &= f_{CT}(\mathbf{x}_e(k), \mathbf{u}_e(k)) \\ X_{ad} &= f_{CT}(\mathbf{x}_o(k), \mathbf{u}_o(k)) \\ X_{bd} &= \mathbf{x}(k+1) \end{aligned} \tag{3.19}$$

$$X_{bd} = \ominus X_{ab} \oplus X_{ad} = X_{ba} \oplus X_{ad} \tag{3.20}$$

The final result will be X_{bd} where all uncertainty is in point d. But instead of getting the corresponding uncertainties for each model (equation 3.8) and then using the covariance part of the spatial operators described in appendix B. The matrices F and G have been derived directly with the Jacobians for the whole model as in equation 3.21. They are however far more complex than for the CT model in the equations 3.9, 3.10 and will therefore not be shown explicit here. The input signals are modelled as white noises like for the CT model which gives the new extended covariance matrix in equation 3.22.

$$\begin{aligned} \mathbf{F} &= \frac{\partial X_{bd}}{\partial \mathbf{x}} \\ \mathbf{G} &= \frac{\partial X_{bd}}{\partial \mathbf{u}} \end{aligned} \tag{3.21}$$

$$\mathbf{Q}_u = \begin{bmatrix} \sigma^2_{v_e} & 0 & 0 & 0 \\ 0 & \sigma^2_{\omega_e} & 0 & 0 \\ 0 & 0 & \sigma^2_{dv_o} & 0 \\ 0 & 0 & 0 & \sigma^2_{d\omega_o} \end{bmatrix} \tag{3.22}$$

After the states v_o, ω_o and *width* have been added the state and input vectors of the filter will have the following appearance:

$$\mathbf{x}(k) = [x(k), y(k), \phi(k), v_o(k), \omega_o(k), W(k)]^T \tag{3.23}$$

$$\mathbf{u}(k) = [v_e(k), \omega_e(k), dv_o(k), d\omega_o(k)]^T \tag{3.24}$$

3.2 Sensor Models

The sensors dealt with in this section are so called exteroceptive sensors meaning that they measures the surroundings of the ego vehicle. In contrast to proprioceptive sensors that measures how the vehicle itself behaves. Measurements done by the later sensor type is typically input signals to the motion models. Both the laser and the radar make measurements of positions in polar coordinates. And the range rate measured by the radar is observed in the direction that the ego vehicle is heading for. There are two parts for describing each sensor in the filter. Both these parts produce an observation but in two different ways. The measurement part is translating the measurements of the reality to a useful observation, $\mathbf{z}$. The laser scanner for example measures several distances, $R = r_1 ... r_N$ for a number of

angles, $\Theta = \theta_1 ... \theta_N$. That are filtered to one position, (r, θ) and thus an observation of where the vehicle is. The other part is predicting an observation, $\hat{\mathbf{z}}$ of the position from the states with the model of the sensor. Equation 3.25 describes the relation between these two parts called the innovation, ν. The innovation is used to calculate the total uncertainty of the observation, including both errors from the measurements and the errors from the prediction. It should be read as the new information that is available from a certain measurement.

$$\nu = \mathbf{z} - \hat{\mathbf{z}} \tag{3.25}$$

3.2.1 Laser Scanner

The laser sensor that has been used in this thesis is a laser range scanner based on Light Detection And Ranging (LIDAR) technology. A more detailed description of it will come in section 4.1. The observation made from the measurements and the predicted observation will be described in the following sections. The laser measurements are basically done in three steps explained in the second section (3.2.3). The predicted observation is shortly described in the next section (3.2.4). But first comes an explanation of the hough transform which is important for the understanding of how the laser measurements are transformed into an observation.

3.2.2 The Hough Transform

The basic idea of the hough transform that is used in this thesis is also the classical purpose with this transform. Namely the problem of detecting lines in an image, or more general for edge detection. There are several methods for solving this problem but most of them requires at least some knowledge about pixels belonging to a certain object in a local environment. Here we only know that we are looking for a scattered line in the global environment. A simple solution would be to find all lines defined by all combinations of pixel pairs and then find all subsets of points that are close to particular lines. This gives however to much computational workload even for a relative small amount of data, coming from a laser scanner. The hough transform on the other hand is a model-based segmentation method that uses smart parametrisation to solve the problem. The basic idea of how the hough transform has been used in this work will be described with an example where two points, A and B are transformed to a line that intersects both points. More about the hough transform and image processing can be found in [3].

Consider a point $A = (x_i, y_i)$ in the xy-plane and the equation of a straight line. There are infinitely many lines passing through A, and all of them satisfies the equation 3.26. But if one consider the km-plane instead (the parameter space) with $x = x_i$ and $y = y_i$, then equation 3.26 yields a single line. And if we consider another point $B = (x_j, y_j)$ in the xy-plane, it too will give a line in the km-plane. Now these two lines in the km-plane will intersect, unless they are parallel. If we assume that the two points A and B in the xy-plane yields lines that intersect in a point (k', m') in the km-plane. Then all the points in the xy-plane between A and B will give lines that intersects the point (k', m') in the km-plane. With this

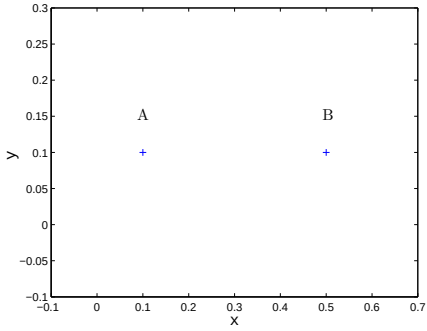


Figure 3.2: Two points A and B in the xy-plane.

approach all points in the xy-plane would correspond to a line in the km-plane. And we could detect lines (points in a row) in the xy-plane by detecting points (intersections of lines) in the km-plane. Note that in practice, detecting points in the km-plane means detect local maximas in the km-plane due to parametrisation. With straight line transformation in equation 3.26 the two points in figure 3.2 will become two lines that intersects as in figure 3.3.

$$y = kx + m \tag{3.26}$$

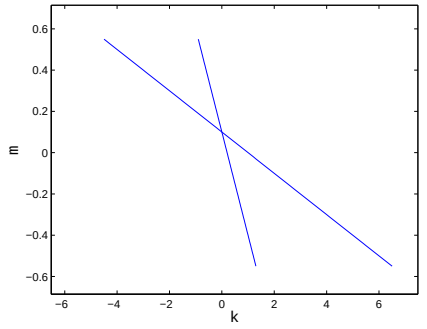


Figure 3.3: A and B transformed to the km-plane.

This approach is suitable for explanation but suffers from limitations when used in practice. One reason is that the slope approaches infinity as the line approaches the vertical axis in the xy-plane ($k \to \infty$), causing problems in the km-plane. Another reason is the non-linear discretisation of k. If the line is represented as in eq. 3.27 we have a hough transform that does not suffer from these limitations. The only thing that needs to be considered here is the discretisation of the parameter space

($\theta\rho$-plane). Which of course will be a trade-off between workload and precision.

$$\rho = x\cos(\theta) + y\sin(\theta) \tag{3.27}$$

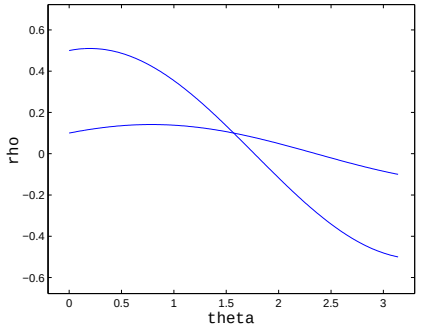


Figure 3.4: A and B transformed to the $\theta\rho$-plane.

