



VDI-Tagung

Fahrerassistenzsysteme und automatisiertes Fahren

Aachen, 17. und 18. Mai 2022

Bildquelle: © RioPatuca Images – fotolia.com

VDI-BERICHTE
Herausgeber:
VDI Wissensforum GmbH

Bibliographische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliographie; detaillierte bibliographische Daten sind im Internet unter www.dnb.de abrufbar.

Bibliographic information published by the Deutsche Nationalbibliothek (German National Library)

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliographie (German National Bibliography); detailed bibliographic data is available via Internet at www.dnb.de.

© VDI Verlag GmbH · Düsseldorf 2022

Alle Rechte vorbehalten, auch das des Nachdruckes, der Wiedergabe (Photokopie, Mikrokopie), der Speicherung in Datenverarbeitungsanlagen und der Übersetzung, auszugsweise oder vollständig.

Der VDI-Bericht, der die Vorträge der Tagung enthält, erscheint als nichtredigierter Manuscriptdruck.

Die einzelnen Beiträge geben die auf persönlichen Erkenntnissen beruhenden Ansichten und Erfahrungen der jeweiligen Vortragenden bzw. Autoren wieder. Printed in Germany.

ISSN 0083-5560

E-ISBN 978-3-18-102394-7

ISBN 978-3-18-092394-9

Inhalt

► Pitch der Innovationen – Impulsvorträge im Plenum

Manipulation von Sensordaten aus Testfahrten zur Analyse und Bewertung implementierter Rückfalllösungen	1
D. Michalik, P. Kohl, A. Kummert, Bergische Universität Wuppertal	

► Sensortechnologien und Perzeption

Radar Target Simulator – Key Technology for AV Development	13
P. Kannan, C. Miquet, IPG Automotive GmbH, Karlsruhe	

► Künstliche Intelligenz (KI), Verhaltensplanung und Kooperation

Realisierung einer querführenden Fahrerassistenzfunktion mithilfe von adaptiver Regelung und neuronalen Netzen	27
L. Witt, D. Münning, H. Oschlies, Volkswagen AG, Wolfsburg;	
S. Schmidt, Otto-von-Guericke-Universität, Magdeburg	

Augmentation von Kameradaten mit Generative Adversarial Networks (GANs) zur Absicherung automatisierter Fahrfunktionen	41
P. Rigoll, P. Petersen, L. Ries, J. Langner, E. Sax, FZI Forschungszentrum Informatik, Karlsruhe	

Kalibrierung von Neuronalen Netzen für Detektionsmodelle	49
F. Küppers, J. Kronenberger, A. Haselhoff, Hochschule Ruhr West, Bottrop;	
J. Schneider, Elektronische Fahrwerksysteme (EFS), Gaimersheim	

Projekt COPE – Collective Perception zur Vermeidung von Kollisionen und gefährlichen Situationen mittels V2X	63
A. Kuhn, M. Lauterbacher, E. Thonhofer, K. Glader, ANDATA, Hallein, Österreich;	
P. Luley, JOANNEUM RESEARCH, Graz, Österreich;	
T. Novak, SWARCO, Perchtoldsdorf, Österreich;	
W. Aigner, HiTec, Wien, Österreich;	
W. Schildorfer, Logistikum FH Steyr, Steyr, Österreich	

► Architekturen für voll- und teilautomatisiertes Fahren

UNICARagil – Disruptive Modular Architectures for Agile Automated Vehicle Concepts 75
T. Woopen, R. van Kempen, B. Lampe, L. Eckstein, Institute for Automotive Engineering (ika) –
RWTH Aachen University

Ein industrieübergreifender Überblick von fehlertoleranten Ansätzen in autonomen Systemen. 89

T. F. Horeis, F. Plinke, Institut für Qualitäts- und Zuverlässigkeitmanagement GmbH, Hamburg;
T. Kain, H. Tompits, Technische Universität Wien, Österreich;
R. C. Rinaldo, DFKI, Bremen;
J. Heinrich, Institut für Qualitäts- und Zuverlässigkeitmanagement GmbH, Wuppertal

Abbildung rechtlicher Anforderungen auf die Architektur komplexer Fahrerassistenzsysteme und autonomer Fahrzeuge. 109

G. Becker, A. Camarinopoulos, A. Papasileka, E. Karamanolis, RISA Sicherheitsanalysen GmbH,
Berlin;
M. Rill, D. Vonderau, B. Liu, V. Pazmino Betancourt, J. Becker,
FZI Forschungszentrum Informatik, Karlsruhe

► Absicherung automatisierter Fahrfunktionen

Real World Scenarios for the Safety Validation and Development of Highly Automated Driving Functions 121

T. Moers, C. Klas, A. Zlocki, fka GmbH, Aachen;
L. Vater, Institute for Automotive Engineering, RWTH Aachen University

Entwicklung, Erprobung und Regelbetrieb automatisierter Fahrzeuge aus datenschutzrechtlicher Perspektive. 131

H. Steege, CARIAD SE, Hannover

► Interaktion mit automatisierten Fahrzeugen und Nutzerakzeptanz

Do you want me to shift lanes? – Investigating automated driving strategies in urban areas 143

P. Roßner, K. Felbel, A. Dettmann, A. C. Bullinger, Chemnitz University of Technology,
Chemnitz

Nachweis der Betriebsbewährung automatisierter und autonomer Fahrzeuge. 159

U. Steininger, TÜV SÜD, München;

J. Mazzega, S. Witkowski, T. Form, Volkswagen, Wolfsburg;

K. Lemmer, Deutsches Zentrum für Luft- und Raumfahrt, Braunschweig



OHNE PROTOTYP GEHT NICHTS IN SERIE.

Unser Podcast ist das Werkzeug, mit dem Sie Ihre Karriere in allen Phasen entwickeln – vom Studium bis zum Chefsessel. Egal, ob Sie Ingenieur*in, Mechatroniker*in oder Wissenschaftler*in sind: Prototyp begleitet Sie. Alle 14 Tage hören Sie die Redaktion von INGENIEUR.de und VDI nachrichten im Gespräch mit prominenten Gästen.

INGENIEUR.de
TECHNIK - KARRIERE - NEWS



PROTOTYP

Karriere-Podcast

JETZT REINHÖREN UND KOSTENFREI ABBESTELLEN:
[WWW.INGENIEUR.DE/PODCAST](http://www.ingenieur.de/podcast)

IN KOOPERATION MIT VDI NACHRICHTEN

Manipulation von Sensordaten aus Testfahrten zur Analyse und Bewertung implementierter Rückfalllösungen

**David Michalik, M.Sc., Per Kohl, M.Sc.,
Prof. Dr.-Ing. Anton Kummert, Bergische Universität Wuppertal**

Zusammenfassung

In diesem Beitrag wird ein Ansatz zur Manipulation von Sensordaten aus realen Testfahrten vorgestellt, mit dem sich potentielle Fehlereinflüsse auf Seiten der Sensorik nachträglich durch entsprechende Algorithmen nachbilden lassen. Die entwickelte Software ist Teil einer flexiblen Testumgebung, welche im Rahmen des Förderprojektes „AFFiAncE“¹ zur Erprobung neuartiger Fahrzeugarchitekturen realisiert wird.

Die erzeugten Fehlerfälle ermöglichen eine schnelle und reproduzierbare Bewertung der implementierten Rückfalllösungen im Falle eines Fehlers auf Seiten der Sensorik und dienen zur Analyse der Zuverlässigkeit und Sicherheit neuartiger Systemarchitekturen für hoch- und vollautomatisierte Fahrzeuge.

Abstract

This paper presents an approach for an algorithm-based manipulation of real-world sensor data from test drives, which allows a subsequent recreation of potential influences due to errors in the sensor system.

The developed software is part of a flexible test environment, that is being designed within the funded project "AFFiAncE" for testing new types of vehicle architectures.

The generated fault cases enable a fast and reproducible evaluation of the implemented fallback solutions in the event of a fault in the sensor system. This enables the analysis of the reliability and safety of novel system architectures for highly and fully automated vehicles.

¹ Dieses Vorhaben wird aus Mitteln des Europäischen Fonds für regionale Entwicklung (EFRE) gefördert.

1. Einleitung

Die Zukunft der Mobilität wird derzeit von vier verschiedenen Entwicklungen beherrscht, die sich unter dem Akronym „CASE“ zusammenfassen lassen. Das Akronym setzt sich dabei aus den Anfangsbuchstaben der Begriffe Connectivity, Autonomous, Shared und Electric zusammen [1]. Es wird erwartet, dass sich vernetzte und automatisierte Verkehrssysteme in den nächsten Jahrzehnten immer stärker durchsetzen und dabei erhebliche Auswirkungen auf die Mobilität, Sicherheit, Umwelt und Gesellschaft insgesamt haben.

Für eine erfolgreiche Einführung spielt das Vertrauen der Gesellschaft in die Sicherheit und Zuverlässigkeit einer Technologie eine zentrale Rolle. Bei der Entwicklung von automatisierten Fahrfunktionen für höhere Automatisierungsgrade stellt sich daher die Frage, wie die Anforderungen an ihre Sicherheit erreicht und in geeigneter Form nachgewiesen werden können. Durch den vollständigen Wegfall des Fahrers als Rückfallebene (ab SAE Level 4) dürfen automatisierte Systeme im Fehlerfall nicht, wie bislang, in einen passiven sicheren Zustand übergehen (Fail-Safe-Prinzip). Ein solches System muss jederzeit in der Lage sein, seine Funktionsfähigkeit zu gewährleisten und das Fahrzeug im Falle eines Fehlers oder einer Störung in einen sicheren, risikominimalen Zustand zu versetzen, z.B. durch den Halt auf dem Seitenstreifen (Fail-Operational-Prinzip).

Hierfür werden neue Fahrzeugarchitekturen benötigt, welche aus technischer Sicht die hohen Anforderungen hinsichtlich ihrer Robustheit und Zuverlässigkeit erfüllen und sich aus ökonomischer Sicht rentabel und finanziert in Serie realisieren lassen.

Da eine Vielzahl an unterschiedlichen Topologien denkbar ist, bedarf es neuer Ansätze zur Analyse, Bewertung sowie dem Vergleich neuartiger Architekturen und Topologien. Im Rahmen des Förderprojektes „AFFiAncE“ entsteht hierfür eine flexible Testumgebung, welche eine Erprobung neuartiger Fahrzeugarchitekturen mit seriennaher Hardware unter realistischen Bedingungen im Labor ermöglicht.

In diesem Beitrag wird ein darin entwickelter Ansatz zur Manipulation von Sensordaten vorgestellt, mit welchem sich potentielle Einflüsse aus Störungen und Fehlfunktionen nachträglich auf bereits vorhandene Sensordaten aus realen Testfahrten einstreuen lassen. Dies ermöglicht die Untersuchung der Relevanz und Auswirkungen möglicher Fehler und Ausfälle seitens der Sensorik. Die daraus gewonnenen Erkenntnisse lassen sich für die Entwicklung neuer Architekturentwicklungen und Fehlerbehandlungsstrategien heranziehen, um sichere, zuverlässige und robuste Fahrzeugarchitekturen zu entwickeln und damit schlussendlich das Vertrauen in das automatisierte Fahren zu steigern.

