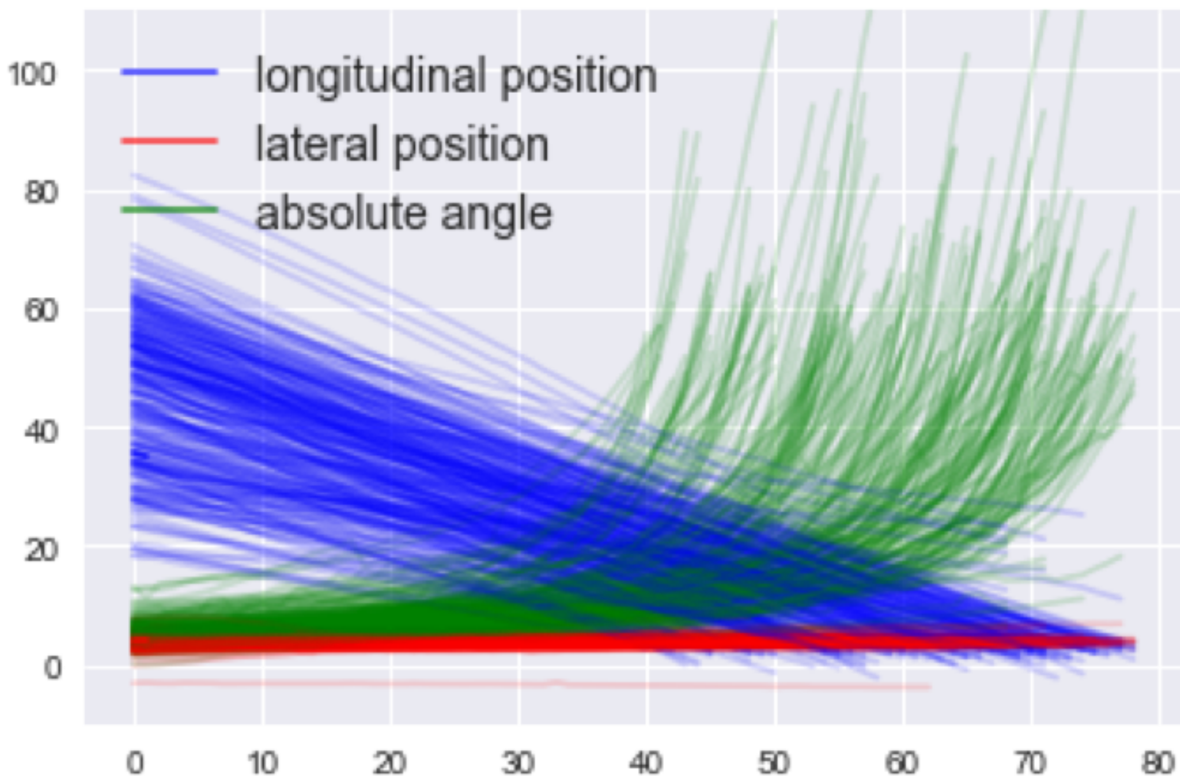
Big Automotive Data Analytics: SEnsor Modeling and Performance Analysis (BADA-SEMPA)

Public report



Project within BADA
Author Nasser Mohammadiha
Date 2019-06-19

FFI Fordonsstrategisk Forskning och Innovation

Content

FFI in short

FFI is a partnership between the Swedish government and automotive industry for joint funding of research, innovation and development concentrating on Climate & Environment and Safety. FFI has R&D activities worth approx. €100 million per year, of which about €40 is governmental funding.

Currently there are five collaboration programs: Electronics, Software and Communication, Energy and Environment, Traffic Safety and Automated Vehicles, Sustainable Production, Efficient and Connected Transport systems.

