DRIVE SWEDEN

Improved Testing of Self-Driving Vehicles in Challenging Traffic Situations

Project leader Katarina Boustedt
ASTAZERO AB | DECEMBER 2020

With support from
VINNOVA Sweden's Innovation Agency
Swedish Energy Agency
FORMAS
Strategic innovation programmes

Table of Content

1. Summary

Today, there are cameras monitoring traffic at strategic locations in the Swedish road system. This study suggests using data from such cameras to automatically create realistic test scenarios for autonomous vehicles. In combination with driver behavior modelling and prediction, these scenarios will improve and simplify the testing of functions for autonomous vehicles. This leads to increased safety and fewer persons killed or seriously injured in traffic.

Sensors placed in the infrastructure, in this case cameras, can gather traffic information which then serves as input to generating scenarios for testing autonomous driving. In essence, real traffic situations are then influencing the assessment of future safety functions in vehicles. The outcome is more realistic testing and thus safer vehicles, both to drivers and passengers and persons outside the vehicle in question. The resulting information could also be valuable to for example authorities and cities when planning new roads and evaluating sections of infrastructure which are accident-prone and should be reconstructed. The hope is to assist the authorities with AD function safety assessment, providing relevant feedback to stakeholders to make a judgement call before releasing AD function to market.
A survey of literature and competition was made in this project. This led to a deepened interest in pursuing this topic further.

2. Swedish Summary

Idag bevakar kameror trafiken på strategiska platser i det svenska vägsystemet. Denna studie föreslår att data från sådana kameror kan användas för att automatiskt skapa realistiska testscenarier för autonoma fordon. I kombination med modellering av förarbeteende kommer dessa scenarier att förbättra och förenkla testningen av funktioner för autonoma fordon. Detta leder till ökad säkerhet och färre döda eller allvarligt skadade i trafiken.

Sensorer som placeras i infrastrukturen, i detta fall kameror, kan samla trafikinformation som sedan fungerar som indata för att generera scenarier för att testa autonom körning. I huvudsak påverkar de verkliga trafiksituationerna sedan bedömningen av framtida säkerhetsfunktioner i fordon. Resultatet är mer realistisk provning och därmed säkrare fordon, både för förare och passagerare och personer utanför fordonet i fråga. Den information som blir resultatet skulle också kunna vara värdefull för till exempel myndigheter och städer när man planerar nya vägar och utvärderar delar av infrastrukturen som är olycksbenägna och bör rekonstrueras. Förhoppningen är att hjälpa myndigheter med säkerhetsbedömning för AD-funktioner, ge relevant feedback som kan bidra till beslutsunderlag till intressenter innan AD-funktioner släpps till marknaden.
En undersökning av litteratur och konkurrenter gjordes i projektet. Detta ledde till ett fördjupat intresse för att driva ämnet vidare.

3. Background

Can output from traffic cameras be used for creating scenarios for testing of autonomous vehicles? This is the fundamental question this short study set out to explore.
It is known that merging is a very complicated situation to autonomous vehicles. Output from cameras placed in real traffic can reveal how drivers behave and this information can

be used for training vehicles. In the next step, the camera data can be implemented in simulations, from which traffic scenarios can be created. In the future, these scenarios can be used to control real or virtual objects in a test control system.

4. Project set up

4.1 Purpose

This feasibility study focuses on the cases of traffic weaving and merging at a highway on-ramp. These cases are selected because they are critical cases, with major impact on both vehicle software and sensor set, influence the sentiment of passengers and other road users, as well as having significant logistical impact on traffic flow.

4.2 Objectives

In order to develop AD and ADAS systems, industry needs a good understanding of complex traffic situations. This can be achieved by recording actual traffic flows in a real field environment.

This project has focused on the lane merging situation and the assumption that it is possible to use modern camera technology and roadside units to record actual traffic flows.

The objective was to define requirements on how an arrangement of equipment should be installed, and what data to capture in order to provide traffic flow information which can be used for repeatable test cases that meet the current and anticipated needs of vehicle manufacturers, regulators such as NCAP, and other stakeholders.

4.3 Project period

1 July 2020 – 31 December 2020

4.4 Partners

AstaZero

To serve its clients with the best possible vehicle testing experience, AstaZero wants to explore the use of real-life information from cameras for generating test scenarios, thereby simplifying testing and making it even more realistic. In the future, AstaZero's test control system may include these features, both in virtual testing and at the proving ground, and be used both in research and testing for clients.