2. Neue Anforderungen an Fahrzeugarchitekturen

Experten, insbesondere aus der deutschen Automobilindustrie, erwarten, dass die Entwicklung des automatisierten Fahrens einen evolutionären Prozess in Form einer stufenweisen Weiterentwicklung der automatisierten Fahrfunktionen beschreiten wird [2]. Die ständige Fortentwicklung in den letzten Jahren hat dazu geführt, dass mit jeder neuen Assistenzfunktion zusätzliche Steuergeräte, Sensoren und Kabelbäume in die Fahrzeugarchitektur integriert wurden. Schätzungen zufolge befinden sich in modernen Fahrzeugen mittlerweile bis zu 120 dezentrale Steuergeräte, die zusammen genommen einen Umfang von mehr als 100 Millionen Codezeilen aufweisen [3]. Redundante Codeteile können bei dieser Herangehensweise nicht ausgeschlossen werden und erhöhen unnötig die Komplexität des Systems, ohne die Möglichkeiten der zahlreich verbauten Informationsquellen und Sensoren, z.B. für die Umweltwahrnehmung, vollständig auszuschöpfen. Bisherige Architekturen stoßen zunehmend an die Grenzen der Skalierbarkeit und führen zu überhöhten Kosten, Platzbedarf und Gewicht.

Für die weitere Entwicklung des automatisierten Fahrens ist diese Vorgehensweise sowohl aus Hardware- als auch aus Softwaresicht ungeeignet. Aktuelle Entwicklungen im Bereich automatisierter Fahrfunktionen zeigen, dass eine steigende Anzahl an internen Daten aus Sensoren sowie externen Daten aus dem Informationsaustausch mit anderen Verkehrsteilnehmenden und der Infrastruktur (V2X) benötigt werden. Die Verarbeitung erfordert höhere Rechenleistungen und eine schnellere Kommunikationsinfrastruktur in den Fahrzeugen. Die Auslegung nach dem Fail-Operational Prinzip erfordert zudem die redundante Auslegung sicherheitsrelevanter Funktionen zur Absicherung des Gesamtsystems. Darüber hinaus müssen die Systemarchitekturen für das automatisierte Fahren skalierbar sein und flexibel für unterschiedliche Fahrzeugkonzepte (PKW, People Mover, etc.) eingesetzt werden können, um Synergieeffekte zu schaffen. Over-the-Air (OTA) Updates sind dabei essentiell, um die Kompatibilität, Sicherheit und Erweiterbarkeit über die gesamten Fahrzeugelebenszyklus gewährleisten zu können.

Um diesen sowie zukünftigen Anforderungen gerecht zu werden, sind neue Konzepte und Ideen für die Fahrzeugarchitektur erforderlich. Jüngste Entwicklungen in der Automobilindustrie konzentrieren sich auf Domänen-, Zonen- und serverbasierte Architekturen, die die Kommunikation, Datenanalyse und Rechenleistung zentralisieren. Dies trägt zur Optimierung und Vereinfachung der Architektur in Bezug auf Funktionalität, Kommunikation und Stromversorgungsstruktur bei und führt letztlich zu Gewichts- und Kosteneinsparungen (vgl. [4], [5], [6]).

Aufgrund der Neuartigkeit dieser Konzepte fehlt es bislang an Erfahrungswerten und empirischen Daten, die einen Vergleich und eine Verifikation unterschiedlicher Topologien ermöglichen. Daher bedarf es neuer Lösungsansätze, um zunächst unterschiedliche Topologien zentralisierter Fahrzeugarchitekturen zu entwickeln, testen und analysieren und schlussendlich durch den Vergleich eine möglichst kostenoptimierte Architektur mit hoher Robustheit bestimmen zu können.

3. Lösungsansatz mittels konfigurierbarem Prüfstand

Um den Aufwand bei der Entwicklung künftiger Fahrzeugarchitekturen zu minimieren sowie standarisierte und reproduzierbare Testprozeduren zu ermöglichen, wurde im Forschungsprojekt „AFFiAncE“ ein Lösungsansatz in Form einer flexiblen, erweiterbaren Testumgebung entwickelt, der die Erprobung neuartiger Architekturen und Topologien mit seriennaher Hardware unter realistischen Bedingungen im Labor ermöglicht (siehe Bild 1) [7].

Das Herzstück der Testumgebung bildet dabei ein Prüfstand, der eine Möglichkeit zur flexiblen physikalischen Konfiguration der Kommunikationsstruktur zwischen einem Sensordaten-Einspeisegerät und drei Rechenplattformen schafft, wodurch unterschiedliche Topologien simuliert werden können. Dieses Open-Loop Testverfahren ermöglicht eine Analyse der Auswirkung möglicher Ausfälle, Degradierungen und Fehler in den unterschiedlichen Bestandteilen der vorliegenden Systemarchitektur, nämlich den Rechensystemen, der Sensorik, der Kommunikations- und Energieversorgungsstruktur sowie der Umgebung des Fahrzeugs, um diese hinsichtlich ihrer Robustheit und Zuverlässigkeit zu bewerten.

Für die Untersuchung der Auswirkungen möglicher Fehler und Ausfälle seitens der Sensorik werden zur Stimulation Sensordaten benötigt, die die potentiellen Störungs- und Fehlerfälle abbilden. Deren Erzeugung stellt eine große Herausforderung dar. Die Nutzung realer Sensordaten garantiert zwar validere Ergebnisse hinsichtlich der Verifikation und Validierung. Das absichtliche Herbeiführen von kritischen Fehlern während des Fahrzeugbetriebes gestaltet sich allerdings als schwierig und ist aus Sicherheitsgründen nur bis zu einem gewissen Maße durchführbar. Zudem ist diese Vorgehensweise kostenintensiv und erfüllt nicht die Anforderungen an die Reproduzierbarkeit. Eine alternative Möglichkeit ergibt sich durch die Generierung synthetischer Sensordaten in virtuellen Simulationsumgebungen. Dieser Ansatz bietet einige Vorteile hinsichtlich Zeit- und Kostensparnis [8]. Gleichzeitig ergeben sich aber auch Ungewissheiten im Hinblick auf die Validität der Ergebnisse, da sowohl die virtuelle Umgebung als auch die Sensormodelle einen hinreichenden Detaillierungsgrad benötigen, um realitätsgenau, physikalisch korrekte und vollständige Daten zu liefern. Um den realen

Daten aus Testfahrten adäquat zu entsprechen, kann dies, je nach Anwendung, beliebig komplex werden und folglich die ursprünglichen Vorteile des Ansatzes aufheben.

Da in diesem Forschungsprojekt der Einsatz seriennaher Hardware im Fokus steht, wurde ein Lösungsansatz gesucht, welcher die Untersuchung basierend auf bestehenden, realen Sensordaten ermöglicht.

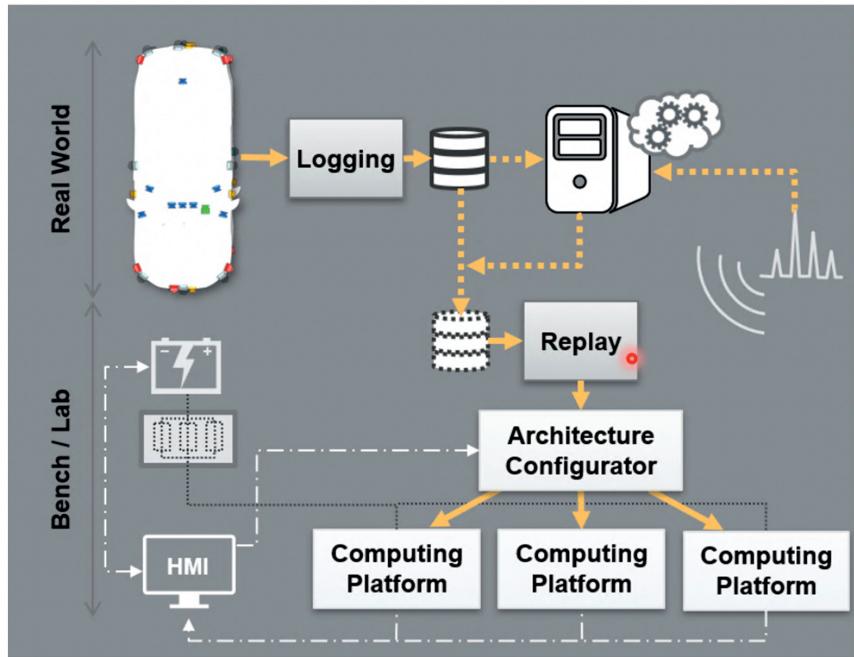


Bild 1: Testumgebung im Forschungsprojekt "AFFiAncE"

4. Manipulation realer Sensordaten zur Szenariensimulation

Einflüsse aus möglichen Stör- und Fehlerquellen können bei der Modellierung des Fahrzeugumfeldes zu Unsicherheiten und falschen Interpretationen führen und damit einen kritischen Einfluss auf die Entscheidungen der Steuerungsalgorithmen nehmen. Im Rahmen der Entwicklung neuartiger Systemarchitekturen ist die Auswahl eines optimalen Sensornetzwerks mit hinreichenden Redundanzen daher ein zentraler Bestandteil, der durch entsprechende Test- und Analysemöglichkeiten optimiert und verifiziert werden muss.

Hierfür wurde eine Software namens „Data Manipulator“ entwickelt, die die nachträgliche Einstreuung potentieller Einflüsse aus Störungen und Fehlfunktionen in bereits aufgezeichnete Sensordaten aus realen Testfahrten ermöglicht. Durch diese Manipulation können am Prüfstand verschiedene Untersuchungen hinsichtlich der Einflüsse auf die vorliegende Topologie durchgeführt werden. Hierfür wurden zunächst mögliche Stör- und Fehlerquellen kategorisiert. Im Anschluss wurden Herangehensweisen zur Ermittlung der Einflüsse möglicher Stör- und Fehlerquellen aufgestellt, um schlussendlich erste Manipulationsalgorithmen zu entwickeln.

4.1. Stör- und Fehlerquellen

Die Stör- und Fehlerquellen, denen die Sensorik in einem Fahrzeug ausgesetzt sind, lassen sich grundsätzlich, unabhängig vom Messprinzip eines Sensors, in vier Kategorien unterteilen:

- **Kommunikation:** Zur Übertragung der Daten werden die Sensoren im Fahrzeug über Leitungen im Kabelbaum mit der Rechenplattform verbunden. Dabei kommen unterschiedliche physikalische Schnittstellen, Leitungstypen und Kommunikationsprotokolle zum Einsatz. Die Elektromagnetische Verträglichkeit (EMV) ist in der Regel für gängige äußere Einflüsse gegeben. Allerdings sind außerordentliche Störungen durch (un-)absichtlich starke elektrische, magnetische oder elektromagnetische Felder nicht ausgeschlossen, die zu Fehlern in der Datenübertragung führen können. Des Weiteren können Beschädigungen oder Brüche an den Kabeln und Steckern zu Störungen bis hin zum kompletten Verlust der Datenübertragung führen.
- **Energieversorgung:** Zwecks Energieversorgung wird die Sensorik über eine Leitung an das Bordnetz des Fahrzeuges angeschlossen. Kurzschlüsse, Kabelbrüche und Lichtbögen können zu Störungen, Ausfällen und Unregelmäßigkeiten wie Spannungsabfällen führen, die einen Einfluss auf das Verhalten des Sensors und dessen Datenübertragung haben können.
- **Komponente:** Der Sensor als Komponente der Fahrzeugarchitektur stellt selbst eine weitere Quelle für Störungen und Fehler dar. Beschädigungen durch Unfälle oder Vandalismus sowie Defekte an internen Bauteilen können unvorhersehbare Störungen in der Mechanik/Hardware hervorrufen und zu einem undefinierten Verhalten am Sensorausgang führen.
- **Umwelt-/Äußere Einflüsse:** Äußere Einflüsse auf den Sensor können einerseits Verschmutzungen sein, die beispielsweise zu Einschränkungen im Sichtfeld eines Sensors führen können. Andererseits können auch starke Witterungseinflüsse wie

Starkregen, Blitzeinschläge, Nebel oder Schnee sowie direkte Sonnen- bzw. Lichteinstrahlung zu einer Degradierung der Wahrnehmungsfähigkeit eines Sensors führen und die Güte der Sensordaten reduzieren.