For more information: www.vinnova.se/ffi

1. Summary

The analysis of sensor data plays a crucial role to build highly automated and autonomous vehicles. Such analysis makes it possible to develop better sensor verification and accurate computer-aided engineering (CAE) simulations, and to better implement the active safety functions. During this project, we focused on four separate and yet very connected problems, namely:

1. Ground truth sensors and matching: The ground truth sensors were improved, and matching algorithms were developed to match ground truth with the production sensors.
2. Sensor modelling: Different modelling methods based on machine learning were used to model autonomous driving sensor errors. We spent quite some time on evaluating the models both functionally and in pure signal processing sense. Some of the models have been superior to others such as approaches based on hidden Markov models.
3. Set up a big data analysis framework such as Spark cluster to speed up the processing.
4. Large sensor error detection and scenario extraction: Anomaly detection methods were implemented using supervised and unsupervised algorithms to find large sensor in large amounts of data collected during expeditions. Scenario detection and clustering problem was also addressed, and an unsupervised algorithm was developed to find important driving scenarios in large amounts of data collected during expeditions.

The results of the project have been published and presented in national and international workshops and conferences.

2. Sammanfattning på svenska

Analysen av sensordata spelar en avgörande roll för att bygga högautomatiserade och autonoma fordon. En sådan analys gör det möjligt att utveckla bättre sensorverifiering och exakta simuleringar genom computer-aided engineering (CAE) och för att bättre genomföra de aktiva säkerhetsfunktionerna. Under detta projekt fokuserade vi på fyra separata och ändå väldigt knutna mål, nämligen:

1. Förbättra referenssensorerna och matcha produktion sensorer och referenssensorerna
2. Modellera olika sensor fel med maskininlärning och statistisk signalbehandling och använda och utvärdera dem i valda applikationer
3. Skapa en stor dataanalysram som Spark cluster för att påskynda behandlingen.
4. Utveckla algoritmer för att detektera stora sensor fel och anomalier, och extrahera scenarier från stora mängder av data samlats under expeditioner

Under projektet har målen varit ganska desamma med vissa ändringar som diskuteras nedan.

1. När det gäller mål 1 uppnådde vi målet och förbättrade referenssystemet till den nivå som var inom ramen för projektet.
2. När det gäller mål 2 kunde vi framgångsrikt utveckla matchnings algoritmer och modellera sensorns fel som en funktion av tillgängliga parametrar som t.ex. fordonsinformation och information om andra fordon och vägar och utvärdera dem för att förstå begränsningarna och bättre motivera fördelarna med det i att utveckla autonom körning. Effekten av vissa externa förhållanden som regn och snö kunde inte modelleras på grund av brist på rena data.
3. Ett Spark-kluster installerades på VCC som för mål 3 och konceptstudier utfördes för att förstå hur användbart det kan vara. Spark-kluster har dock inte visat sig vara mycket användbar för problemen mi detta projektet.

4. När det gäller mål 4 utvecklades flera anomalitetsdetekterings algoritmer och jämfördes. På grund av förändringar i intressen och betydelsen av scenario-detekteringsproblem för AD-verifiering, spenderade vi tid på att klustra och upptäcka scenarier i massor av data som samlats under expeditioner. En framgångsrik algoritm utvecklades i detta ämne och publicerades på en konferens.

Resultaten av projektet har publicerats och presenterats i nationella och internationella workshops och konferenser.

3. Background

Today, automated vehicles are envisioned to take over more and more driving task from the human driver with the development rapidly approaching the possibility of launching completely autonomous vehicles. As the majority of traffic related casualties and injuries can be attributed to human error, the advent of such high level of autonomy is predicted to have a paramount effect on reducing the number of casualties caused by traffic. In order to achieve this goal, Active Safety (AS) and Autonomous Driving (AD) functions have to be implemented to use the information provided by the sensors and take decisions to replace the driver decisions. As a result, the success of autonomous vehicles heavily relies on the quality of the environmental information that the sensors provide. In short, we need to replace, not only the brain of the driver but also her eyes and ears.