AstaZero has served as coordinator, project manager and provider of key competence in testing, operation of test beds and test control in this project.

Viscando

Viscando is currently providing cities and road authorities with detailed insights on traffic volumes and flow, behaviors and conflicts based on measurements from its internally developed stereovision and AI-based sensor OTUS3D. Based on OTUS3D measurements, as well as the expertise in traffic data collection and analysis, Viscando has an ambitious goal to supply AV developers with naturalistic data and data-driven real-life scenarios and behavior models, thus reducing the development time and cost of AV projects. Understanding the needs and requirements for data and models, increasing the quality of collected data, and finally developing expertise in scenario extraction and behavior modelling, are paramount to reach this goal.

Viscando has provided this project with expertise in automatic systems (e.g., cameras and other systems combined by edge computing) for real-time measurement, data collection and post processing of all types of traffic: pedestrians, cyclists, cars and heavy traffic.

Zenseact, formerly Zenuity
Zenseact's purpose is to make safe and intelligent mobility real, for everyone, everywhere. Zenseact is a leading company in development of the complete software stack for ADAS and AD, from sensing to actuation where the focus is to build a single cutting-edge software platform in order to serve various levels of autonomy and offer unequalled scalability at the same time. Scenario-based verification is one of the tools that has been used in Zenseact to accelerate the development and also to prove the safe behavior of the AD software.

In this project, Zenseact has contributed with its competence in the need for field test environments characteristics and improvement potentials for verification of AD. .

5. Method and activities

This feasibility study will provide analysis and initial clarification of three functions:

- Boundaries and design of field test environment for enabling tests and analyses in relation to the operations at AstaZero
- What skills and levels should the field test environment develop, carry out research on and make available?
- What is the certification process to qualify for use of the environment?
  It is likely that gradual development of ability and offerings will take place, where weaving and merging are the first use cases. The core business is to provide enhanced intelligence to test performers (e.g., position and vehicle movements).

Topics to investigate will include

- Sensor environment design and location with focus on safety
- Test control
- Data collection
- Proof of Concept (weaving and merging)

6. Results and Deliverables

Unsupervised autonomous driving (AD) functions need to be evaluated and proven to handle all relevant traffic scenarios in their defined operational design domain (ODD[1]) before deployment. Naturalistic Field Operational Test (NFOT) is a widely used evaluation approach in which prototype cars are directly tested on public roads. However, due to the low exposure rate of safety critical scenarios, several billion kilometers of supervised road

[1] Operational Design Domain (ODD) (SAE J3016_201806): Operating conditions under which a given driving automation system or feature thereof is specifically designed to function, including, but not limited to, environmental, geographical, and time-of-day restrictions, and/or the requisite presence or absence of certain traffic or roadway characteristics.

testing is needed to confidently validate a safe behavior using this approach [1]. Therefore, road testing requires a large fleet of vehicles, long test duration and a large budget. In addition, the lack of repeatability for safety critical scenarios can be crippling during development.

Driving several billion kilometers is considered unfeasible by most leading companies. Thus, there is an urgent need for more efficient ways of evaluating self-driving features in order for them to reach the market. Virtual testing in simulation environments is one way of verifying the AD vehicles. The newly developed or updated AD software is exposed to a feasible variation of scenarios which may occur in the ODD before its integration into production vehicles. This approach, widely referred to as scenario-based verification, heavily relies on data from naturalistic driving to derive a complete catalogue of scenarios, with statistically feasible variations, necessary to reliably characterize the ODD.

Common practice is to collect the naturalistic data by means of the onboard sensing capabilities in the customer vehicles or reference sensor systems installed on test vehicles. However, the infrequent exposure to critical scenarios and road stretches with complex traffic scenarios, such as highway on/off ramps, reduces the efficiency of the data collection and hinders a deep understanding of the dynamics of such challenging scenarios.

To circumvent this limitation, several complementary measurement technologies have recently been proposed to enrich the naturalistic data needed for safety validation of automated vehicles (AV), for example drone monitoring [2]. Another approach uses stationary AI-powered traffic sensors which collect traffic data in selected locations during extended periods of time [2, 3, 4]. This results in lower cost and higher rate of relevant traffic interactions recorded compared to road vehicles. However, this technology has not been used for the purpose of full-scale data collection for AV testing before. Hence, a detailed understanding of the accuracy and availability of collected data is lacking.