When the intersection in the $\theta\rho$-plane in figure 3.4 is transformed back to the xy-plane with equation 3.27. It will end up as a line going through the points A and B as shown in figure 3.5.

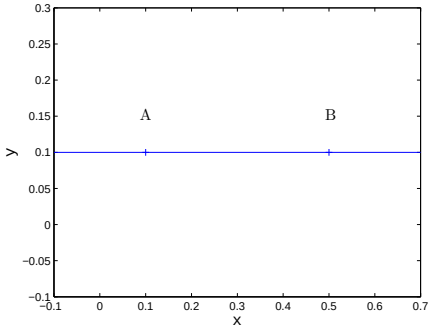


Figure 3.5: Intersection in the $\theta\rho$-plane transformed back to the xy-plane.

Another benefit using the hough transform is that a natural pre-filtering is already done in this case when dealing with laser scanner data. When using the hough transform in image processing one has to do some thresholding in order do get a series of dots where the lines where, to be able to apply the hough transform. Otherwise there would be to many hough transformations to be an attractive method.

3.2.3 Laser Measurements

To illustrate how the measurements from the laser are transformed into an observation, an example is shown that is taken from one time step of the data used in this work. At first the laser hits close to the position of the tracked vehicle looks as in figure 3.6. In this example the tracked vehicle is close to zero meters in y direction. Where the back of the vehicle is close to 37 meters in x direction. I.e. the tracked vehicle is driving approximately 37 meters in front of the ego vehicle. In direction from left to right.

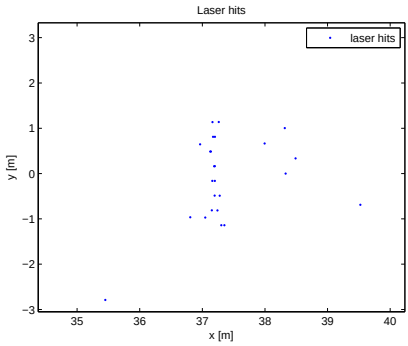


Figure 3.6: Laser hits before filtering.

In the first step the predicted position is used to place a window that captures the most interesting laser hits. The window size depends on the standard deviation in x and y direction. The window size in y direction also depends on the width of the vehicle in front. A similar approach can be found in [14] called range weighted hough transform (RWHT). In figure 3.7 the window is shown and the laser hits inside it is displayed with a cross.

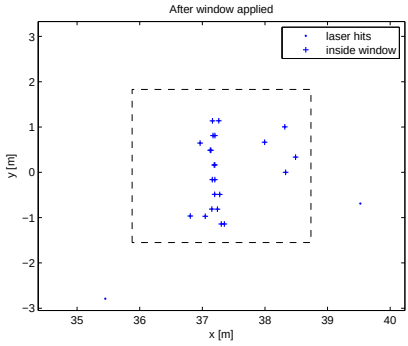


Figure 3.7: After window applied on laser hits.

The second step contains the hough transform applied on the set of hits resulted from step one. Ideally the hough transform would result in a maximum number of hits for a certain bin and angle. In this thesis however this were not the case most times. Instead a method of using the union set of points for several bins and angles were used. This works well for this application as long as the resolution for the bins and angles are high enough. After step two it is determined which points are to be the ones that will define the vehicle. This means that it is possible to calculate the relative position and the width. In figure 3.8 the laser hits that has been filtered out with the hough transform are displayed with rings.

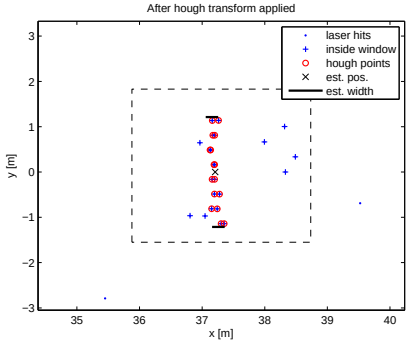


Figure 3.8: After hough transform applied on laser hits inside window.

The final step is a basic least square fit operation done on the set of hough points from step two. From the fitted line a slope is given which is the measured relative orientation. In figure 3.9 the estimated slope is shown as a line.

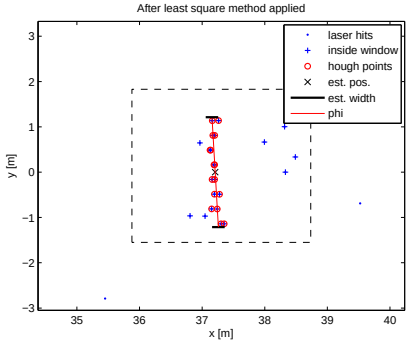


Figure 3.9: After least square fit operation applied on hough points.

In this case the two hits that are filtered out in step one are probably ground hits or some other spuriousness. From the hits taken out by the hough step, the four to

the right are probably hits under the vehicle reflected by the ground. However the three to the left are more tricky. They could either be ground hits or be a result of a rough surface of the vehicle, like a bumper for example. Even for the hits considered to be a part of the vehicle some irregular pattern can be distinguished. This is due to that the hits corresponds to different layers of the laser scanner on the vertical axis. After the three steps the laser measurement has produced an observation according to equation 3.28.

$$\mathbf{z}_l(k) = [r(k), \theta(k), \phi(k), W(k)]^T \tag{3.28}$$

This observation can be expressed as a filter g_l dependent on both the current state $\mathbf{x}(k)$ and the raw measurements $\mathbf{Z}_l(k)$.

$$\mathbf{z}_l(k) = g_l(\mathbf{x}(k), \mathbf{Z}_l(k)) \tag{3.29}$$

3.2.4 Laser Observation Model

The part of the observation model that predicts an observation from the states is modelled as follows:

$$\hat{\mathbf{z}}_l(k) = h_l(\hat{\mathbf{x}}(k|k-1)) \tag{3.30}$$

Where h_l is a function of the states. This model for the prediction is described by equation 3.31.

$$\mathbf{z}_l(k) = h_l(\mathbf{x}(k)) + \mathbf{v}(k) \tag{3.31}$$

The predicted observation for the laser is straightforward but not trivial because a transformation from Cartesian to polar coordinates is needed. The laser scanner measures in polar coordinates and the relative position in the state vector is represented in Cartesian coordinates.

$$\begin{aligned} \hat{r} &= \sqrt{\hat{x}^2 + \hat{y}^2} \\ \hat{\theta} &= \tan^{-1} \frac{\hat{y}}{\hat{x}} \end{aligned} \tag{3.32}$$

The predictions of ϕ and *width* are however trivial since they are states in the filter and will therefore pass through the model unprocessed.

3.2.5 Radar

The radar that has been used in this thesis is a Long Range Radar. It will be described further in section 4.1. Section 3.2.6 describes how the measurement observation, $\mathbf{z}$ is derived. Whereas the last section describes how the predicted observation, $\hat{\mathbf{z}}$ is derived.