In order for any decision algorithm to react correctly to any situation on the road, it is essential to have reliable and accurate sensor performance in real traffic. Although sensor technology for automotive applications has undergone a huge development in recent years, it is not at the point where it can perform as good as the human sensory system. Since we cannot rely on having perfect sensors in the near future, we need to increase our knowledge about the performance of the sensors and understand how their limitations affect the decisions made by the AS and AD functions.

The increase in computational power and storage capacity allows for collecting massive amounts of information about the traffic environments during field tests. The bandwidth of each individual sensor can easily pass 10 MB/s, resulting in a huge amount of sensor data. The analysis of this sensor data is currently often done by recognizing the relatively infrequent situations where the sensor error, e.g., error in localizing the surrounding objects or missed objects, has resulted in undesired behaviour of AS or AD functions. As a result, the majority of sensor data, which carries information about the sensor performance, is not utilized.

The main goal of this project is to build big data analysis knowledge on how to automatically analyse the data that is generated by sensors used in the automotive industry. The project focuses on developing methods that are scalable for data that comes with a large volume and variety. We stress that the developed methods will utilize the collected big data to reveal structure and correlations that are very hard to discover by today's manual methods. This will lead to statistically significant descriptions of the limitations of the used sensors and safety critical sensor errors, methods for automated sensor verification and fault detection, for predicting the sensor performance in different conditions, and in the future, for better sensor fusion and AS function implementation.

Such capabilities will also allow for a development of sensor models in the sense of mathematical descriptions of the sensors, based on the observed data. Having a method with such models as output will provide the industry with a quantitative tool to simulate the sensor performance in various conditions and will result in reduced project lead times due to more accurate computer-aided engineering (CAE) environments. Moreover, sensor models can be used to identify critical conditions for which more focused test driving is needed, which can reduce the test-driving time significantly.

4. Purpose, research questions and method

The analysis of sensor data plays a crucial role to build highly automated and autonomous vehicles. Such analysis makes it possible to develop better sensor verification and accurate computer-aided engineering (CAE) simulations, and to better implement the active safety functions. The main research questions addressed by the project were related to improving ground truth, developing sensor models, and semi-, un-supervised data analysis. During the project, we have achieved the main goals and developed different analysis methods to improve the verification of AD.

The main research questions addressed by the project were:

1. How to improve the performance of the reference sensing system by developing new or improving existing signal processing and machine learning methods utilized in the reference sensors.
2. How to use the reference sensors to obtain a measure of sensor error and how to develop big data analysis methods to investigate the nature of this error, inter-dependencies of the error, and the influence of external conditions such as environmental and traffic conditions on the error.
3. How to develop unsupervised methods to analyse the big sensor data in order to get a reliable measure of sensor performance. Such methods can partially utilize the ground truth data, i.e., semi-supervised learning, but such ground truth may not be available for all the logged sensor data.

   Different machine learning methods based on regression analysis, hidden Markov models (HMM), deep neural networks (DNN), Gaussian Processes (GP), support vector machines (SVM) etc were developed in investigating the above research questions. New derivations of the algorithms were developed in order to make them suitable for large data. Methods based on hidden Markov models and deep neural networks have been the most successful methods for most problems, especially sensor modelling and scenario clustering. For details and more discussions on the developed methods, please refer to the published articles.

5. Objective

The main goal of this project has been to build a framework and knowledge on big sensor data processing that is collected from various types of sensors used in active safety and AD. Four main objectives have been within the project scope:

1. Develop and improve reference sensing systems using Velodyne LiDAR
2. Develop statistical sensor models to be able to artificially produce realistic sensor errors given a specific environment, surrounding traffic, weather, and road conditions.
3. Set up a big data analysis framework such as Spark cluster to speed up the processing.
4. Develop large sensor error detection methods without having access to ground truth data.

Additionally, the developed methods have been aimed to be used within relevant applications for ADAS and AD, such as critical sensor error analysis and CAE, and be qualitatively and quantitively evaluated.

During the project the objectives have remained quite the same with some modifications, discussed below.