Moreover, virtual testing of AV heavily relies on the degree to which the simulation environment resembles the real-world scenarios. A higher fidelity simulation environment facilitates the test process and increases the chance to identify driving situations that the AD software under test fails to safely handle. With well-defined and quantitative models that accurately describe the behavior of human-driven vehicles, it is possible to evaluate the AD vehicle in mixed-traffic environments where its performance depends crucially on the interplay with human-driven vehicles. This is especially important at highway on/off-ramps, where vehicles negotiate their ways onto the highway, to preserve traffic flow and avoid safety-critical conflicts.

A supplementary approach to safety verification of AD software is directed physical testing over a catalogue of challenging real-life scenarios in the test track. Organizations like UNECE and Euro NCAP try to enforce safety in vehicle functions by accrediting independent proving grounds to evaluate vehicles in realistic and controlled environments. It is foreseen that scenario-based evaluation of functions will become a more integral part of this process, where a subset of all simulated scenarios is run on a test track. Comparable results between simulation and test track are therefore paramount to assess the results of testing. The impact of having a toolchain which has its origin in real traffic environment could lead to more accurate testing of autonomous driving functions, which in turn could accelerate the introduction of AD vehicles to market and at the same time reduce the fatalities and seriously injured in traffic.

State-of-the-art

Infrastructure-based traffic sensors

Using Infrastructure-based traffic sensors to collect data has been done before, such as in the public dataset collected by US Department of Transportation, NGSIM [5] which is used in traffic behavior research [4]. However, this dataset is just 90 minutes in total, too short for any statistically certain conclusions on scenario parameter occurrences. Furthermore, it does not offer possibility to acquire up-to-date traffic data from new locations, a requirement for AV projects.

Traffic surveillance cameras have been used to extract scenarios and behaviors [3]. This gives large amount of data from multiple cameras but has serious drawbacks: low quality of video and limited number of camera locations, and the longitudinal position and speed accuracies of detected objects using single cameras are low.

This inaccuracy has been mitigated in a recent study [6] where an experimental setup based on data fusion between camera and traffic surveillance radar was developed, achieving improved accuracies of object tracks in a highway merge scenario, at prototype level.

Several companies offer off-the-shelf systems for automated traffic data collection and analysis, for example Bosch [7]. However, these systems are primarily dimensioned for traffic analysis rather than recording single trajectories, which limits their potential usage for the AV data collection.

A suitable system for advanced data collection is Viscando's stereovision and AI based sensor OTUS3D for 3D measurement of traffic and extraction of accurate trajectories of different object types, such as light and heavy vehicles, bikes and pedestrians. The design of this system makes it hard to locate by the road users, eliminating any influence on their behavior, which resolves a reported drawback of infrastructure-based sensors [2].

Driver behavior models

The use of driver models in vehicular system testing was revolutionized in the 2010s and has since then been a very important tool for quantitative simulation of driver–vehicle–environment (DVE) scenarios. Chalmers, together with Volvo Cars and other actors, has since then participated actively in the research, especially through successful FFI projects such as QUADRA, dnr 2009-02766, and QUADRAE, dnr 2015-04863, for the purpose of increasing the understanding of human behavior related to active safety and complex driving scenarios.

Important contributions from Chalmers have been both to map the state-of-the-art [8] specifically related to critical near-crash scenarios, and to find fundamental properties of driver behavior in novel basic research [9, 10]. Important findings were both related to the actual control of the vehicles, solving the 70-year mystery related to the *remnant* in human control [11], as shown to originate from intermittent *reaching* patterns when controlling the vehicle, and the importance of visual cues related to retinal flow and its special case for longitudinal control referred to as looming.

These findings are still considered state-of-the-art, and research is currently active to further explore the biological roots of driver behavior, for the purpose of building quantitative driver models. One such active track of research is the influence of active driver gazing [12], and how that relates to retinal flow and driver behavior. Another example of a similar research direction is how to better understand reaction times [13], where they would no longer be an effect of a symbolic and atomic event in the environment, but rather an effect of an accumulation of deviations from the expected. From these recent findings, it is clear that low-level visual ques are crucial for the understanding of human behavior, rather than the traditional high-level object and trajectory-based approach. In the traditional approach, the behavior of a single actor was very much seen from the perspective of a full cognitive awareness of the traffic environment and its future developments, where each other actor was clearly represented in the cognitive perceptual domain. However, by using the recent findings from retinal flow and active gaze modelling it can clearly be seen that driver action do exist without object-based perception, and that the combination of the retinal-flow field may influence driver action in ways that would not be predictable from a traditional objectification of the driving scene.