3.2.6 Radar Measurements

In the case of the measurements from the radar, $\mathbf{z}_r$ is trivial to derive since the measurements done by the radar corresponds directly to an observation. I.e. the observation and the measurements both include the same quantities as in equation 3.33. Of course some pre-filtering is performed in the radar to give this observation. This is though the data that is interesting for a work like this. The raw data is also available from the sensor but then a lot more work has to be made on getting a useful observation.

$$\mathbf{z}_r(k) = [r(k), v_{rel}(k), \theta(k)]^T \tag{3.33}$$

3.2.7 Radar Observation Model

The model h_r for predicting a radar observation is similar to the one used for the laser (h_l). But for the radar case the model also depends on the input signal $\mathbf{u}$. The model is described in equation 3.35.

$$\mathbf{z}_r(k) = h_r(\mathbf{x}(k), \mathbf{u}(k)) + \mathbf{v}(k) \tag{3.34}$$

And the prediction can be written as:

$$\hat{\mathbf{z}}_l(k) = h_r(\hat{\mathbf{x}}(k|k-1), \hat{\mathbf{u}}(k)) \tag{3.35}$$

The predictions of r and θ are done in the same way as for the laser with equation 3.32. The prediction of v_{rel} is more complicated since the radar measures the relative speed in the coordinate system of the ego vehicle. Therefore a transformation from the tracked vehicles coordinate system to the ego vehicle coordinate system is needed. This is performed according to equation 3.36.

$$\hat{v}_{rel} = \cos\hat{\theta}(\hat{v}_o - T\hat{dv}_o) - \hat{v}_e \tag{3.36}$$

Where v_o is the fourth state and θ is derived from the first two states via equation 3.32. Whereas dv_o and v_e is coming from the input signal vector $\mathbf{u}$. Even though dv_o has been zero in this work the need of v_e still makes the model a hybrid dependent on both state and input signals.

Note that in the model used in this work for the radar, the measured speed, v_{rel} is assumed to be $\dot{x}_{rel}$. I.e. the relative speed is measured in the current direction that the ego vehicle is heading for. But if the radar measures the relative speed in the same direction as the distance, $v_{rel} = \dot{r}_{rel}$ then the model will become according to equation 3.37.

$$\hat{v}_{rel} = \cos\hat{\theta}(\cos\hat{\phi}(\hat{v}_o - T\hat{dv}_o) - \hat{v}_e) - \sin\hat{\theta}\sin\hat{\phi}(\hat{v}_o - T\hat{dv}_o) \tag{3.37}$$

Chapter 4

Concluding Remarks

4.1 Experiments

The results in this work is the fused data resulting from sensor data recorded on a Scania truck. Especially one set of data has been studied. The scenario is a test track at Scaina CV in Södertälje. Were the tracked vehicle is a bus driving ahead of the truck. The width of the bus is 2.55 meters. The vehicles are driving in the same direction but the distance between them differs during the run. Both vehicles are driving in a speed range of about 40-60 km/h (see figure 4.7). The weather is clear and the road is made of regular asphalt. The laser sensor is placed almost in the middle of the front and approximately 70 cm above ground. Whereas the radar sits 22 cm to the right and 13 cm above the laser seen from the front of the truck. The 22 cm horizontal difference between the sensors is compensated for in the filter by a shift of 22 cm in y-direction of the radar measurements. In this way the observations from both sensors seems to come from the position where the laser is placed. If the filter is extended to three dimensions there will still not be a need of a shift in the vertical direction since the radar does not give measurements with information about this dimension. The measurements from the sensors have been used to make observations in the filter.

Figure 4.1: The placement of the laser (red) and the radar (blue) on the truck.

The laser sensor that has been used here is a ibeo LUX® laser range scanner. It is based on Light Detection And Ranging (LIDAR) technology. The laser scans the environment with several rotating laser beams and calculates the time-of-flight of the received echoes. As mentioned in the background section the laser is working with direct detection. This means that it is not able to measure the range rate which is the case for a coherent detection laser. The data from this laser scanner consists of distance, angle and echo pulse width information. It can also detect up to three echoes per transmitted laser pulse and uses four scan levels separated with 0.8 degrees in vertical direction. In the xy-plane (horizontal plane) the central working range reaches from $+35\,^\circ$ to $-50\,^\circ$ using all four scan levels. But in the lateral working range from $+50\,^\circ$ to $+35\,^\circ$ and from $-50\,^\circ$ to $-60\,^\circ$ only two scan levels are being used. In figure 4.2 the working range of the laser scanner is illustrated. There are several working modes for this sensor when it comes to scan frequencies and angular resolution. The data that has been recorded for this thesis has been recorded with a scan frequency of 25 Hz and a constant angular resolution of 0.25 degrees in the whole working range. Comparisons between other laser scanners can be found in [2] and [8].

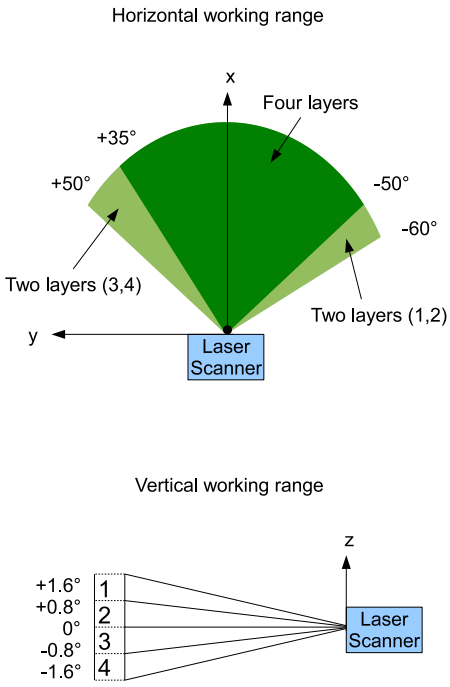


Figure 4.2: The working range of the laser scanner.

The radar that has been used in this thesis is a Long Range Radar from Knorr-Bremse. It is working with frequencies around 77 GHz. The main difference in properties of the radar relative to the laser scanner (seen from table 1.1) is that the radar measures the range rate very well but the laser does not. The radar also measures azimuth angle less accurate than the laser. Though this property has been difficult to see in this work due to the heavy pre-filtering in the radar. As mentioned earlier in section 1.5 this thesis has been limited to situations where the radar has locked on a target.

Some additional data has been recorded via the CAN network from the available standard on-board sensors on the truck. The yaw rate, ω_e and the speed at the frontal wheel axis, v_e. These quantities from the ego vehicle has been used as input signals in the filter.

The data logs have been recorded with a program called RTMaps. Where the information from the radar and the ego vehicle have been logged via the CAN network of the ego vehicle. This data has been recorded together with the information from the laser. The recorded data has been saved in a format that Matlab can handle. The filter has been implemented and simulated in Matlab. Where the raw data from all sensors were tagged with time tics. I.e. one time tag for each data that has been measured and saved at a specific time instance with a resolution of microseconds. These time tics have been rounded to tens of milliseconds in order to get feasible time length when looping trough the whole scenario.

The filter loop checks for eventual new data every tenth millisecond. When there is any data with a matching time tic this data is used to update the current state. If the data comes from the laser or radar the corresponding sensor model is used to make a measurement update. Or if the data comes from the internal sensors the input data vector is updated. The later data type is however not necessarily used. I.e. if the input data vector is updated again before any time update has been made. Every time new data is available from the radar or the laser a time update is made with the time elapsed from the last measurement until the soon to be made measurement. I.e. a prediction is made by the motion model based on the current state and the elapsed time between the measurements. In figure 4.3 the filter loop is illustrated. Also some pseudo code is presented to explain the blocks in the figure.