1. As for Objective 1, we achieved the goal and improved the ground truth to the level that was in the scope of the project.
2. As for Objective 2, the effect of some external conditions such as rain and snow could not be modelled due to lack of clean data. Most of the data is collected in clear weather, there is data in rain and snow also for example, but it turned out to come without ground truth association so that it was not possible to apply supervised analysis on that. But we could successfully do the matching and model sensor errors as a function of available parameters such as ego vehicle information and information about other vehicles and roads and evaluate them to understand the limitations and to better justify the benefits of that in developing autonomous

driving. It is still very interesting to study the effect of external conditions such as light and rain and snow on the sensor performance and it should be possible in short future with newly collected data.

3. A small Spark cluster was set up at VCC as for Objective 3 and proof of concept studies were carried out to understand how useful that can be. A thesis work was also done to analyse the benefits of the cluster. Considering the applications and methods that have been developed, especially that the methods have been quite complex, the cluster did not turn out to be a very useful tool. The main reason for this is that most of the developed methods have had complex and time-consuming training, which happens only one time, but relatively faster inference phases. Since one need to re-write the algorithms in a Spark language, the achievable gain was not big enough to spend the requited time on this part.
4. As for Objective 4, several anomaly detection algorithms were developed and compared. Improving and integrating the algorithms into the verification chain was found to be challenging. The main reasons for this were the difficulty in defining the anomalies and not-so-good performance of the developed methods. In principle, this is a such a difficult question that it seems unfeasible to be solved with a good performance for time-being. Therefore, the objective was changed slightly, and a new question on unsupervised scenario clustering was investigated. This question was chosen due to the increasing importance of creating scenarios banks to develop and validate AD. A successful algorithm was developed on this subject and was published in a conference.

6. Results and deliverables

Improving the reference system

The most intuitive way to assess the performance of any sensor is to compare its outputs with the outputs of a reference sensor, which provides the same information but with much higher fidelity than the sensor under study. An example of such a reference system, is light detection and ranging (LIDAR) systems, which typically exhibit very small range error (in the order of 2 cm). The LIDAR systems are traditionally used for position measurements. However, the sensors used for AS and AD output more information about the surroundings, e.g., speed and acceleration of the target vehicles, road information, etc. In order to use these sensors as ground truth, various processing units such as clustering, classification, and tracking algorithms have to be developed so that all the required information can be provided by the sensor with a high fidelity. Within the project we improved the algorithms that have been developed by VCC.

Matching the reference and production sensors

With available high-fidelity reference systems, it is possible to obtain a reliable measure of sensor error by comparing the sensor output with the corresponding outputs from the reference system. With this capability, it is possible to annotate each sensor output with its corresponding error value, with the goal to assess and model the sensor performance in terms of, e.g., accuracy and missed/false detections.

In [R1] we propose an offline method to match tracked objects from two sensors in complex and real-life traffic scenarios. Detected objects in each of the sensors are described by a dynamic state vector representing their position, speed etc and the physical shape of the object. The proposed method is based on a weighted Euclidean distance between a set of features derived from these descriptors. The performance of the developed method is evaluated using a manually labeled data set, where the obtained overall true positive and negative rates are around 85% and 95%, respectively, while substantially better results were obtained considering specific traffic situations such as highways. Our experiments show that the performance of the matching method improves substantially for specific traffic scenarios such as highway driving, where around 98% true positive and negative rates are obtained.

Sensor modelling

Once we are able to match the objects from reference sensors to the production sensors, we can compute the error between the reference and production sensors. Thereby we can train a supervised ML model and then automatically generate sensor errors from that.