The modern approach of seeing human behavior as an effect of low-level sensation, without mandatory conceptualization of the traffic environment, is promising for (1) generating state-of-the-art realism in the model behavior, and (2) to further bring the research and state-of-the-art forward.

Closed-loop simulations for development and verification of AD functions

Computer simulations play a crucial role in the development and verification of AD algorithms as it is impractical or even impossible to prove their safety by real driving [1]. Closed-loop simulations allow for virtually driving a great number of traffic scenarios by running in parallel on computer clusters, enabling efficient testing of AD functions. The use of such simulation environments is now common practice among automotive OEMs and start-ups aiming at the commercialization of AD products [14].

Because of the many different players and methodologies involved in AD testing, the adoption of open standards is of key importance as it enables a shared understanding of the safety argumentation among industries and regulatory bodies. OpenDRIVE [15] and OpenSCENARIO [16] standards are among the most established standards to represent the static and dynamic content of simulation scenarios. The suitability of these two standards for industrial and research use cases was investigated by the FFI project Simulation Scenarios [17]. One of the deliverables of the project was the scenario engine esmini [18], an open-source, cross-platform application to run OpenSCENARIO and OpenDRIVE files. The application is currently being maintained and improved in functionality by the open-source community, with major contributions from Swedish partners.

The traffic scenarios investigated by the Simulation Scenarios project include interaction between the host vehicle and other traffic actors in terms of scripted, pre-determined interactions (e.g., triggered maneuvers, platooning), but do not investigate how to include accurate driver models in the OpenSCENARIO formulation. Future work should address this gap by investigating how to include accurate driver models into the OpenSCENARIO formulation and how to execute AD-relevant scenarios in the esmini environment.

Scenarios and standards

Scenario based verification and validation of autonomous driving (AD) functions is a popular area undergoing a lot of change. Large efforts are made towards harmonization of test and evaluation processes for AD function validation. An example of this is the Pegasus project [19] which was completed in the autumn of 2019. A result of the project was a process for function verification called “The Pegasus Method” [20]. Based on data collected from real world test driving and accident databases, scenarios can be reconstructed and represented in a common format. The scenarios can then be used as part of validation test schemes aimed at safety assurance argumentation, where the AD functions are evaluated on a set of parameterized scenarios. The project aligned with other scenario format initiatives, such as the aforementioned OpenDRIVE [15] and OpenSCENARIO [16] to make use of already existing open-source formats. Today, both these formats are developed by ASAM [21] where the first official version has been released and ongoing work for the second version is on its way.

In 2019, a new ISO standard working group was formed in order to streamline the work towards safety assurance argumentation. The working group ISO/TC 22/SC33/WG09 "Test scenarios of automated driving systems" [22] has since then continued to work towards a scenario-based approach of verifying and validating AD functions, using the results from the Pegasus project. The group is looking to release four standards by 2023.

Work done in the Swedish national FFI research project ESPLANADE [23] has shown that certification of AD functions through already existing standards, such as ISO 26262 [24], is not feasible. The project, which delivered its final report in 2019, argued that a plausible way forward in evaluating safety is to apply a method called quantitative risk norm (QRN). QRN defines an acceptable frequency of incidents with different levels of severity and a mapping of incidents to different classes of definitions. This type of argumentation could lead to faster exposure of AVs on national roads, where an acceptable number of incidents are allowed. A scenario-based approach could potentially support this method by examining the incidents occurring in a defined set of testable scenarios in an ODD, evaluated in simulation and test track.

Tools for testing autonomous functions have been prototyped in the FFI project Chronos part 2, Dnr 2017-05501 [25], where the unpredictability of AVs must not make testing unfeasible. Based in scenarios, the tools developed showed capability of synchronously adapting actors to the action of the VUT on the test track, thus enabling the running of complex scenarios that could only be set up in simulations before, in a safe and repeatable way. Work developing the tools will be continued in the successor FFI application "Phanes", coordinated by AstaZero, should it be approved. Using these tools for test track execution will be an important part of the scenario format and driver behavior model evaluation.

7. Conclusions, Lessons Learnt and Next Steps

In conclusion, this study has shown that there is a great potential in automating scenario generation using camera output, both for traffic planning and simulation and for vehicle testing.