Input Update:

```
u = [ve,we,0,0];
```

Time Update:

```
[x,p] = time_update(x,p,u,T,...);
```

Radar Update:

```
zr = [r,v_rel,theta];
[x,p] = radar_update(x,p,u,zr,...);
```

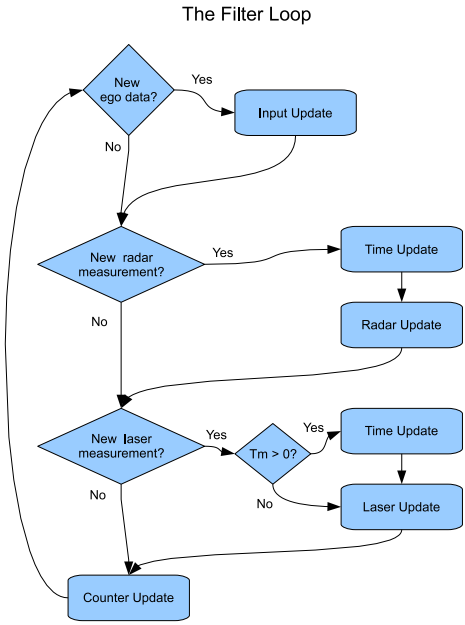


Figure 4.3: The implementation of the filter.

Laser Update:

```
zl = laser_measurement_update(r,theta,x,p);
[x,p] = radar_update(x,p,zl,...);
```

Counter Update:

```
Ttot = Ttot + Ts;
```

New data/measurement?:

```
if (Ttot == time_tic); ...;
```

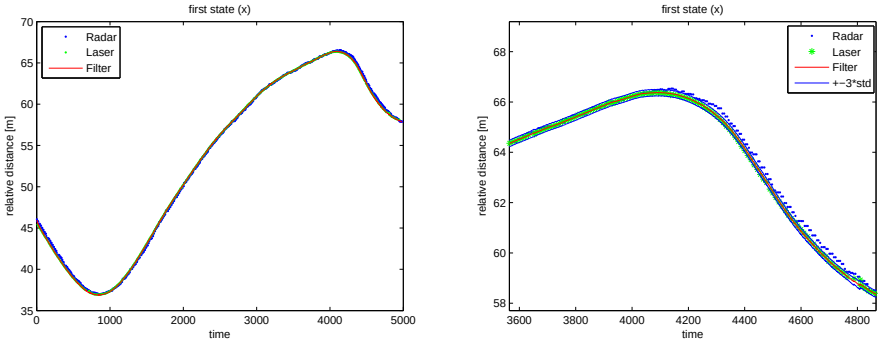
Where x is the state vector and p is the covariance matrix. Ttot is the total time that is compared with the time tics associated with the corresponding data for that time instant. In this work Ts is 10 ms. The Tm block is for preventing a time update of the model (Time Update). If there are both radar and laser data for the same time in the loop. I.e. if there is radar data for a certain time in the loop the Time Update will make an update of the model for the elapsed time since the latest measurement where done. But if there also is laser data available for the same time step in the loop the time elapsed between the measurements will be zero. And therefore must not a Time Update be performed. On the other hand if the only available data is coming from the laser for a certain time in the loop a time update has to be made for the model.

4.2 Results

The results presented here comes from the scenario described in the beginning of section 4.1. Note that the filter will be presented by both the predictions and the measurement updates. In the following text and figures each state in the state vector will be presented along with the observations made by the sensors. Except the speed and the yaw rate of the tracked vehicle, v_o and ω_o who are not observed by any sensor. Instead it is compared with the yaw rate of the ego vehicle, ω_e in figure 4.8. The state vector looks like in equation 4.1

$$\mathbf{x} = [x, y, \phi, v_o, \omega_o, W] \tag{4.1}$$

The first state is the relative distance between the vehicles, x defined in longitudinal direction seen from the ego vehicle. In figure 4.4a the result for the whole run is presented. In figure 4.4b a zoomed version is shown.



(a) (b)

Figure 4.4: Relative distance in longitudinal direction from observations and filter.

In figure 4.5 the second state, y is presented. It is the relative distance in lateral direction seen from the ego vehicle. Where positive values correspond to left. I.e. in the middle of the sequence the tracked vehicle is around one and a half meter to the left of the ego vehicle.

In figure 4.6 the relative orientation, ϕ is displayed. I.e. how the vehicles are oriented relative each other. If the tracked vehicle is driving in front of the ego vehicle as in this scenario and the tracked vehicle decides to turn left, then the relative orientation will become positive. In other words it has the same sign convention as the lateral relative distance (y).

The fourth state is shown in figure 4.7. This is the velocity of the tracked vehicle, v_o. For comparison the velocity of the ego vehicle, v_e is shown as well. However the data presented as v_o "hybrid data" in the figure. Is not just observations because it depends on the state vector as well as the radar observations of the relative

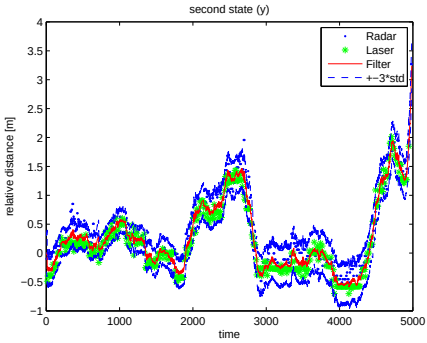


Figure 4.5: Relative distance in lateral direction from observations and filter.

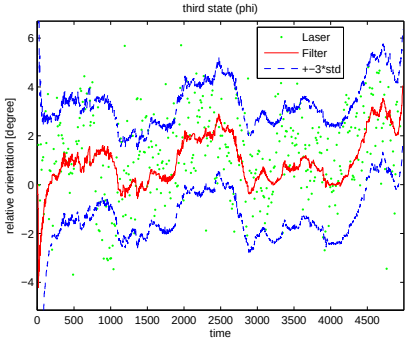


Figure 4.6: Relative orientation from observations and filter.

velocity. It can be seen more like a hint of how the filter should act. Of course any observation must not be taken too serious because it is more or less accurate but it concerns this "hybrid data" even more.

The fifth state is the yaw rate of the tracked vehicle, ω_o which has the same sign convention as y and ϕ. In figure 4.8 it is presented along with the yaw rate of the ego vehicle, ω_e. The ego vehicle is turning right in the end of the sequence causing a decrease of its yaw rate.

The final state is the width of the tracked vehicle, presented in figure 4.9. The dotted line shows the actual width of the bus which is 2.55 meters.

In figure 4.10a and 4.10b the innovation of the observations are shown for the radar and the laser respectively. Together with three times the standard deviation. The innovation covariance for the laser concerning the measured width have a different appearance compared to the others because it depends on the distance to the target. I.e. this innovation covariance grows with longer distance.

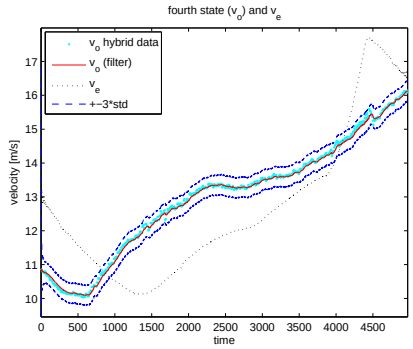


Figure 4.7: Velocity of tracked vehicle from "hybrid data" and filter.

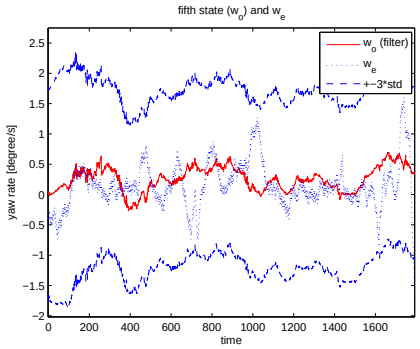


Figure 4.8: Yaw rate of tracked vehicle (filter) and ego vehicle (ego data).