In [R3], we focus on sensor modelling on object level and propose sensor independent time series models based on HMM to effectively model sensor errors. We specifically consider our observations that sensor errors change slowly over time and that the errors are usually correlated with parameters describing the environment or ego-vehicle motion. To achieve a model that delivers these properties, we use an autoregressive input-output HMM (AIOHMM) to learn the sensor output time t by conditioning on an input vector and the previous output Large amounts of collected data from real traffic on country roads and highways in Europe is used and we perform simulations to show that the proposed methods can satisfactorily capture the sensor behaviors. Noticeably, we enable 30 times faster training of the original AIOHMM while still maintaining just as good. These models are sensor independent and can be used in automotive simulation environments in order to efficiently and effectively model realistic sensor outputs. The proposed data-driven models are evaluated and compared both qualitatively and quantitatively to collected data from real test drives. Our evaluations show that the AIOHMM is the best model, able to capture the temporal aspects of the sensor characteristics better than the other two models. The AIOHMM is through auto regression and the time inhomogeneous transition probabilities able to capture the temporal structures of the sequences and the dependency of the error to other parameters describing the scenario.

In [R4] and [R6], a GPS sensor model has been developed by the BADA-SEMPA team where different factors affecting GPS performance which are available in the logged data are utilized in a statistical framework to generate artificial GPS errors. For this purpose, we propose a state-based semi-stochastic model and an efficient learning algorithm, where stochastic parts are modelled using auto-regression and Gaussian mixture models. The resulting model successfully mimics the distributions of the absolute error and the first difference, for data with ideal GNSS conditions. The results indicate that the developed method is capable of producing realistic sensor errors which are of practical use in developing AD.

Large sensor error prediction

In [R5], several machine learning methods are exploited for sensor error prediction. We studied error prediction and anomaly detection methods based on Neural networks, where using the history data as well as other available information, we develop supervised regression and classification methods and evaluate their performance on real sensor data. Our experiments show that the selection of the best model depends on the number of anomalies available in the data where either regression or classification alternative might be preferred. The evaluation of the methods turns out to be application-dependent to weigh the importance between different aspects such as sensitivity and precision.

Scenario clustering

The safety of autonomous vehicles needs to be verified and validated by rigorous testing. It is expensive to test autonomous vehicles in the field, and therefore virtual testing methods are needed. Generative models of manoeuvres such as cut-ins, overtakes, and lane-keeping are needed to thoroughly test the autonomous vehicle in a virtual environment. To train such models we need ground truth manoeuvre labels and obtaining such labels can be time-consuming and costly. In [R2], we use a mixture of hidden Markov models to find clusters in manoeuvre trajectories, which can be used to speed up the labelling process. The manoeuvre trajectories are noisy, asynchronous and of uneven length, which make hidden Markov models a good fit for the data. The method is evaluated on labelled data from a test track consisting of cut-ins and overtakes with favourable results. Further, it is applied to natural data where many of the clusters found can be interpreted as driver manoeuvres under reasonable assumptions. We show that mixtures of hidden Markov models can be used to find motion patterns in driver manoeuvre data from highways and country roads.

We show that the method can be used to find typical motion patterns in a data set of multivariate motion trajectories where each motion trajectory consists of the change in relative position and velocity of the host vehicle and the tracked vehicle on the road, observed at discrete time steps from the time when it enters the host vehicle's field of view until it exits its field of view. We have shown that mixtures of hidden Markov models can be used to find typical patterns in relative motion trajectories of driver manoeuvres on highways and country roads. Our results suggest that the choice of the feature vector affects the type of clusters found. In particular, the relative velocity between the host vehicle and other tracked vehicles on the road seems like a good feature choice when looking for typical driver manoeuvres. A natural next step is to segment the manoeuvres into "partial-manoeuvres" such as left-turn, right-turn, and keeping straight, and so on using an MHMM fitted to windowed sequences of the data. The clustering can be used to segment the manoeuvres into partial-manoeuvre categories, and the learned segments can be used to train an HMM for manoeuvre generation. Other potential application areas for the method are complex driving behaviours in other kinds of traffic scenarios, for example, free, congested, and saturated traffic, and driving behaviours at signalized intersections.