The project partners soon realized that driver behavior modeling would be greatly beneficial and hence contact was made with Chalmers. Furthermore, it would be beneficial to include the end-user perspective in future work, and therefore Volvo Cars was approached and has shown great interest in furthering this research area.

This extended group has created a project proposal submitted to Vinnova FFI EMK for the December 2020 call. The aim of the proposed project is to explore using output from traffic cameras as input to scenario generation for virtual and real-life testing of autonomous vehicle systems.

8. Dissemination and Publications

As per the purpose of this short study, no publications as such have been generated. However, the project has made initial contacts with Trafikverket and will submit a short text towards a future project application on the strategic aspects of implementing cameras for traffic analysis as described in this report.

References

1. N. Kalra, S. M. Paddock. "Driving to safety: How many miles of driving would it take to demonstrate autonomous vehicle reliability?", Transportation Research Part A: Policy and Practice 94 (2016): 182 - 193.
2. R. Krajewski, J. Bock, L. Kloeker, L. Eckstein, "The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems," 2018 21st International Conference on Intelligent Transportation Systems (ITSC), Maui, HI, 2018, pp. 2118 - 2125
3. X. Ren, D. Wang, M. Laskey, K. Goldberg, "Learning Traffic Behaviors by Extracting Vehicle Trajectories from Online Video Streams", *14th IEEE Int. Conf. on Automation Science and Engineering (CASE)*, Munich, 2018, pp. 1276-1283
4. B. Coifman and L. Li, "A critical evaluation of the Next Generation Simulation (NGSIM) vehicle trajectory dataset," Transportation Research Part B: Methodological, vol. 105, pp. 362 – 377, 2017
5. J. Colyar and J. Halkias, NGSIM - US Highway 101 Dataset, https://www.fhwa.dot.gov/publications/research/operations/07030/07030.pdf
6. D. Notz, F. Becker, T. Kühbeck, D. Watzenig, (2020). Extraction and Assessment of Naturalistic Human Driving Trajectories from Infrastructure Camera and Radar Sensors. arXiv preprint arXiv:2004.01288.
7. Bosch Smart City, https://cds.bosch.us/smart-city
8. G. Markkula, O. Benderius, K. Wolff, M. Wahde, A Review of Near-Collision Driver Behavior Models, Human Factors: The Journal of the Human Factors and Ergonomics Society, Vol. 54 iss. 6, 1117-1143, 2012
9. O. Benderius, G. Markkula, Evidence for a fundamental property of steering, Human Factors: The Journal of the Human Factors and Ergonomics Society, Volume: 58 issue: 1, 884 - 888, 2014
10. G. Markkula, O. Benderius, M. Wahde, Comparing and validating models of driver steering behaviour in collision avoidance and vehicle stabilisation, Int. J. of Vehicle Mechanics and Mobility, 52:12, 2014
11. O. Benderius, Modelling driver steering and neuromuscular behaviour, Ph.D. Thesis, Chalmers U. Tech., 2014
12. S. Tuhkanen, J. Pekkanen, P. Rinkkala, C. Mole, R. M. Wilkie, O. Lappi, Humans Use Predictive Gaze Strategies to Target Waypoints for Steering, Scientific Reports volume 9, Article number: 8344 (2019
13. G. Markkula, J. Engström, J. Lodina, J. Bärgman, T. Victor, A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies, Accident Analysis & Prevention, Volume 95, Part A, October 2016, pp. 209 - 226
14. Waymo blog post, https://blog.waymo.com/2020/10/revealing-our-approach-to-safety.html
15. OpenDRIVE, https://www.asam.net/standards/detail/opendrive/
16. OpenSCENARIO, https://www.asam.net/standards/detail/openscenario/
17. FFI project Simulation Scenarios, ref. 2016-05495, https://www.vinnova.se/en/p/simulation-scenarios/
18. Esmini, https://github.com/esmini/esmini
19. The Pegasus project, https://www.pegasusprojekt.de/en/about-PEGASUS
20. The Pegasus method, https://www.pegasusprojekt.de/en/pegasus-method
21. ASAM, https://www.asam.net
22. ISO/TC 22/SC33/WG09 "Test scenarios of automated driving systems", https://www.iso.org/committee/5383785.html
23. FFI Project ESPLANADE, https://esplanade-project.se
24. ISO 26262, https://www.iso.org/standard/43464.html
25. Chronos part 2, https://www.astazero.com/chronos-part-2/