4.2. Herangehensweisen zur Ermittlung der Einflüsse

Der „Data Manipulator“ basiert auf Manipulationsalgorithmen, die die potentiellen Einflüsse aus Störungen und Fehlfunktionen nachträglich auf bereits vorhandene Daten einstreuen. Für deren Bestimmung lassen sich drei Herangehensweisen, abhängig von der Realitätsnähe und dem einhergehenden Aufwand, herleiten:

- **Logische/Hypothetische Einflüsse:** Manipulationsalgorithmen lassen sich mit geringem Aufwand aus theoretisch möglichen Stör- und Fehlerfällen und dem daraus logisch bzw. hypothetisch folgenden Einfluss auf das Messprinzip bzw. die Messung des Sensors ableiten. Um realistische Annahmen zu treffen kann der Einbezug von Expertenwissen helfen, um bekannte Fehlerfälle zu reproduzieren sowie neue, theoretisch mögliche Fehler herzuleiten und die Auswirkungen auf die Sensordaten richtig zu beurteilen. Diese Herangehensweise ermöglicht schnelle, kosteneffiziente Untersuchungen, um erste Tendenzen beim Vergleich von Topologien zu ermöglichen. Für fundierte Aussagen müssen diese Einflüsse allerdings durch Messungen nachgewiesen werden.
- **Metaanalysen:** Mögliche Einflüsse lassen sich zudem mithilfe einer Metaanalyse aus bereits erfolgten Untersuchungen und Tests ableiten. Im Rahmen diverser Projekte, wurde bereits eine Vielzahl an wissenschaftlichen Veröffentlichungen publiziert, in denen vor allem die Einflüsse bei schlechten Wetterverhältnissen untersucht worden sind (z.B. [9]–[11]). Den Idealfall stellt eine Untersuchung dar, bei welcher das gleiche Sensormodell zum Einsatz gekommen ist. Aber auch Ergebnisse aus abweichenden Sensorkonfigurationen lassen Rückschlüsse auf die allgemeinen Veränderungen durch die Einflüsse zu, welche in Manipulationsalgorithmen übertragen werden können.
- **Untersuchungen/Tests:** Durch die Durchführung von Testkampagnen im Realbetrieb können die Sensoren unter kontrollier- und reproduzierbaren Bedingungen möglichen Umwelteinflüssen und Störungen ausgesetzt werden. In einer anschließenden Analyse können die Sensordaten aus regulären Fahrten (Ground-Truth) mit den Daten unter dem jeweiligen Einfluss verglichen werden, um die Auswirkung auf die Daten zu ermitteln. Die Ergebnisse solcher Tests lassen sich im Anschluss in Manipula-

tionsalgorithmen überführen. Diese Herangehensweise erfordert den größten Aufwand, erzeugt allerdings auch die validesten Ergebnisse, da die Auswirkungen der Einflüsse exakt für die jeweils vorliegende Sensorkonfiguration ermittelt werden können.

4.3. Umsetzung der Manipulationsalgorithmen

In einem ersten Schritt wurde in der Zusammenarbeit mit Experten des Konsortiums ein Fehlerkatalog mit insgesamt zehn Testfällen erarbeitet, um die Eignung des „Data Manipulators“ zu verifizieren (Proof-of-Concept). Die Auswirkungen möglicher Störeinflüsse wurden entsprechend der Herangehensweise „Logische/Hypothetische Einflüsse“ ermittelt und abgeleitet. Der Fehlerkatalog umfasst Testfälle für potentielle Fehler mit einer kurzen Beschreibung des Einflusses auf die Daten, den notwendigen sowie optionalen Parametern zur Anpassung des Einflusses sowie einer abstrakten Beschreibung der Realisierung des Manipulationsalgoritmus. Hierzu gehören u.a. folgende Beispiele:

Testfall – Informationsverlust in physikalischen Sensordaten

Äußere Einflüsse sowie Schäden am Sensor können zu einer Degradation der Wahrnehmung führen. Daraus folgt ein Verlust an Informationsgehalt innerhalb der Daten. Unterschiedliche Ursachen können dabei unterschiedliche Auswirkungen auf die Informationen selbst haben, z.B.:

- **Beispiel: „Verschiebung des Sensors“**

Tatsächliche Verschiebung des Sensors z.B. durch Unfall oder einen Fehler im Kalibrierungsprozess. Hierdurch verschiebt sich das Sichtfeld des Sensors.

- **Beispiel: „Ausblendung eines Bereiches im Sichtfeld“**

Durch Verdreckung und Vandalismus (Graffiti, Zerkratzen etc.) kann ein bestimmter Teil eines Sensors verdeckt werden, was zum Verlust der Detektionen innerhalb des Abschnitts führt.

- **Beispiel: „Gischteffekt“**

Die Gischt, die bei regennasser Fahrbahn durch aufgewirbeltes Spritzwasser entsteht, führt abhängig vom Messprinzip sowie der Regen- und Windstärke zu fehlerhaften Messungen.

- **Beispiel: „Rauschen“**

Störungen, z.B. durch EMV-Einflüsse, führen zu einem Rauschen innerhalb der Daten oder zur Unterbrechung der Übertragung.

Testfall – Datenstromfehler

Störungen auf der Kommunikations- und Energieversorgungsebene können dazu führen, dass es zu Fehlern innerhalb der Datenströme kommt. Diese Einflüsse können unterschiedliche Auswirkungen haben, z.B.:

- **Beispiel „Zeitstempel / Taktzeit“**

Im normalen Zustand werden die Daten der Sensoren mit einer festen Taktzeit versendet und vom Empfänger beim Eingang mit den anderen Daten synchronisiert. Mögliche Störeinflüsse können zu instabilen Taktzeiten beim Ver- sand und Empfang führen und damit die Synchronisation der Daten beeinflus- sen.

- **Beispiel: „Datenverlust“**

Bedingt durch die steigende Leistungsdichte der Übertragungskanäle, können Störungsursachen, wie z.B. äußere EMV-Einflüsse, Auswirkungen auf die Datenübertragung haben. Dies kann zum Verlust einzelner Nachrichten bis hin zum kompletten Verlust der Verbindung führen.

Für die vorliegenden Testfälle wurden Softwarealgorithmen entwickelt, welche die Daten der einzelnen Sensortypen derart manipulieren, dass der Einfluss der Störung bzw. des Fehlers in die bestehenden Sensoraufnahmen eingebettet wird. Je nach Ausprägung des Einflusses kommen für die Umsetzung der Algorithmen regelbasierte, parametrierbare sowie stochasti- sche Methoden und Filter zum Einsatz. Der dafür benötigte Prozess der Datenverarbeitung besteht aus insgesamt drei Schritten. Die aufgezeichneten Sensordaten im Datenformat des jeweiligen Übertragungsprotokolls werden in einem ersten Schritt dekodiert, um die physika- lischen Messwerte, Detektionen und Informationen der Sensoren zu extrahieren. Die zu ma- nipulierenden Informationen werden im Anschluss an die Software übergeben, welche die Daten zwecks generischer Programmierung in einer standardisierten Klasse ablegt. Im An- schluss werden die Manipulationsalgorithmen parametriert und auf die jeweilige Instanz der Klasse angewandt. Die nun manipulierten Sensordaten werden schlussendlich wieder in ihr ursprüngliches Datenformat enkodiert am Prüfstand zur Einspeisung zur Verfügung gestellt.

4.4. Anwendungspotentiale

Mit Hilfe des „Data Manipulators“ lassen sich innerhalb der Testumgebung unterschiedliche Teile der Fahrzeugarchitektur testen. Zum einen können verschiedene Sensorkonfiguratio- nen und deren Robustheit im Falle eines Fehlers verglichen werden. Hierfür wird zunächst ein identischer Fehler in die Datenaufnahme zweier Konfigurationen eingebettet. Im An-

schluss kann das Ergebnis der Modellierung des Umfeldes als Grundlage zur Bewertung herangezogen werden. Damit lässt sich beispielsweise überprüfen, inwiefern eine redundante Auslegung eines Sensors zur Reduzierung der Unsicherheiten im Fehlerfall beiträgt.

Des Weiteren kann es zum Testen und Vergleichen unterschiedlicher Topologien einer Architektur eingesetzt werden. Die vorgestellte Testumgebung bietet die Möglichkeit, per Brute-Force Algorithmus alle logisch sinnvollen Topologien, unter Einbezug bestimmter Kriterien, zu erstellen, indem die entsprechenden physischen Verbindungen zwischen dem Sensordaten-Einspeisegerät und den drei Rechenplattformen automatisch konfiguriert werden. Anschließend können diese unter diversen Gesichtspunkten, wie z.B. Skalierbarkeit, Komplexität und Kosten, bewertet werden, um schlussendlich die optimale Lösung wählen zu können. Im weiteren Verlauf kann der Ansatz auch zum Testen weiterer, einzelner Teile der Software adaptiert werden. Insbesondere die Fehlerbehandlungsstrategien spielen für den Fail-Operational Betrieb eine zentrale Rolle. Je nach Auslegung ist ein möglicher Einsatz allerdings nur im Rahmen eines Closed-Loop Verfahrens sinnvoll, um die Reaktionen des Systems über einen längeren Zeitraum einfließen zu lassen. Denkbar wäre eine direkte Integration der Manipulationsalgorithmen in das Testverfahren, sodass die Fehler im laufenden Testbetrieb in die Daten eingebettet werden. Dies wäre einerseits virtuell mithilfe von Simulationsumgebungen und synthetischen Daten möglich, aber auch unter gesicherten Bedingungen auf realen Testumgebungen im Fahrzeug selbst denkbar.

Schlussendlich bietet der Ansatz viele Potentiale, die im Rahmen der Entwicklung neuartiger Fahrzeugarchitekturen für hoch- und vollautomatisierte Fahrzeuge für die Analyse der Zuverlässigkeit und Sicherheit genutzt werden können.