7. Dissemination and publications

7.1 Dissemination

How are the project results planned to be used and disseminated?	Mark with X	Comment
Increase knowledge in the field	x	
Be passed on to other advanced technological development projects	x	
Be passed on to product development projects	x	
Introduced on the market		
Used in investigations / regulatory / licensing / political decisions		

7.2 Publications

Findings are presented and discussed in different conferences, e.g.:

[R1] J. Florbäck, L. Tornberg, N. Mohammadiha, "Offline Object Matching and Evaluation Process for Verification of Autonomous Driving", ITSC, 2016.

[R2] J. Martinsson, N. Mohammadiha, A. Schliep, "Clustering Vehicle Motion Trajectories Using Finite Mixtures of Hidden Markov Models", ITSC 2018.

[R3] E. L. Zec, N. Mohammadiha, A. Schliep, "Modelling Autonomous Driving Sensors Using Hidden Markov Models", ITSC 2018.

[R4] E. Karlsson, N. Mohammadiha, "A Statistical GPS Error Model for Autonomous Driving", IV 2018.

[R5] A. Tashvir, J. Sjöberg, N. Mohammadiha, "Sensor Error Prediction and Anomaly Detection Using Neural Networks", The first Swedish Symposium on Deep Learning (SSDL), 2017.

[R6] E. Karlsson, N. Mohammadiha, "A Data-driven Generative Model for GPS Sensors for Autonomous Driving", Software Engineering for AI in Autonomous Systems (SEFAIAS), 2018.

Furthermore, public presentations have been given and methods and results of the projects have been discussed, for example:

1. Srikar Muppirisetty, "BADA-SEMPA: big data analysis to aid verification of autonomous driving", AstaZero Researchers' Day, 2017.
2. Srikar Muppirisetty, "Big data analytics for verification of autonomous driving within BADA-SEMPA project", KTH ACCESS Data Analytics Workshop,15th May 2017.
3. Nasser Mohammadiha, "Machine Learning and Data Analysis Aiding Autonomous Driving", Swedish Workshop on Data Science, 13th December 2017.

8. Conclusions and future research

Verification of safety of the self-driving cars is a challenging step of putting these vehicles in the streets. Even with lots of collected logged data and many hours of test driving, it is still not sufficient and possible to test all the corner cases. Therefore, an efficient verification approach utilizes the benefits of virtual as well as real driving testing. To have reliable virtual testing, one needs to model all the involved components so that they behave as real components. One of the important such components is the sensor unit. In addition to that, finding the scenarios and corner cases for which AD should be tested is still a great challenge and ongoing work in academia and industry.

During the BADA-SEMPA project, we addressed several aspects of the above-mentioned problems. We developed different modelling methods based on machine learning to model autonomous driving sensor errors. We spent quite some time on evaluating the models both functionally and in pure signal processing sense. This evaluation turned out to be a great problem itself due to the stochastic nature of the underlying time series. Some of the models have been superior to others such as approaches based on hidden Markov models. Future research is needed to study the effect of external factors such as light and weather on the sensor performance.

Additionally, the ground truth sensors were improved during the project, and matching algorithms were developed to match ground truth with the production sensors. We also addressed the problem of large sensor error detection and several methods for this purpose were implemented and compared. However, the definition of anomaly, such that it can be of practical use, turned out to be hard. Moreover, none of the developed methods were able to solve this difficult problem sufficiently good. This is a question that might be addressed in the future research.

Finally, scenario detection and clustering problem was addressed within the project, and an unsupervised algorithm was developed to find driving scenarios in large amounts of data collected during expeditions. This is a very important topic in developing autonomous driving and should be investigate in follow up research projects.

9. Participating parties and contact persons

Volvo Cars: Srikar Muppirisetty (Srikar.muppirisetty@volvocars.com)

Zenuity: Nasser Mohammadiha, (Nasser.mohammadiha@zenuity.com)

Chalmers: Devdatt Dubhashi (dubhashi@chalmers.se)

ÅF: Christian U Larsson (christian.u.larsson@afconsult.com)