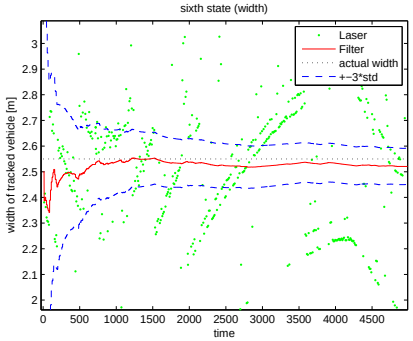
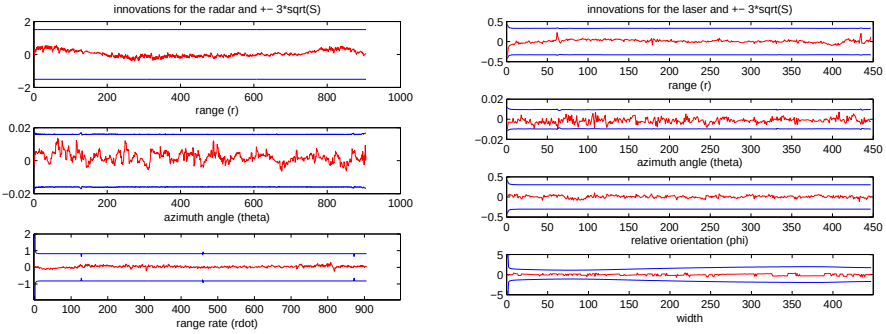


Figure 4.9: The width of the tracked vehicle from observations and filter together with the actual width.



(a) (b)

Figure 4.10: The innovations for the observations made by the radar and the laser.

4.3 Discussion

The main limitation in this work is the lack of verification data to make a qualitative analysis of the models in the filter. Some recorded data of speed and yaw rate from the tracked vehicle is needed to see if the models are good enough. First after a validation one can start to make finer adjustments.

An interesting scenario to record would be that the tracked vehicle is driving slalom in front of the ego vehicle. In this way the model would be tested for large variations in yaw rate. An option if no data from the tracked vehicle would be available could be to drive after a vehicle in a turn that is known and constant. Or if both vehicles are driving through a arbitrary turn where the measured yaw rate from the ego vehicle is very accurate and reliable. In this way a form of validation could be done if the vehicles are driving with the same speed.

The width of the tracked vehicle is estimated in a way that is not the most statistically correct approach. It is pessimistic because it assumes that one laser beam on each side of the object always misses, i.e. worst case is always assumed. A more statistically correct estimation would be to assume that one laser beam misses. I.e. that half an average distance between two laser beams misses on each side. However with the later approach the estimated width gets approximately two decimetres shorter than the actual width for the scenario used in this work. One possible reason why the pessimistic approach works better here could be that the tracked vehicle here, a scania bus has no sharp edges but rounded sides. Which will make the laser beams hitting the sides of the vehicle to reflect in other directions than back to the detector in the laser scanner unit.

The rounding of the time information to a resolution of ten milliseconds may be a little bit optimistic to give a fair view of how the filter would work in a real

time application. It is however good enough to show the qualities of the models in this scenario. Rounding to a resolution of milliseconds also leads to computational difficulties with the amount of data used in this work.

4.4 Conclusion

It is difficult to make a conclusion of the results for two reasons. The first is the lack of validation data. The second is that the tracked vehicle is driving quite straight ahead in front of the ego vehicle. Something that points to that the motion model is reasonable is that the estimated yaw rate of the tracked vehicle is around the same magnitude as the measured yaw rate of the ego vehicle. See figure 4.8. The ego vehicle yaw rate is though somewhat more noisy than would be expected from the actual movements of the truck.

However if the observations made by the sensors are reliable. As the innovations indicates in figure 4.10a and 4.10b. Then the filter does a good job to get the most possible information from the sensors. With the initialisation and the settings of the noise parameters made for this scenario, the filter relays more on the observations made by the laser when it comes to the relative position. Which is desirable since we know that the laser is able to measure the range with higher accuracy than the radar. This can be seen in figure 4.10b and 4.10a where the innovation for the range is smaller for the laser than for the radar. Also in figure 4.4b and 4.5 can one see how the filter follows the laser more than the radar. The states for the relative distance in lateral direction, y and orientation, ϕ are highly correlated since no lateral slip is allowed. A shift in relative lateral direction must be the result of a relative pose change. This is also the case for the states in the filter. Particularly the peaks in the middle and in the end sections are visible in figure 4.5 and 4.6.

An important notation is that the approach used in this work is sensitive to the amount of laser hits that is received for the tracked vehicle. At least around ten hits are needed if the pre-filter g_l is going to be able to give a reasonable observation of some quantities. Another effect is that a small change in position of some of the involved vehicles from one time instant to another, can have a great impact on which laser beams that will hit the object. Thereby some observed quantities will change dramatically. For example if one laser beam hits just at the side of a vehicle for some period and then a small move makes the beam to miss the vehicle. Then the width of the vehicle suddenly may seem to be one or two decimetres shorter. At least for a longer distance between the vehicles. The conclusion is that both the number of laser hits and also where they hit, has a great effect on the estimated states. However the relative distance in x direction is less sensitive for these effects. The greatest impact is seen for the estimation of y, ϕ and W. Especially in the observations of the width this is visible as jumps between otherwise a quite regular pattern. The two effects mentioned above are the major reasons why the observations of ϕ are noisy. This will indirect affect ω_o through the model.

Another interesting notation about this work is that it is inspired by the area of robotics rather than vehicular systems. So far the active safety systems that have been developed for automotive vehicles have been focused on the behaviour of the ego vehicle. One example is warning systems where the driver is alerted if the vehicle is crossing a road line. But in this work the focus is on what happens in front of the ego vehicle when it comes to the behaviour of other vehicles and objects. This is the main reason why the models in this thesis have more in common with models that are often used in the area of moving robots. Where the problems of relative motions involving transformations of uncertainties have been studied for a long time. With the CT model used in this work the motions of the vehicles are feed with a velocity, v and angular velocity, ω. In this way predictions can be made of what will probably happen in the next time instant with the vehicles/objects in front of the ego vehicle. Another benefit with this model compared to for example a road model is that it describes the motions of a vehicle that are independent of how the road looks like. This makes the model more general in the way that it is applicable everywhere vehicles or other objects are moving in front of the ego vehicle and therefore not dependent on external information about the road. However since the vehicles/objects live their own life independent of the road, this model is in the same time more customised.

4.5 Further Improvements/Future work

In order to know what improvements that can be done in this work. A real verification needs to be performed. But there are some possible improvements that can be done.

The ego vehicle model can be modelled with a more accurate vehicle model that include dynamics. When geometric parameters and weight etc are known a model can be built that captures the moves of the truck in a better way. Also lateral slip can be added to the motion model as shown in section 3.1.3. When lateral slip is added sideways motions are modelled as well as the longitudinal motions.

One important usage for the approach that has been tested in this work is the concept of active safety systems. Where the driver gets a warning if a vehicle is approaching the ego vehicle alarmingly fast. Or warn if an object is behaving in some other way that could be dangerous in the next time instant. This active safety system could also be used to warn if a potential crash between two vehicles in front of the ego vehicle is likely to happen. Regardless of what the primary objective would be with an active safety system there would be a need of capability to make qualitative predictions with a longer time horizon then just between the measurements. One could for example introduce acceleration of the tracked vehicles as a new state to be able to make predictions with constant acceleration as well as constant speed. As said in section 1.5 the approach here has been to test a concept independent of external information from GPS and maps, or com-

munication between vehicles. Nevertheless it can be a part of a greater system where this kind of information is available. Then problems concerning localisation and mapping could be investigated as well.