5. Zusammenfassung und Ausblick

In diesem Beitrag wurde ein Ansatz zur nachträglichen Manipulation von Sensordaten aus realen Testfahrten vorgestellt, mit dem potentielle Fehlereinflüsse auf Seiten der Sensorik durch entsprechende Algorithmen nachgebildet werden können. Seitens der Sensorik wurden hierfür diverse Fehlerquellen identifiziert und kategorisiert sowie deren Einflüsse auf die Daten näher untersucht. Im Anschluss wurden Algorithmen entwickelt, die es ermöglichen, einen Fehlerfall zu simulieren und Sensordaten entsprechend zu manipulieren. Die eigens entwickelte Anwendungssoftware „Data Manipulator“ ist dabei Teil einer flexiblen Testumgebung für die Erprobung neuartiger Fahrzeugarchitekturen mit seriennaher Hardware unter realistischen Bedingungen im Labor, welche im Rahmen des Förderprojektes „AFFiAncE“ realisiert wird. Die Ergebnisse und gewonnenen Erkenntnisse der Tests werden künftig für die Entwicklung sicherer, zuverlässiger und robuster Fahrzeugarchitekturen und neuer

Fehlerbehandlungsstrategien herangezogen, um die Entwicklung im Bereich des automatisierten Fahrens voranzutreiben und (mehr) Vertrauen in die Technik zu schaffen.

Im weiteren Verlauf des Projektes erfolgt die Durchführung des Proof-of-Concept anhand zweier Szenarien. Im Rahmen einer Testkampagne werden dafür exemplarisch die Einflüsse von Gischt auf die vorliegende Sensorkonfiguration im Testfahrzeug untersucht und deren Auswirkung auf die Daten der Sensoren ausgewertet, um daraus geeignete, verifizierte Manipulationsalgorithmen abzuleiten. Anschließend werden unterschiedliche Topologien einer ersten Systemarchitektur exemplarisch am entwickelten Prüfstand getestet, um die Wirksamkeit der Manipulation von Sensordaten aus realen Testfahrten nachvollziehbar nachzuweisen und damit den vorgestellten Ansatz zu verifizieren.

Literatur

- [1] D. C. Ehlers, „Mobility of the future – Connected, Autonomous, Shared, Electric“, S. 4, 2018.
- [2] A. Cacilo *u. a.*, *Hochautomatisiertes Fahren auf Autobahnen - Industriepolitische Schlussfolgerungen*. 2015.
- [3] D. Zerfowski und J. Crepin, „Vehicle Computer - Automotive-Softwareentwicklung neu gedacht“, *ATZ Elektron*, Bd. 14, Nr. 7, S. 36–41, Juli 2019, doi: 10.1007/s35658-019-0071-z.
- [4] L. Bauer, „Smart Vehicle Architecture - A sustainable approach to building the next generation of vehicles“, Aptiv, Whitepaper, März 2020. [Online]. Verfügbar unter: https://www.aptiv.com/docs/default-source/white-papers/2020_aptiv_whitepaper_sva.pdf?sfvrsn=18b5c13e_2
- [5] M. A. Shankar, „Future Automotive E/E Architecture“, Bd. 14, Nr. 3, S. 6, 2019.
- [6] M. Kovač, D. Reinhardt, O. Jesorsky, M. Traub, J.-M. Denis, und P. Notton, „European Processor Initiative (EPI)—An Approach for a Future Automotive eHPC Semiconductor Platform“, in *Electronic Components and Systems for Automotive Applications*, Cham, 2019, S. 185–195. doi: 10.1007/978-3-030-14156-1_15.
- [7] N. Potthoff, T. Feismann, R. Denkelmann, M. Kiffmeier, M. Ruebartsch, und S. Frei, „Optimization of Power and Signal Distribution Systems for Advanced Safety Features“, in *AmE 2020 - Automotive meets Electronics; 11th GMM-Symposium*, März 2020, S. 1–7.
- [8] D. Michalik, P. Kohl, M. Krause, A. Kummert, und T. Meisen, „Synergiepotenziale von Virtual City Twins im Bereich automatisiertes Fahren – Beschleunigung der technischen Entwicklung und Überwindung von Akzeptanzbarrieren“, in *Making Connected Mobility Work: Technische und betriebswirtschaftliche Aspekte*, H. Proff, Hrsg. Wiesbaden: Springer Fachmedien, 2021, S. 113–133. doi: 10.1007/978-3-658-32266-3_6.
- [9] W. Ritter, M. Bijelic, T. Gruber, M. Kutila, und H. Holzhüter, „DENSE: Environment Perception in Bad Weather—First Results“, in *Electronic Components and Systems for Automotive Applications*, J. Langheim, Hrsg. Cham: Springer International Publishing, 2019, S. 143–159. doi: 10.1007/978-3-030-14156-1_12.
- [10] S. Hasirlioglu, I. Doric, A. Kamann, und A. Riener, „Reproducible Fog Simulation for Testing Automotive Surround Sensors“, in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)*, Sydney, NSW, Juni 2017, S. 1–7. doi: 10.1109/VTCSpring.2017.8108566.
- [11] M. Jokela, M. Kutila, und P. Pyykönen, „Testing and Validation of Automotive Point-Cloud Sensors in Adverse Weather Conditions“, *Applied Sciences*, Bd. 9, Nr. 11, S. 2341, Juni 2019, doi: 10.3390/app9112341.

Radar Target Simulator – Key Technology for AV Development

Prasanna Kannan, Charles Miquet,
IPG Automotive GmbH, Karlsruhe

Zusammenfassung

Die Anzahl der Sensoren und die Datenmengen, die für Fahrerassistenzsysteme und autonome Fahrfunktionen verarbeitet werden müssen, steigen kontinuierlich an. Die Entwicklung und der Test dieser Systeme sind mit dem realen Fahrversuch alleine nur schwer umzusetzen. Vor allem aufgrund der nahezu unbegrenzten Menge möglicher Szenarien ist das Testen und Absichern im virtuellen Fahrversuch wichtiger denn je. Im Bereich ADAS/AV ist die Umgebungserfassung dabei eine der größten Herausforderungen. Unvorhergesehenes Verhalten von Verkehrsteilnehmern oder schwierige Umweltbedingungen dürfen zu keinen Fehlfunktionen führen.

Der Radarsensor ist ein Schlüsselement, um autonome Fahrfunktionen realisieren zu können. Zusammen mit weiteren Sensoren wie Kamera oder Lidar bildet er die Grundlage für automatisierte Fahrfunktionen. In diesem Vortrag werden die aktuellen Herausforderungen bei der Prüfung von Radarsensoren dargelegt. Sensortechnologien wie das Bildgebungsradar (bessere Winkelauflösung, Azimut- und Elevationsfähigkeit) erfordern innovative Testlösungen. Dabei soll unter anderem die aktuell existierende Lücke beim „Over-the-Air“-Test unter Verwendung modernster Technologien diskutiert werden.

Die Autoren stellen in diesem Beitrag eine umfassende Methodik zum Testen von Radarsensoren in einer virtuellen Fahrzeugumgebung vor. Das Ziel ist die Entwicklung einer Testumgebung, die jede Entwicklungsphase unterstützt – von der Konzept- und Vorserienphase bis hin zur Serienreife.

Abstract

The number of sensors and the amounts of data that need to be processed for driver assistance systems and autonomous driving functions are continuously increasing. The development and testing of these systems is difficult to implement with real driving tests alone. Especially due to the almost unlimited number of possible scenarios, testing and validation in virtual driving tests is more important than ever. In the field of ADAS/AV, environment detection is

one of the greatest challenges. Unforeseen behavior of road users or difficult environmental conditions must not lead to malfunctions.

The radar sensor is a key element for implementing autonomous driving functions. Together with other sensors such as camera or lidar, it forms the basis for automated driving functions. This presentation will outline the current challenges in testing radar sensors. Sensor technologies such as imaging radar (better angular resolution, azimuth and elevation capability) require innovative test solutions. Among others, the currently existing gap in "over-the-air" testing using state-of-the-art technologies will be discussed.

In this paper, the authors present a comprehensive methodology for testing radar sensors in a virtual vehicle environment. The goal is to develop a test environment that supports every phase of development, from the concept and pre-production phase to series production.

Autonome Fahrzeuge

Da autonome Fahrzeuge über eine große Menge unterschiedlicher Sensoren und Systeme verfügen, müssen diese untereinander kommunizieren können. Die Komplexität der Fahrzeuge muss daher im Verbund betrachtet werden. Gleichzeitig steigen der Absicherungsaufwand und der Testbedarf für den Anwender, da viele Wechselwirkungen zwischen den verbauten Komponenten auftreten. Für eine fehlerfreie Funktionsweise ist es erforderlich, dass alle Sensoren, Aktuatoren und Softwarefunktionen im Fahrzeug problemlos zusammenarbeiten können. Aufgrund der daraus resultierenden Komplexität muss schon innerhalb der Konzeptphase die Nutzung der virtuellen Fahrzeugentwicklung ermöglicht werden. Mit ihr kann etwa getestet werden, ob die Sensorik das Fahrzeugumfeld optimal erfasst oder ob es nötig ist, die Position beziehungsweise Konfiguration der Sensoren anzupassen.

Aufgrund der großen Anzahl verschiedener Technologien, deren Zusammenspiel reibungslos funktionieren muss, sowie der Anzahl der integrierten Systeme spricht man von einem „System of Systems“. Im Hinblick auf die Tatsache, dass bei automatisierten Fahrfunktionen der Level 3+ menschliche Fahrer keine Rückfallebene für das System bilden können, steigt der Testbedarf darüber hinaus in bisher unbekannten Größenordnungen weiter an.

Der Radarsensor

Die Radartechnologie wird dabei immer intelligenter und komplexer. Im Automobilbereich bieten Radarsensoren diverse Vorteile wie etwa eine sehr gute Mess- und Trennfähigkeit von Relativgeschwindigkeiten. Mithilfe eines Radars ist es möglich, Abstände zu anderen Verkehrsteilnehmern (Targets) präzise zu bestimmen, mehrere Ziele voneinander zu unterscheiden und deren Bewegungen nachzuverfolgen. So stellen Radarsensoren eine ideale Basis

dar, um Systeme wie Abstandsregeltempomaten oder autonome Notbremssysteme zu realisieren.

Aus diesem Grund ist es besonders wichtig, den Radarsensor so gründlich wie möglich zu testen und abzusichern. Die Simulation bietet dafür eine unverzichtbare Grundlage, indem sie die Umgebung, die erfasst wird, extrem realitätsnah emuliert.

Beim Test von Radarsensoren ist es wichtig, den Sensor selbst sowie die Fahrzeugintegration auf verschiedenen Ebenen zu testen. Mithilfe der Simulation wird die Umgebung, die der Radarsensor wahrnimmt, in einer Form nachgebildet, mit der der Sensor umfassend getestet werden kann. Die Radarsimulation ermöglicht den Test der gesamten Wirkkette, wodurch die Anzahl der nötigen realen Fahrversuche verringert werden kann. Eine dafür geeignete Technologie ist die Radar-Target-Simulation mithilfe der Over-the-Air-Methode. Diese ermöglicht es, Radarsensoren in den Validierungsumfang zu integrieren, ohne in den hochfrequenten Teil der Radareinheit eingreifen zu müssen.

Die in diesem Vortrag vorgestellte Radar-Testumgebung erfüllt die Anforderungen der verschiedenen Zielgruppen und ermöglicht eine Verifizierung unter Verwendung synthetischer Datenquellen, die auf phänomenologischen Modellen sowie auf Ground Truth- und Rohsignal-Modellen basieren. Darüber hinaus wird mögliche Hard- und Software vorgestellt, die zur Realisierung des Testziels beitragen kann.

Abstraktionslevel

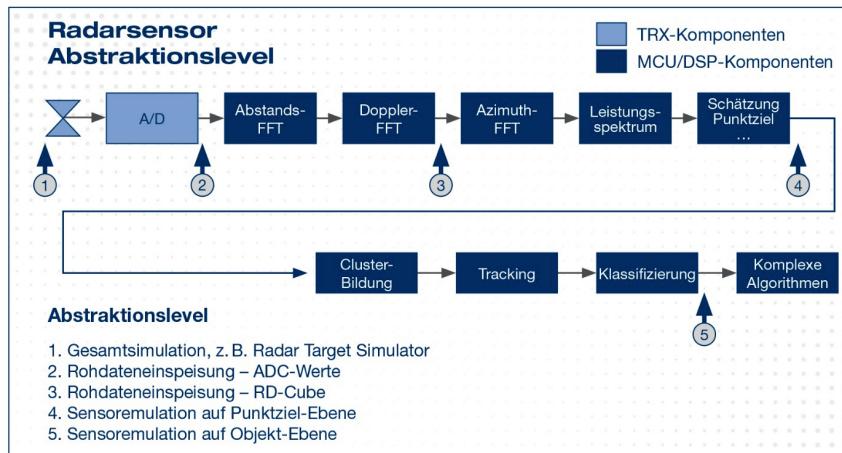


Bild 1: Abstraktionslevel des Radarsensors

Abbildung 01 veranschaulicht die typische Prozessierungskette eines Radarsteuergeräts im Automobilbereich. Das Steuergerät verfügt über verschiedene Komponenten, die für die hochfrequente Signalverarbeitung zuständig sind. Weitere Komponenten übernehmen die Prozessierung. Wie in der Abbildung ersichtlich ist, kann der Sensor auf verschiedenen Abstraktionslevels getestet werden. Für Abstraktionslevel 1 ist der sogenannte Radar Target Simulator notwendig, während bei Level 2 und 3 die Raw Signal Injection verwendet wird. Level 4 und 5 hingegen laufen über die Restbussimulation, also die Emulation auf Basis der Echtzeit-Hardware. In dieser Arbeit liegt der Fokus auf dem Abstraktionslevel 1, für den das Radar-in-the-Loop Interface entwickelt wurde. Die Modellierungsgüte der verschiedenen Sensormodellklassen wird im folgenden Kapitel vorgestellt.

Modellierungsgüte

Um verschiedene Modellierungstiefen bieten zu können, verfügt die offene Integrations- und Testplattform CarMaker über drei Sensorklassen für die verschiedenen Sensortypen, siehe Abbildung 02. Durch sie können einerseits Subsysteme mit speziell angepassten Informationen versorgt werden, andererseits wird aber auch zu jeder Zeit die höchste Komplexität der Sensorsimulation für den Test des Gesamtsystems gewährleistet. Im Folgenden werden die unterschiedlichen Sensorklassen, ihre Anwendungsgebiete sowie die jeweiligen Alleinstellungsmerkmale erläutert.

Sensoren		
	Ideal	HiFi
Anwendungs-fälle	<ul style="list-style-type: none"> grundlegende Funktionstests Referenzsensor 	<ul style="list-style-type: none"> Funktionstests Fail-Safe-Tests
Eigen-schaffen	<ul style="list-style-type: none"> Output: Objektliste Ground-Truth-Informationen technologieübergreifend einfache Parametrierbarkeit 	<ul style="list-style-type: none"> Output: Objektliste technologiespezifische Fehler physische Effekte inkl. Nachbearbeitung/ Zielauswahl 
		<ul style="list-style-type: none"> Entwicklung des Perzeptionsalgorithmus Test der Signalkette 

Bild 2: Die verschiedenen Sensormodellklassen in CarMaker

Klasse 1: Ideale Sensormodelle



Bild 3: Abbildung des idealen Sensormodells

Der Zweck idealer Sensormodelle (siehe Abbildung 03) ist die Ausgabe einer Liste von relevanten detektierten Objekten. Eine Extraktion aus dem Simulationsmodell stellt die dafür benötigten Informationen bereit. Auf dieser Basis wird das Umfeld auf eine technologieunabhängige, ideale Art und Weise erfasst – die Sensorphysik wird dabei nicht im Detail betrachtet und sensortypische Fehler werden nicht in der Objektliste abgebildet.

Ideale Sensormodelle bieten verschiedene Vorteile, etwa eine sehr hohe Performanz und eine sehr einfache Parametrierbarkeit. Ihr Zweck ist es, Fehler in der Perzeption auszuschließen. Aufgrund dessen sind sie insbesondere für grundlegende Tests der verschiedenen Fahrzeugfunktionen geeignet. Zudem bieten sie die Möglichkeit, für den Test der Perzeption und Sensorfusionsalgorithmen als Referenz (Ground Truth) eingesetzt zu werden. Auf diese Weise kann eine aufwendige manuelle oder teilautomatisierte Annotation von Sensordaten vermieden werden. Simulationsszenarien müssen für die Verwendung mit idealen Sensormodellen

zudem nicht sehr detailliert sein. Hier ist es ausreichend, wenn relevante Objekte wie Verkehrsteilnehmer, Schilder oder Straßenmarkierungen in geringer Modellgenauigkeit vorhanden sind.

Klasse 2: Hi-Fi-Sensormodelle

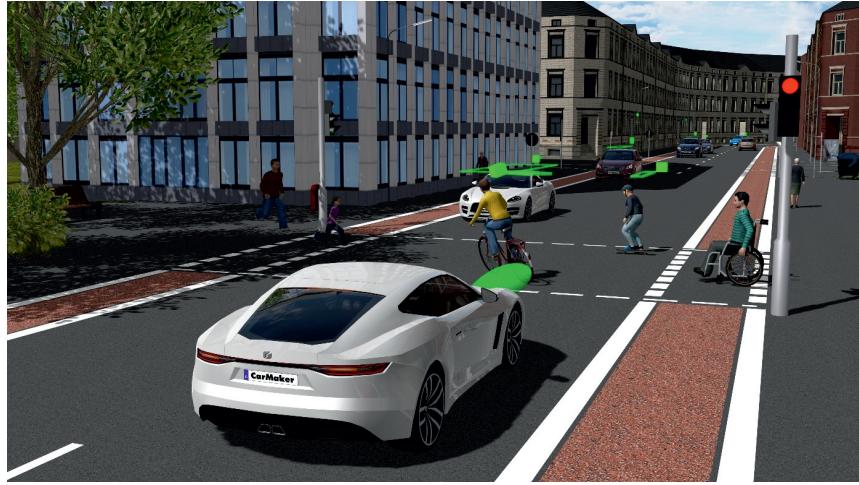


Bild 4: Abbildung des Hi-Fi-Sensormodells

Wie die zuvor erläuterten idealen Sensormodelle geben auch die Hi-Fi-Sensormodelle bzw. phänomenologischen Sensormodelle (siehe Abbildung 04) eine Objektliste aus. Im Gegensatz zu einer reinen Extraktion aus dem Simulationsmodell werden in diesem Fall die Informationen durch physikalische Effekte angereichert. Auch technologiespezifische, teils stochastische Fehlermodelle können berücksichtigt werden. Auf diese Weise kann eine realitätsnahe Objektliste für die jeweilig verwendete Sensortechnologie ausgegeben werden.

Da die Signalausbreitung und -verarbeitung durch vereinfachte Modelle abgebildet werden, wird im Regelfall eine hohe Performanz erzielt. Während physikalische Effekte im Mittel korrekt abgebildet werden, kann es im Einzelfall dennoch zu Abweichungen kommen.

Ein mögliches Einsatzgebiet der Hi-Fi-Sensormodelle sind etwa Robustheitstests von Fahrzeugfunktionen. Diese sollen überprüfen, ob Funktionen korrekt ausgeführt werden, obwohl Sensoren fehlerhafte Informationen bereitstellen.

Um die Detektion der Objekte zu beeinflussen bzw. False Negatives oder False Positives hervorzurufen, sind ergänzend weitere Objekte nötig. Im Falle der Radarsensoren sind insbeson-

dere Leitplanken und Tunnel relevant, da die Reflexionen der Radarstrahlen dazu führen können, dass sogenannte Geisterobjekte auftreten. Anspruchsvolle Situationen für die Objektdetektion können allerdings mit dieser Sensormodellklasse nur eingeschränkt untersucht werden. Die Ursache hierfür liegt in vereinfachten Signalausbreitungsmodellen oder zufällig auftretenden Effekten.

Klasse 3: Raw Signal Interfaces, RSI

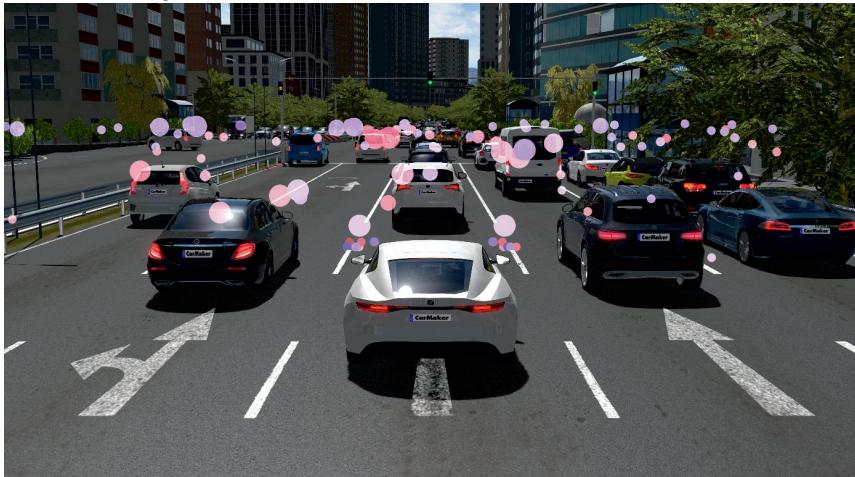


Bild 5: Abbildung des Raw Signal Interfaces

Raw Signal Interfaces (siehe Abbildung 05) sorgen für die Bereitstellung von Eingangsdaten für die Perzeptionsalgorithmen des Sensors. Dies können etwa Bilddaten für die Kamerasimulation oder die sogenannte Kanalimpulsantwort für Radar sein.

Bei der Generierung von Rohsignalen werden detaillierte physikalische Effekte sowie die Materialeigenschaften der Objekte berücksichtigt. Es ist also zwingend notwendig, dass die Umgebung entsprechend detailliert modelliert wurde – alle Oberflächen müssen mit den jeweiligen technologieabhängigen Materialeigenschaften versehen werden. Dies führt zu deutlich höheren Anforderungen an die Hardware, insbesondere im Hinblick auf den Grafikprozessor (Graphics Processing Unit, GPU).

Die Rohsignalschnittstelle kann etwa für den Funktionstest einer Sensorkomponente bzw. des dahinterstehenden Detektionsalgorithmus eingesetzt werden. Der Detaillierungsgrad ist in diesem Fall der höchste, der in Verbindung mit einer Echtzeitsimulation zu erzielen ist. Die von

der Rohsignalschnittstelle gelieferten Radarsignale können dazu genutzt werden, zu überprüfen, ob die Signalverarbeitung in der Lage ist, daraus die korrekte Fahrzeugklasse sowie die Geschwindigkeits- und Abstandswerte abzuleiten. Verschlüsselungsmechanismen und standardisierte Schnittstellen unterstützen dabei, auch zugekauft Sensoren beim Anwender einzubinden, ohne Zugang zum geistigen Eigentum des Zulieferers zu benötigen.