As mentioned in section 1.5 the data association problem in this work has been reduced to the ability of tracking several moving objects by the radar. Even if the approach here would be refined, still the filter would probably need the radar to initialize where to start looking in the laser data. But if the radar loses one or several targets in a couple of seconds there could be a need for other methods to associate the laser hits to the right object. One way to do this could be to differentiate the laser data and in this way find high probabilities for moving objects. However for this to work even higher demands are required on the number of laser hits of the object and that they are clean hits. Another related issue is how to scale the window in g_l to make the spatial filtering of the laser data. In this work the scaling is based on the uncertainty of the position in x and y direction and also by the estimated width of the object affecting the scaling in y direction (lateral direction). But the window could also be scaled with distance to the target.

With the setup of sensors used in this work the laser range scanner has been working with a greater geometrical working area than the radar. In order to use more of the given information from the laser in lateral direction, two radars could be used, one on each side of the front for example. There is also a difference in range capability between the sensors in favour of the laser. On the other hand fewer laser beams will hit the target at those distances (around 100 meters) which makes the observations made by the laser to be very poor. The laser scanner used in this work can be used in a mode where the angular resolution for the $\pm 10\,^\circ$ working range is increased to $0.125\,^\circ$. However the scan frequency will drop to 12.5 Hz which is a big drawback and the angular resolution for the working range between $30\,^\circ$ and $50\,^\circ$ as well as between $-60\,^\circ$ and $-30\,^\circ$ will be decreased to $0.5\,^\circ$.

The models could be extended to three dimensions to capture more dynamics (if dynamic models are used) and if the vertical working range of the laser is extended one could estimate the height of the vehicle as well as the width. The interesting with height is that it is easier to classify a vehicle if also the height can be estimated. However a camera would probably be more sufficient for an extension with an extra dimension.

Bibliography

[1] Lars A. A. Andersson. *Multi-robot Information Fusion*. PhD thesis, Linköping University, 2008. ISBN 978-91-7393-813-6.

[2] J.E. Naranjo J.G. Zato F. Aparício J.M. Armingol A. de la Escalera F. García, F. Jiménez. Analysis of lidar sensors for new adas applications. usability in moving obstacles detection, 2009.

[3] Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Prentice Hall, 3rd edition, 2009.

[4] Fredrik Gustavsson. *Statistical Sensor Fusion*. Studentlitteratur AB, 2010.

[5] Per Holmberg. Sensor fusion coordinated mobile robots. Master's thesis, Linköping University, 2003.

[6] IEEE. *Continuous Mobile Robot Localization: vision vs. laser - Robotics and Automation*, May 1999.

[7] Harold Josephs and Ronald L. Huston. *Dynamics of mechanical systems*. CRC Press, 2002.

[8] Kyeong-Hwan Leea and Reza Ehsani. Comparison of two 2d laser scanners for sensing object distances, shapes, and surface patterns. *Computers and Electronics in Agriculture*, 60:250–262, 2008.

[9] X. Rong Li and Vesselin P. Jilkov. Survey of maneuvering target tracking. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4):1333–1364, october 2003.

[10] Christian Lundquist. *Automotive Sensor Fusion for Situation Awareness*. PhD thesis, Linköping University, 2009.

[11] Christian Lundquist, Umut Orguner, and Thomas B. Schön. Tracking stationary extended objects for road mapping using radar measurements. In *Proceedings of the IEEE Intelligent Vehicle Symposium (IV)*, Xi'an, Shaanxi, China, June 2009.

[12] Christian Lundquist and Thomas B. Schön. Estimation of the free space in front of a moving vehicle. In *Proceedings of the SAE 2009 World Congress*, Detroit, MI, USA, April 2009. SAE.

[13] Randall Smith, Matthew Self, and Peter Cheesman. Estimating uncertain spatial relationships in robotics. In *Proceedings of the 2nd Annual Conference on Uncertainty in Artificial Intelligence (UAI-86)*, pages 435–461, New York, NY, 1986. Elsevier Science Publishing Company, Inc.

[14] Johan Forsberg Ulf Larsson and Åke Wernersson. Mobile robot localization: Integrating measurements from a time-of-flight laser. *IEEE Transactions on Industrial Electronics*, 43(3):422–431, june 1996.

Appendix A

Models

All models presented here are assumed to be discrete to handle the nature of the problem with discrete-time measurements and calculations. The models are also non-linear and needs to be linearised to be incorporated with the Kalman Filter. This means that the transition functions and the corresponding Jacobians after being derived, needs to be recalculated for each time step. An important notation is how the double time index should be read. $(n|m)$ means "at time n given all information up to time m". For example $\mathbf{x}(k+1|k)$ means the state evaluated at time $k+1$ given all measurements until time k.

A.1 Motion Model

Here comes a more detailed description of a general representation of a motion model. Starting with a non-linear, discrete-time state transition function in equation A.1.

$$\mathbf{x}(k) = f(\mathbf{x}(k-1), \mathbf{u}(k)) + \mathbf{w}(k) \tag{A.1}$$

Where the next state, $\mathbf{x}(k)$ is given by the previous state, $\mathbf{x}(k-1)$ the model input, $\mathbf{u}(k)$ and some additional noise, $\mathbf{w}(k)$. In order to do the linearizing of the transition function a Taylor expansion is used according to equation A.2.

$$\mathbf{x}(k) = f(\hat{\mathbf{x}}(k-1), \hat{\mathbf{u}}(k)) + \nabla f_x \tilde{\mathbf{x}}(k-1) + \nabla f_u \tilde{\mathbf{u}}(k) + O(\tilde{\mathbf{x}}^2(k-1), \tilde{\mathbf{u}}^2(k)) + \mathbf{w}(k) \tag{A.2}$$

Where the state and signal error is described by the first and second equation and the expectation mean is given by the last equation.

$$\begin{aligned} \tilde{\mathbf{x}}(k|k) &= \mathbf{x}(k) - \hat{\mathbf{x}}(k|k) \\ \tilde{\mathbf{u}}(k) &= \mathbf{u}(k) - \hat{\mathbf{u}}(k|k) \\ \hat{x} &= \mathbf{E}[x] \end{aligned} \tag{A.3}$$

The Taylor series is made about the working point defined by $\hat{\mathbf{x}}(k|k)$ and $\hat{\mathbf{u}}(k)$. So the Jacobian ∇f_x is evaluated at $\mathbf{x}(k) = \hat{\mathbf{x}}(k|k)$ and ∇f_u is evaluated at

$\mathbf{u}(k) = \hat{\mathbf{u}}(k)$. If the Taylor series is truncated after first order like in equation A.2 and used together with equation A.3 the following state prediction is given.

$$\begin{aligned}\hat{\mathbf{x}}(k|k-1) &= \mathbf{E}[\mathbf{x}(k)|\hat{\mathbf{x}}(k-1|k-1), \hat{\mathbf{u}}(k)] \\ &\approx \mathbf{E}[f(\hat{\mathbf{x}}(k-1|k-1)), \hat{\mathbf{u}}(k) + \nabla f_x(\mathbf{x}(k-1) - \hat{\mathbf{x}}(k-1|k-1)) \\ &+ \nabla f_u(\mathbf{u}(k) - \hat{\mathbf{u}}(k)) + \mathbf{w}(k)|\hat{\mathbf{x}}(k-1|k-1)] \\ &= f(\hat{\mathbf{x}}(k-1|k-1), \hat{\mathbf{u}}(k))\end{aligned} \tag{A.4}$$