Der zuvor erwähnte Abstraktionslevel 1 bei der Prozessierungskette eines Radarsteuergeräts verwendet alle drei Sensorklassen, während in Abstraktionslevel 2, 3 und 4 die Raw Signal Interfaces eingesetzt werden. Level 5 hingegen nutzt die Hi-Fi-Sensormodelle.

HIL

Die Sensor-Emulation mithilfe der CarMaker-Sensormodelle wird von der Autonomous Drive Emulation-(ADE-)Architektur [1] von Keysight Technologies unter Verwendung der Hardware-Plattform Xpack4 in Echtzeit berücksichtigt. Über die Video Interface Box X, die für die Emulation des Bildsensors zuständig ist, kann die Kameraemulation erfolgen. Unter Verwendung eines Over-the-Air-(OTA-)Simulators – dem Radar Target Simulator (RTS) – wird die Radarsensorsimulation durchgeführt.

Bestimmte Fahrzeugregler benötigen zusätzliche Eingangssignale, zum Beispiel Geschwindigkeit, Gieren oder Nicken. Diese Signale können ebenfalls mithilfe der Xpack4-Plattform verarbeitet werden. Über die Echtzeit-Hardware werden die Aktuatorsignale den entsprechenden Subsystemen im Fahrzeugmodell zugespielt, wobei alle notwendigen Technologien wie CAN-FD, FlexRay oder SOME/IP unterstützt werden. So wird eine möglichst realitätsnahe Emulation des Fahrzeugaufbaus ermöglicht.

Der gesamte Radarsensor wird dabei als eine Art Blackbox angesehen. Demnach werden keine Annahmen über das Steuergerät benötigt, die in der Modellierung zu einer Verfälschung führen könnten. Weil außerdem keine Schnittstelle im Steuergerät vorausgesetzt ist, gestaltet sich die Integration der Sensoren in das Testsystem als verhältnismäßig einfach. Durch eine unmittelbare Verknüpfung der Keysight-Hardware [2] mit der Umfeldsimulation CarMaker können auf diese Weise beliebig viele Targets und Objekte dargestellt werden, ohne dass ein mechanischer Ansatz nötig ist.

Radar-in-the-Loop Interface: Überblick

Das Radar-in-the-Loop Interface wurde entwickelt, um aufbereitete Zieldaten für RTS bereitzustellen, die über mehrere Antennenelemente verfügen. Diese Zieldaten werden entweder

von einem virtuellen Objektsensor oder einem Raytracing-basierten Radarsensormodell generiert, das von CarMaker bereitgestellt wird. Hauptmerkmal des Radar-in-the-Loop Interfaces ist die Auswahl, Verteilung und Übertragung von Zieldaten in einem segmentbasierten Format. Dabei können die Anzahl bzw. die Winkel der Segmente, die Zielauswahl und andere radar-spezifische Variablen über Parameter eingestellt werden.

Abbildung 06 veranschaulicht die Komponenten einer RTS-Testsystemumgebung. Die Car-Maker-Hauptsimulation läuft auf einem Echtzeitsystem, das eine konstante Zykluszeit garantiert und außerdem die IO-Hardware für die DUT-Restbussimulation bereitstellt, welche für Closed-Loop-Anwendungen erforderlich ist.

Sowohl die Erstellung von Testszenarien als auch die Bedienung der Simulation erfolgen über die CarMaker/HIL-GUI, die auf einem Front-End-PC installiert ist. Ein GPU-Sensor-Framework ermöglicht es dem Hauptsimulationsprogramm, die GPU-Rechenleistung des Front-End-PCs für komplexe Radarsensorberechnungen zu nutzen. Die vom Radar-in-the-Loop Interface verarbeiteten Zieldaten werden über UDP/IP-Pakete über eine Gigabit-Ethernet-Verbindung an den Radar Target Simulator übertragen.

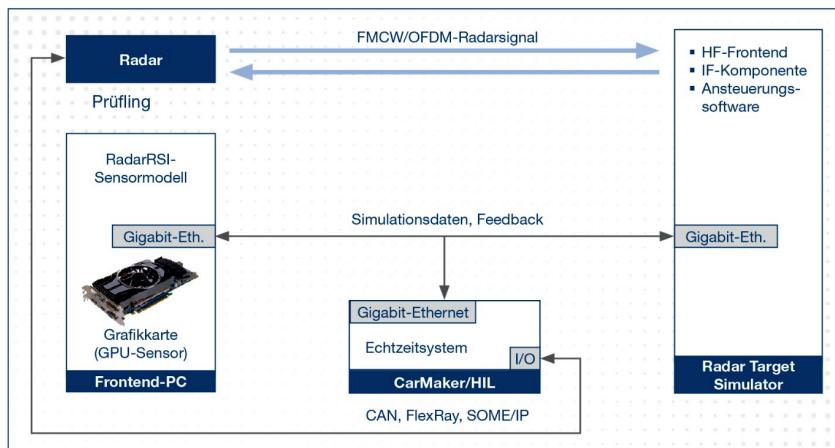


Bild 6: Die Komponenten einer RTS-Testsystemumgebung

Sensortypen des Radar Target Simulators:

Objektsensorbasierte Zieldaten

Objektbasierte Zieldaten, die aus dem in CarMaker enthaltenen Objektsensor Ground Truth generiert werden, sind ein guter Ausgangspunkt für eine reproduzierbare und leicht verständliche Simulation. Der Objektsensor detektiert die Objekte im Sichtfeld des Sensors und stellt

sie als Objektliste zur Verfügung. Dieser Ansatz erzeugt ein Ziel pro Verkehrsobjekt. Jedes detektierte Objekt trägt die Objektattribute Reichweite (R), Geschwindigkeit (v), Azimutwinkel (φ) Elevation (θ) entweder am Referenzpunkt oder am nächstgelegenen Punkt der Verkehrsobjekte. Um die Radar Cross Section (RCS) des Verkehrsobjekts zu berechnen, wird der Objektsensor außerdem durch ein phänomenologisches Modell ergänzt. Die Attribute jedes qualifizierten Objekts werden an den Radar Target Simulator gesendet, um die virtuelle Umgebung für den Radarsensor Under Test (RUT) zu simulieren, siehe Abbildung 07.

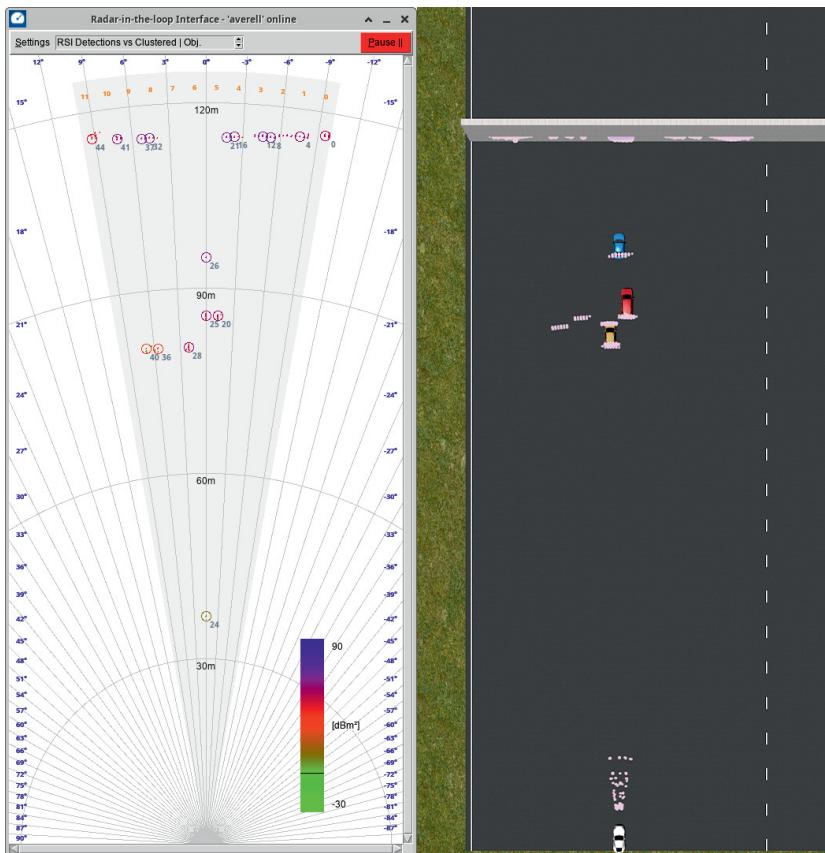


Bild 7: GUI der Radar-in-the-Loop Interface

Raytracing-basierte Zieldaten

Die auf Raytracing basierenden Zieldaten werden aus dem Sensormodell des Raw Signal Interface (RSI) erzeugt. Der Rohsignalsensor simuliert die polarimetrische Ausbreitung der elektromagnetischen Welle mittels Raytracing. Das RSI-Sensormodell in CarMaker wird mit den gleichen Randbedingungen wie das RUT parametrisiert, z. B. Field of View (FoV), Reichweite oder Einbaulage des Sensors. Zusätzlich zu den Standardparametern werden auch Mittenfrequenz, Sendeleistung, Antennencharakteristiken, Auflösung und Genauigkeit in Bezug auf Reichweite, Geschwindigkeit und Winkel des Radarmodells entsprechend der Spezifikation des RUT parametrisiert. Das RSI-Sensormodell liefert eine Radarpunktwolke als Ausgabe, die die Attribute Reichweite (R), Geschwindigkeit (v), Azimutwinkel (ϕ) und Empfangsleistung (P)/RCS enthält. Ein adaptiver Segmentierungs- und Clustering-Algorithmus wird auf die Punktwolke angewandt, wobei die einzelnen Ziele anhand der Segmente, in denen sie sich befinden, sortiert werden. Die Anzahl der Segmente ergibt sich aus der Anzahl der auf der RTS-Hardware vorhandenen Front-Ends oder kann frei parametrisiert werden, wenn der RTS den Ankunftsinkel (Angle of Arrival, AoA) durch Manipulation der Phase des IF-Signals simuliert. Jedes Segment deckt einen bestimmten Bereich von Azimutwinkeln innerhalb des FoV des Radars ab. Die Identifizierung der Segmente, zu denen die Ziele gehören, erfolgt auf Grundlage ihrer Azimutwinkel. Wenn die Anzahl der Ziele pro Segment innerhalb einer Stichprobe das parametrisierte Maximum überschreitet, werden nur die Ziele mit signifikanter RCS ausgewählt. Sobald die Ziele segmentweise sortiert sind, werden sie zu Clustern zusammengefasst. Die Ziele mit der größten oder durchschnittlichen Leistung in jedem Cluster werden aufgezeichnet und an den RTS gesendet. Der RTS erzeugt dann das Radar-Echo auf der Grundlage der Radar-Signatur dieses virtuellen Ziels.