Now if equation A.2 and A.4 are combined according to equation A.3 the approximated prediction error $\tilde{\mathbf{x}}(k|k-1)$ will become like in equation A.5.

$$\begin{aligned}\tilde{\mathbf{x}}(k|k-1) &= \mathbf{x}(k) - \hat{\mathbf{x}}(k|k-1) \\ &= f(\hat{\mathbf{x}}(k-1), \hat{\mathbf{u}}(k)) + \nabla f_x \tilde{\mathbf{x}}(k-1) + \nabla f_u \tilde{\mathbf{u}}(k) + O(\tilde{\mathbf{x}}^2(k-1), \tilde{\mathbf{u}}^2(k)) \\ &+ \mathbf{w}(k) - f(\hat{\mathbf{x}}(k-1), \hat{\mathbf{u}}(k)) \\ &\approx \nabla f_x \tilde{\mathbf{x}}(k-1) + \nabla f_u \tilde{\mathbf{u}}(k) + \mathbf{w}(k)\end{aligned} \tag{A.5}$$

The uncertainty of the prediction is related to the prediction error from equation A.5. If we assume that the prediction $\mathbf{P}(k|k+1)$ is equal to the conditional mean. The uncertainty of the error can be approximated as the square of the conditional mean error $\tilde{\mathbf{x}}(k+1|k)$ according to equation A.6.

$$\begin{aligned}\mathbf{P}(k|k-1) &\triangleq \mathbf{E}[\tilde{\mathbf{x}}(k|k-1)\tilde{\mathbf{x}}^T(k|k-1)|\hat{\mathbf{x}}(k-1|k-1)] \\ &\approx \mathbf{E}[(\nabla f_x(k)\tilde{\mathbf{x}}(k-1) + \nabla f_u(k)\tilde{\mathbf{u}}(k) + \mathbf{w}(k)) \\ &(\nabla f_x \tilde{\mathbf{x}}(k-1) + \nabla f_u \tilde{\mathbf{u}}(k) + \mathbf{w}(k))^T] \\ &= \nabla f_x \mathbf{E}[\tilde{\mathbf{x}}(k-1|k-1)\tilde{\mathbf{x}}^T(k-1|k-1)]\nabla^T f_x \\ &+ \nabla f_u \mathbf{E}[\tilde{\mathbf{u}}(k)\tilde{\mathbf{u}}^T(k)]\nabla^T f_u + \mathbf{E}[\mathbf{w}(k)\mathbf{w}^T(k)] \\ &= \nabla f_x \mathbf{P}(k-1|k-1)\nabla^T f_x + \nabla f_u \mathbf{Q}_u(k)\nabla^T f_u + \mathbf{Q}_w(k)\end{aligned} \tag{A.6}$$

If the errors in the driving parameters are modelled as a stochastic process with zero mean we can write the noises according to equation A.7.

$$\begin{aligned}\mathbf{Q}_u(k) &= \mathbf{E}[\tilde{\mathbf{u}}(k)\tilde{\mathbf{u}}^T(k)] \\ \mathbf{Q}_w(k) &= \mathbf{E}[\mathbf{w}(k)\mathbf{w}^T(k)] \\ \mathbf{E}[\tilde{\mathbf{u}}(k)] &= 0 \\ \mathbf{E}[\mathbf{w}(k)] &= 0\end{aligned} \tag{A.7}$$

With the substitution $\mathbf{F}_x(k) = \nabla f_x$ and $\mathbf{G}_u(k) = \nabla f_u$ the model can be written in the form presented in equation A.8.

$$\begin{aligned}\hat{\mathbf{x}}(k|k-1) &= f(\hat{\mathbf{x}}(k-1|k-1), \hat{\mathbf{u}}(k)) \\ \mathbf{P}(k|k-1) &= \mathbf{F}_x(k-1)\mathbf{P}(k-1|k-1)\mathbf{F}_x^T(k-1) \\ &+ \mathbf{G}_u(k-1)\mathbf{Q}_u(k-1)\mathbf{G}_u^T(k-1) + \mathbf{Q}_w(k-1)\end{aligned} \tag{A.8}$$

This general model fits almost all kinematic models since it can introduce noise from both external and internal driving variables. For example in the CT model used in this thesis the external noise $\mathbf{w}(k)$ in equation A.1 is replaced with a driving noise in the input signal, $\mathbf{u}$.

A.2 Sensor Model

Here a general sensor model is derived that has been used as role model for the sensor models in this thesis. The model will become like in equation A.9.

$$\mathbf{z}(k) = h(\mathbf{x}(k)) + \mathbf{v}(k) \tag{A.9}$$

Where $\mathbf{v}(k)$ is an observation error with the following characteristics:

$$\begin{aligned} \mathbf{E}[\mathbf{v}(k)] &\approx 0 \\ \mathbf{E}[\mathbf{v}(k)\mathbf{v}^T(k)] &= \mathbf{Q}_v(k) \end{aligned} \tag{A.10}$$

In the general case it is assumed that the observation is non-linear. As with the motion model, Taylor expansion is used to get an approximately linear model. Where the errors and expected mean still holds from equation A.3.

$$h(\mathbf{x}(k)) = h(\hat{\mathbf{x}}(k|k-1)) + \nabla h_x \tilde{\mathbf{x}}(k|k-1) + O(\tilde{\mathbf{x}}^2(k|k-1)) + \mathbf{v}(k) \tag{A.11}$$

Here ∇h_x is the Jacobian of $h(\mathbf{x}(k))$ evaluated at $\mathbf{x}(k) = \hat{\mathbf{x}}(k|k-1)$. The observation $\hat{\mathbf{z}}$ is predicted as:

$$\hat{\mathbf{z}}(i|j) \triangleq \mathbf{E}[\mathbf{z}(i)|\mathbf{Z}^j], \qquad \mathbf{Z}^j = \{z_1 ... z_j\} \tag{A.12}$$

Where all observations until time j are taken into account. As before the Taylor series is truncated to the first order. And all observations up to time $k-1$ are taken into account which gives an observation prediction as below.

$$\hat{\mathbf{z}}(k|k-1) \approx \mathbf{E}[h(\hat{\mathbf{x}}(k|k-1)) + \nabla h_x \tilde{\mathbf{x}}(k|k-1) + \mathbf{v}(k)|\mathbf{Z}^{k-1}] \tag{A.13}$$

$$= h(\hat{\mathbf{x}}(k|k-1)) \tag{A.14}$$

There are two types of errors in an observation. The first is the error from the sensor, $\mathbf{v}$. The other one is the difference between the real sensor and the model of the sensor, $h(\hat{\mathbf{x}}(k+1|k))$. I.e. one measurement error and one model error. Both these errors are included in ν and usually referred to as the Innovation.

$$\nu(k|k-1) = \mathbf{z}(k) - h(\hat{\mathbf{x}}(k|k-1|)) \tag{A.15}$$

If the Innovation is approximated with a Taylor expansion one gets the following expression:

$$\begin{aligned} \nu(k|k-1) &= \mathbf{z}(k) - h(\hat{\mathbf{x}}(k|k-1|)) \\ &= h(\hat{\mathbf{x}}(k|k-1|)) + \nabla h_x \tilde{\mathbf{x}}(k|k-1) + O(\tilde{\mathbf{x}}^2(k|k-1)) + \mathbf{v}(k) \\ &\quad - h(\hat{\mathbf{x}}(k|k-1|)) \\ &= \nabla h_x \tilde{\mathbf{x}}(k|k-1) + O(\tilde{\mathbf{x}}^2(k|k-1)) + \mathbf{v}(k) \end{aligned} \tag{A.16}$$

The total uncertainty of the measurements is calculated with the Innovation. In this way both errors from the measurements and errors in the prediction will be taken into account.

$$\begin{aligned}\mathbf{S}(k|k-1) &= \mathbf{E}[\nu(k|k-1)\nu^T(k|k-1)] \\ &\approx \mathbf{E}[(\nabla h_x \tilde{\mathbf{x}}(k|k-1) + \mathbf{v}(k))(\nabla h_x \tilde{\mathbf{x}}(k|k-1) + \mathbf{v}(k))^T] \\ &= \nabla h_x \mathbf{E}[\tilde{\mathbf{x}}(k|k-1)\tilde{\mathbf{x}}^T(k|k-1)]\nabla^T h_x + \mathbf{E}[\mathbf{v}(k)\mathbf{v}^T(k)] \\ &= \nabla h_x \mathbf{P}(k|k-1)\nabla^T h_x + \mathbf{Q}_v(k)\end{aligned} \tag{A.17}$$

And with the substitution $\mathbf{H}_x(k) = \nabla h_x$ the innovation covariance can be written as:

$$\mathbf{S}(k|k-1) = \mathbf{H}_x(k-1)\mathbf{P}(k|k-1)\mathbf{H}_x^T(k-1) + \mathbf{Q}_v(k-1) \tag{A.18}$$

A notification about the radar model should be stressed. The model that has been derived above is a general model for exteroceptive sensors that makes relative observations from the measurements of the surrounding environment. In the case of the radar model however the observations made of the tracked vehicles speed, v_o are dependent on the ego vehicle speed and thus the input vector, $\mathbf{u}$. The fact that h_r also depends on the input signal makes this model incomplete. A more correct innovation covariance for the radar sensor would look like the following equation:

$$\mathbf{S}(k) = \mathbf{H}_x(k)\mathbf{P}(k|k-1)\mathbf{H}_x^T(k) + \mathbf{J}_u(k)\mathbf{Q}_u(k)\mathbf{J}_u^T(k) + \mathbf{Q}_v(k) \tag{A.19}$$

Where $\mathbf{J}_u$ is the Jacobian with respect to the input signal ($\mathbf{J}_u = \nabla h_u$). Equation A.19 can be derived in analogy with the motion model in appendix A, where Jacobians with respect to both state and input vectors are used.

Appendix B

Stochastic Relationships

When it comes to relative motion between different frames as in this work. There will be a need of dealing with the spatial uncertainty from the frames involved. And to transform the uncertainty information from one frame to another. In the following sections two operations are presented that can be used to express the resultant of any sequence of relationships. These two dimensional spatial relationships are often used in robotics applications. They are described in [13] along with examples from the robotics society. Here X_{ij} means a relationship from i to j. In this work it can be a move from i to j or a measurement of a relative position j from position i, where $X_{ij} = [x_{ij}, y_{ij}, \phi_{ij}]^T$.

B.1 Compounding

Suppose that we want to add two spatial relationships X_{ij} and X_{jk} to get the resultant relationship X_{ik}. The following operation is called compounding. It is used to calculate the resultant relationship from two given relationships which are arranged head-to-tail.

$$X_{ik} \triangleq X_{ij} \oplus X_{jk} = X_{ij} + \mathbf{R}(\phi_{ij}) X_{jk} = \begin{bmatrix} x_{jk}\cos\phi_{ij} - y_{jk}\sin\phi_{ij} + x_{ij} \\ x_{jk}\sin\phi_{ij} + y_{jk}\cos\phi_{ij} + y_{ij} \\ \phi_{ij} + \phi_{jk} \end{bmatrix} \tag{B.1}$$

This operation could be used to determine the location of an object after a sequence of relative motions. And as seen this is not just vector addition but involves rotation as well. To understand how this operation affects the uncertainty the Jacobian needs to be derived. If the first-order estimate of the mean of the compounding operation is:

$$\hat{X}_{ik} \approx \hat{X}_{ij} \oplus \hat{X}_{jk} \tag{B.2}$$

Then the first-order estimate of the covariance becomes:

$$\mathbf{P}(X_{ik}) \approx J_{\oplus} \begin{bmatrix} \mathbf{P}(X_{ij}) & \mathbf{P}(X_{ij}, X_{jk}) \\ \mathbf{P}(X_{jk}, X_{ij}) & \mathbf{P}(X_{jk}) \end{bmatrix} J_{\oplus}^T \tag{B.3}$$

And the Jacobian of the compounding operation is given by:

$$J_{\oplus} \triangleq \frac{\partial(X_{ij} \oplus X_{jk})}{\partial(X_{ij}, X_{jk})} = \begin{bmatrix} 1 & 0 & -(y_{ik} - y_{ij}) & \cos\phi_{ij} & -\sin\phi_{ij} & 0 \\ 0 & 1 & x_{ik} - x_{ij} & \sin\phi_{ij} & \cos\phi_{ij} & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \tag{B.4}$$

In this work however the two relationships being compounded are independent, which makes the cross terms to be zero ($\mathbf{P}(X_{ij}, X_{jk}) = 0$). And the first-order estimate can therefore be rewritten as:

$$\mathbf{P}(X_{ij}) \approx J_{1\oplus}\mathbf{P}(X_{ij})J_{1\oplus}^T + J_{2\oplus}\mathbf{P}(X_{jk})J_{2\oplus}^T \tag{B.5}$$

where $J_{1\oplus}$ and $J_{2\oplus}$ are the left and right halves of the compounding Jacobian $J_{\oplus}$:

$$J_{\oplus} = \begin{bmatrix} J_{1\oplus} & J_{2\oplus} \end{bmatrix} \tag{B.6}$$

B.2 The Inverse Relationship

To be able to express the spatial uncertainty of one frame in another frame one more operation is needed namely the inverse relationship. Given a relationship X_{ij}, the inverse relationship X_{ji} is given by:

$$X_{ji} \triangleq \ominus X_{ij} \triangleq \begin{bmatrix} -x_{ij}\cos\phi_{ij} - y_{ij}\sin\phi_{ij} \\ x_{ij}\sin\phi_{ij} - y_{ij}\cos\phi_{ij} \\ -\phi_{ij} \end{bmatrix} \tag{B.7}$$

And with the first-order mean estimate like we did before we get:

$$\hat{X}_{ji} \approx \ominus \hat{X}_{ij} \tag{B.8}$$

Then the first-order covariance estimate becomes:

$$\mathbf{P}(X_{ji}) \approx J_{\ominus}\mathbf{P}(X_{ij})J_{\ominus}^T \tag{B.9}$$

Where the Jacobian of the reversal operation is:

$$J_{\ominus} \triangleq \frac{\partial X_{ji}}{\partial X_{ij}} = \begin{bmatrix} -\cos\phi_{ij} & -\sin\phi_{ij} & y_{ji} \\ \sin\phi_{ij} & -\cos\phi_{ij} & -x_{ji} \\ 0 & 0 & -1 \end{bmatrix} \tag{B.10}$$

With $J_{\ominus}$ the uncertainty is not inverted but expressed from the opposite point of view.