Szenario

Neben geeigneten Sensormodellen werden für den Test von Fahrerassistenzsystemen und autonomen Fahrfunktionen zwingend realitätsnahe Verkehrsszenarien benötigt. Ein Szenario stellt dabei ein Abbild einer realen Verkehrssituation dar und lässt sich in fünf verschiedene Ebenen unterteilen. Die erste Ebene bildet die Straßenebene, in der grundlegende Daten wie die Anzahl der Fahrspuren und der Straßenverlauf definiert werden. Die zweite Ebene beschreibt die Straßeninfrastruktur, etwa Geschwindigkeitsbegrenzungen oder Verkehrsschilder. Ebene drei stellt temporäre Beeinflussungen der ersten beiden Ebenen dar, beispielsweise durch Baustellen oder Straßensperrungen. Bewegliche Objekte wie andere Verkehrsteilnehmer werden in Ebene vier definiert, während Ebene fünf Umweltbedingungen wie Wetter und Tageszeit umfasst.

Da Szenarien unabhängig vom Testfahrzeug sind, lassen sie sich vielseitig verwenden. Auch ist es möglich, bereits existierende Daten in die Simulation zu übertragen, um sie als Basis für die Szenarienerstellung zu nutzen. Hierfür können etwa Unfalldatenbanken oder bereits durchgeführte Fahrversuche dienen.

Jedes Verkehrsobjekt hat festgelegte Manöver, die seine Bewegung innerhalb des Szenarios festlegen. Diese Manöverabläufe lassen sich dabei für die Längs- und Querdynamik frei definieren. Als Alternative zu festgelegten Manövern können sich Verkehrsteilnehmer auch autonom im Szenario bewegen, indem sie die entsprechenden Verkehrsregeln einhalten und auf andere Verkehrsteilnehmer reagieren. Von jedem Szenario können verschiedene Variationen erzeugt werden, die dabei unterstützen, die Robustheit des Systems durch die Erweiterung des Testraums zu gewährleisten.

3-D-Welt

Abhängig vom Detailgrad der Sensorsimulation bieten sich verschiedene Abstraktionsgrade für die geometrische Modellierung der Umwelt an. Beispielsweise benötigen einfachere Sensormodellklassen keine besonders detaillierte Umgebung, da sie geometrische Effekte abstrahiert abbilden.

Beim RSI-Sensormodell hingegen wird vorausgesetzt, dass die 3-D-Umgebung einen hohen Detailgrad bzw. eine entsprechende Güte aufweist und entsprechend für das Sensormodell angepasst ist. Für die präzise Sensorsimulation werden hier Raytracing- und Rasterisierungsverfahren eingesetzt. Für die korrekte Funktionsweise der Sensormodelle bzw. ein realistisches Sensormodellverhalten ist es elementar wichtig, dass sämtliche Materialien korrekt parametriert sind (Beispiel: Stoßstange aus Kunststoff, restliche Bauteile aus Blech), um etwa die Reflexionen von Radarstrahlen korrekt berechnen zu können. Mithilfe der Objektdatenbank in CarMaker, die sehr viele Modelle von Gebäuden, Fahrzeugen etc. enthält, können Szenarien mit geringem Aufwand entsprechend ausgestattet werden.

Zusammenfassung

Der Radarsensor bildet in der Automobilindustrie ein Schlüsselement für autonome Fahrfunktionen. Um diese Fahrfunktionen im virtuellen Fahrversuch testen und absichern zu können, wird eine geeignete Testumgebung benötigt.

In diesem Paper wurde die Entwicklung eines entsprechenden OTA-Simulators (Radar Target Simulator, RTS) sowie der dazugehörigen Schnittstelle für CarMaker erläutert. Das Ziel dieser Schnittstelle – dem Radar-in-the-Loop Interface – ist es, die Umgebung so detailliert wie möglich darzustellen, wobei die physikalischen Grenzen des RTS berücksichtigt werden.

Die dabei entstandene Testmethode ermöglicht sowohl den reproduzierbaren Test als auch die Absicherung autonomer Fahrfunktionen in beliebig komplexen Verkehrssituationen. So kann die Lücke zwischen realem Fahrversuch und der reinen Simulation geschlossen werden.

Referenzen

- [1] Autonomous Drive Emulation, Abgerufen am 13. Oktober 2021, von <https://www.keysight.com/de/de/cmp/2020/autonomous-drive-emulation.html>
- [2] Automotive Radar Signal Analysis and Generation Solutions, Abgerufen am 15. Oktober 2021, von <https://www.keysight.com/de/de/products/application-specific-test-systems/automotive-radar-signal-analysis-generation-solutions.html>

Realisierung einer querführenden Fahrerassistenzfunktion mithilfe von adaptiver Regelung und neuronalen Netzen

Laura Witt, M. Sc., Dr.-Ing. Daniel Münning,
Dr.-Ing. Hendrik Oschlies,
Volkswagen AG, Wolfsburg;
Priv.-Doz. Dr.-Ing. S. Schmidt,
Otto-von-Guericke-Universität, Magdeburg

Zusammenfassung

Die Nachfrage und Leistungsfähigkeit aktueller Fahrerassistenzsysteme steigt stetig an, wobei deren Applikation eine umfangreiche und anspruchsvolle Aufgabe darstellt [1 bis 3]. Um ein optimales Kundenerlebnis zu ermöglichen, ist eine fahrzeugspezifische Applikation der Fahrerassistenzfunktion notwendig, welche sowohl Unterschiede aufgrund von Bauteiltoleranzen als auch Veränderungen über die Lebensdauer eines Fahrzeugs berücksichtigt. Adaptive Regelungsverfahren stellen dafür eine Realisierungsmöglichkeit dar.

In diesem Beitrag wird ein Ansatz zur Realisierung einer adaptiven Regelung für ein querführendes Fahrerassistenzsystem mithilfe von maschinellem Lernen vorgestellt, welches im Realversuch erprobt wird. Durch einen Vergleich der Mittenabweichung wird gezeigt, dass mithilfe eines neuronalen Netze eine vergleichbar gute Querführungsgüte wie mit klassischen Querführungsassistenzfunktionen erreicht werden kann.

1. Einleitung

Fahrerassistenzsysteme (FAS) sind elektronische Systeme, die den Fahrzeugführer bei seiner Fahraufgabe unterstützen. Bereits in den vergangenen Jahren zeigte sich eine deutliche Reduzierung der Todesopfer bei Unfällen durch die Entwicklung von Fahrerassistenzsystemen, wie beispielsweise ESP, ACC und Spurhalteassistenten. Da sowohl die Nachfrage nach diesen Funktionen als auch deren Performance steigen, werden die Systeme vermehrt in aktuellen Fahrzeugen verbaut [1,2].

Auf dem Weg zum autonomen Fahren stellen Fahrerassistenzsysteme eine Zwischenstufe dar. Während beispielsweise der Autonomous Vehicles Readiness Index (AVRI) der KPMG aus dem Jahr 2020 zeigt, dass in Deutschland vor allem im Bereich der Infrastruktur und Kundenakzeptanz noch deutlicher Aufholbedarf auf dem Weg zum autonomen Fahren notwendig

ist, zeigt eine Umfrage der Kampagne „Bester Beifahrer“ des Deutschen Verkehrssicherheitsrates (DVR), dass 2020 rund 80% der Befragten ein Fahrzeug mit mindestens einem FAS gekauft haben und dieses überwiegend auch regelmäßig nutzen .

Laut Gasser et al. in sind Fahrerassistenzsysteme in drei Kategorien einteilbar . Funktionen der Kategorie B steigern vor allem den Fahrkomfort und sind jederzeit vom Fahrer übersteuerbar. Das System nimmt hierbei direkten Einfluss auf das Fahrzeugverhalten. Spurhalteassistenten, als Beispiel für diese Kategorie, unterstützen den Fahrer aktiv dabei das Fahrzeug in der Fahrspurmitte zu halten.

Zur Realisierung einer solchen Funktionalität werden Informationen aus verschiedenen fahrzeuginternen Quellen verwendet, welche bereits einzeln betrachtet sehr komplex sind. Infolgedessen ist die Applikation eines FAS eine anspruchsvolle Aufgabe, da für jede Funktion eine Vielzahl an Parametern festzulegen ist . Aufgrund der umfangreichen Konfigurationsmöglichkeiten, z. B. aus Motorisierung, Lenkungsart und Reifenmodell, ist eine kombinationsindividuelle Applikation sowohl in Bezug auf die Kosten als auch auf die Zeit bislang nicht möglich. Um jedoch ein optimales Kundenerlebnis in jedem Fahrzeug zu ermöglichen, sollte jedes Fahrzeug eine individuelle Anpassung der Parameter erhalten. Mithilfe einer solchen fahrzeugspezifischen Applikation können neben der angesprochenen Kombinationsvielfalt aufgrund der Fahrzeugkonfigurationsmöglichkeiten ebenfalls die tolerierbaren Streuungen der Bauteile, beispielsweise in Zusammenhang mit der Lenkung oder den Fahrzeugachsen, berücksichtigt werden. Um zusätzlich Veränderungen der Bauteile über die Lebenszeit oder auch äußere Einflüsse wie beispielsweise Änderungen in der Fahrzeugbeladung zu berücksichtigen, müssen die Reglerparameter kontinuierlich angepasst werden.

Dieser Beitrag betrachtet einen Ansatz des maschinellen Lernens, um eine querführende Fahrerassistenzfunktion fahrzeugspezifisch zu realisieren. Dabei wird zuerst auf den aktuellen Stand der Technik im Bereich von adaptiven Regelungsverfahren im Zusammenhang mit der Fahrzeugregelung eingegangen. Im Anschluss erfolgt eine Vorstellung des verwendeten Versuchsträgers sowie des Softwaresetups. Basierend auf dem Stand der Technik wird ein möglicher Lösungsansatz zur Realisierung einer querführenden Fahrerassistenzfunktion mithilfe von adaptiver Regelung durch künstliche neuronale Netze vorgestellt. Abschließend werden erste Ergebnisse aus einer Versuchsfahrt gezeigt.

2. Stand der Technik

Im Bereich der quer- und längsführenden Regelung von Fahrzeugen verschiedener Art existieren unterschiedliche Ansätze von adaptiven Regelungen. Im Rahmen dieser Arbeit werden adaptive Ansätze in *klassische adaptive Regelung* und *adaptive Regelung mit maschinellen Lernverfahren* unterteilt.

Klassische adaptive Regelungsverfahren

Im klassischen Bereich der adaptiven Regelung sind vor allem modellbasierte Verfahren, wie beispielsweise Model Reference Adaptive Control (MRAC), vertreten. Bei MRAC wird ein Modell verwendet, welches parallel zur Regelung geschaltet wird. Der prinzipielle Aufbau hierzu ist in Bild 1 zu sehen.

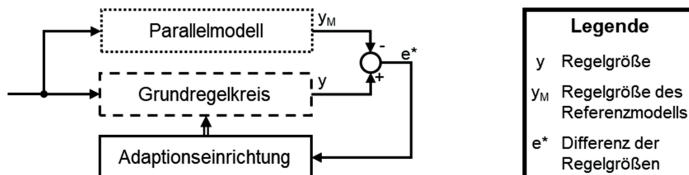


Bild 1: Model Reference Adaptive Control angelehnt an

Basierend auf identischen Eingangssignalen ermittelt das Modell die Regelgröße y_M und der Grundregelkreis die Regelgröße y . Die Differenz, bezeichnet als e^* , wird als Eingangssignal der Adaptionseinrichtung verwendet. Bei diesem Verfahren wird das Verhalten der Regelung somit an das fest vorgegebene Verhalten des Parallelmodells adaptiert.

Model Reference Adaptive Control wurde in den vergangenen Jahren beispielsweise von Ahmadian et al. zur Realisierung eines stabilen Fahrzeugverhaltens verwendet. Unbekannte Veränderungen in der Fahrzeugdynamik sollen hier bei schnellen, kritischen Spurwechselmanövern betrachtet werden. Als Referenzmodell der Regelung dient ein Einspurmodell mit zwei Freiheitsgraden. In allen untersuchten Fahrszenarien zeigt sich durch MRAC in der Simulation eine Verbesserung des Fahrzeugverhaltens, sowohl in Bezug auf das Fahrzeughandling als auch auf die -stabilität. Aufgrund des verwendeten Einspurmodells wurde nur eine begrenzte Anzahl von Einflüssen auf das Fahrzeug untersucht.

Die vorgestellte Anwendung von MRAC sowie mehrere Forschungsbeiträge im Bereich der Drohnen und Unterwasserfahrzeugregelung zeigten in der Simulation, dass mithilfe dieses adaptiven Verfahrens eine Performanceverbesserung verschiedener Fahrzeuge erreicht werden kann. Dabei ist unter anderem die Nachbildung eines Referenzmodells ausschlaggebend.

Da bei einer querführenden Fahrerassistenzfunktion viele verschiedene Sensoren und Komponenten zusammen agieren, ist eine genaue Modellierung aller notwendigen Komponenten schwierig und aufgrund der hohen Variantenvielfalt der Fahrzeuge aufwendig. Zusätzlich werden vor allem Umwelteinflüsse in Simulationen häufig nicht ausreichend berücksichtigt. MRAC stellt somit im Rahmen der Arbeit für die betrachtete Anwendung kein geeignetes adaptives Regelungsverfahren dar.

Adaptive Regelungsverfahren mit maschinellem Lernen

Im Gegensatz zu dem vorherigen klassischen Ansatz werden bei Verwendung von adaptiven Regelungsverfahren mit maschinellem Lernen keine Referenzmodelle benötigt.

Maschinelles Lernen ist ein Teilgebiet der künstlichen Intelligenz. Basierend auf Daten bzw. vergangenen Erfahrungen werden Problemstellungen gelöst, indem die Parameter der zuvor definierten Modelle anhand einer gegebenen Metrik angepasst werden. Eine mögliche Modellklasse sind dabei künstliche neuronale Netze (KNN), welche mithilfe verschiedener Lernverfahren auf Grundlage von Datenpaaren trainiert werden können.

Viele Forschungsarbeiten beschäftigen sich mit dem direkten Lernen der Stellgröße und ersetzen somit den gesamten klassischen Regelansatz durch maschinelle Verfahren. Nachfolgend wird eine Übersicht verschiedener Veröffentlichungen in diesem Bereich dargestellt.

Erste Erfolge bezüglich des Erlernen eines diskreten Lenkwinkels für ein autonomes Landfahrzeug erzielte Pomerleau mit ALVINN (Autonomous Land Vehicle In a Neural Network) bereits 1989 in . ALVINN präzisiert mit einem dreischichtigen KNN die Fahrtrichtung basierend auf einem Kamerabild und Laserdaten, um einem Straßenverlauf zu folgen. Die Fahrtrichtung wird hierbei durch 45 diskrete Lenkwinkel repräsentiert. Im Rahmen der Testbedingungen konnte ein vergleichbar gutes Verhalten wie bei einem Algorithmus ohne KNN erzielt werden, wobei die benötigte Rechenleistung geringer war. Eine allgemeine Aussage bzgl. der Performance in allen Fahrsituationen wird nicht getroffen, da für diese ausgiebigeres Erproben notwendig ist.

In wird das Training von Pomerleau anhand von Live-Fahrerdaten durchgeführt. Durch Manipulation der Eingangsbilddaten entsteht eine größere Trainingsdatenvarianz, die zur vollständigen Repräsentation der Fahraufgabe beitragen soll. Auf Basis dieser Daten konnten erfolgreiche Fahrten mit ALVINN auf verschiedenen Strecken realisiert werden, bei denen die Berechnungen viermal schneller als bei anderen Systemen ohne KNN ist.

Bojarski et al. verwenden in ein Convolutional Neural Network (CNN), um aus Kamerabildpixeln ein diskretes Lenkommando in Form einer Krümmung zu präzisieren. Als Trainingsdaten

dienen hierbei Aufnahmen von menschlichen Fahrern, welche um zusätzlich generierte Fahr-situationen, beispielsweise außerhalb der Spurmitte, erweitert wurden, um die Trainingsdaten-varianz zu steigern. Basierend auf diesen Daten konnte ein künstliches neuronales Netz tra-i-niert werden, welches sowohl in Simulations- als auch in Realfahrten die Umweltwahrneh-mung, Bahnplanung und Fahrzeugregelung erfolgreich übernehmen kann. Auf verschiedenen Strecken und unter verschiedenen Umweltbedingungen war es möglich, dass Fahrzeug na-hezu komplett ohne Eingriffe des Fahrers in der Fahrspur zu halten.

Hecker et al. erweitern in die bisher überwiegenden Ansätze mit frontgerichteter Kamera auf ein Rundum-Kamera-System mit vier Kameras sowie um einen Routenplaner. Motiviert wird diese Erweiterung dadurch, dass einem menschlichen Fahrer ähnliche Informationen eben-falls zum Lösen seiner Fahraufgabe zur Verfügung stehen. Mithilfe verschiedener CNNs wer-den aus den einzelnen Kamerabildern Features extrahiert, welche anschließend in mehreren Long Short-Term Memories (LSTMs) weiterverarbeitet werden. Die Fusion der entsprechen-den zeitabhängigen Bildinformationen mit den Positionsinformation des Routenplaners erfolgt durch ein vorwärtsgerichtetes KNN, welches einen Lenkradwinkel zwischen $\pm 720^\circ$ und eine Geschwindigkeit prädiziert. In experimentellen Vergleichen anhand von aufgenommenen Da-ten konnte gezeigt werden, dass durch die zusätzlich verwendeten Kameras eine Verbesse- rung der Prädiktionsgenauigkeit für Lenkwinkel und Geschwindigkeit erreicht wird. Zusätzlich wurde gezeigt, dass ein Routenplaner eine Verbesserung der Performance bewirken kann.

Die vorgestellten Paper zeigen, dass die Realisierung einer Fahrzeugquerföhrung mithilfe von neuronalen Netzen möglich ist. Dabei werden überwiegend Bilddaten und folglich CNNs ver-wendet. In zeigt sich die Möglichkeit die Performance durch Fusion mit anderen Daten die Performance zu steigern.

Im Gegensatz zu den vorgestellten Arbeiten sollen im Rahmen dieses Beitrags keine Bilddaten sondern fahrzeuginterne Signalquellen verwendet werden, um ein kontinuierliches Lenkungs-signal zu prädizieren.

3. Hardware- und Softwaresetup

Zur Umsetzung einer Querführungsregelung durch neuronale Netze in einem realen Fahrzeug wurde ein Setup entwickelt, welches nachfolgend vorgestellt wird.

Hardwaresetup

Zur realen Erprobung des implementierten Ansatzes für eine Querführungsassistenzfunktion durch neuronale Netze wird ein Versuchsträger mit einem prototypischen Steuergerät, einer MicroAutoBox II (MABXII) der Firma dSPACE, ausgestattet. Der Aufbau ist in Bild 2a zu sehen. Die Verbindung zwischen Modell und Fahrzeug wird hierbei mittels Controller Area Network (CAN) und Ethernet hergestellt.

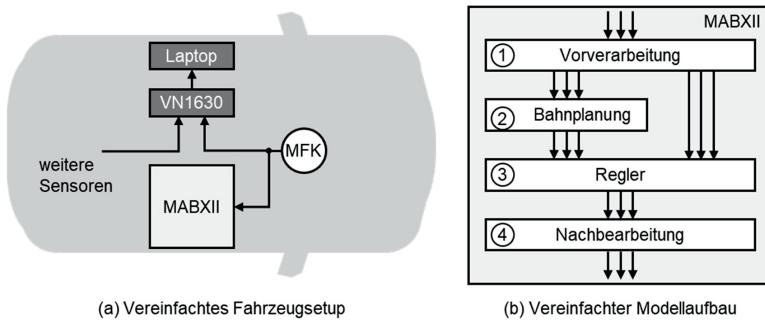


Bild 2: Vereinfachter Aufbau des Fahrzeugs und des Modells der Fahrerassistenzfunktion

Für die Spurerkennung im Rahmen der Fahrerassistenzfunktion ist das Steuergerät der Multi-funktionskamera (MFK) relevant. Der integrierte Algorithmus extrahiert die Informationen zu den Spurmarkierungen aus den aufgenommenen Bilddaten und sendet diese über eine Ethernetverbindung an das Prototypensteuergerät. Zusätzlich werden an dieses per CAN Informationen aus anderen Sensoren, wie beispielsweise die Fahrzeuggeschwindigkeit oder die Gierrate, gesendet. Über ein Bus-Interface VN1630 der Firma Vector Informatik GmbH werden die CAN-Busse aus dem Fahrzeug sowie CAN-Busse aus dem Modell zeitsynchron auf einem Laptop aufgezeichnet. Die entsprechend gewonnenen Daten werden zum einen zur Auswertung und zum anderen zum Training der neuronalen Netze verwendet.

Softwaresetup

Die adaptiven Regelungsverfahren zur Realisierung einer querführenden Assistenzfunktion werden in eine gegebene Umgebung integriert, deren Aufbau vereinfacht in Bild 2b dargestellt ist. Das beschriebene Modell liegt als Simulink-Modell vor. Entsprechend wird das Training des KNNs ebenfalls mithilfe von Matlab realisiert. Im Fahrzeug läuft der Programmcode des Modells, wie bereits im vorherigen Abschnitt erwähnt, auf der MABXII. Nachfolgend soll der vereinfachte Modellablauf in vier Schritten erläutert werden.

Nach dem Einlesen der fahrzeuginternen Signale werden diese für die nachfolgenden Schritte 2 bis 4 vorbereitet, indem beispielsweise Formatkonvertierungen vorgenommen werden. Zusätzlich erfolgen weitere Analysen bezüglich des aktuellen Fahrer- sowie Fahrzeugzustandes und der gesamten Fahrsituation. In Schritt 2, der Bahnplanung, wird basierend auf einer Fusion der Spurdaten aus der Multifunktionskamera und aus Daten anderer Fahrzeugsensoren eine Bahn berechnet, anhand welcher das Fahrzeug durch die nachfolgende Regelung in Schritt 3 geführt werden soll. Abschließend in Schritt 4 werden die Signale nachbereitet. Neben unterschiedlichen Absicherungen erfolgt beispielsweise erneut die Konvertierungen der Signale, bevor diese an das Fahrzeug gesendet werden.

Aktuelle Regelung

Bevor in Abschnitt 4 auf den betrachteten Ansatz zur Verwendung von adaptiven Regelungsverfahren für ein querführendes Fahrerassistenzsystem eingegangen wird, wird nachfolgend die aktuelle Regelung vorgestellt. Der prinzipielle Aufbau ist in Bild 3 zu sehen.

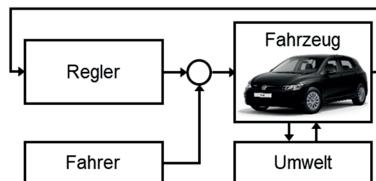
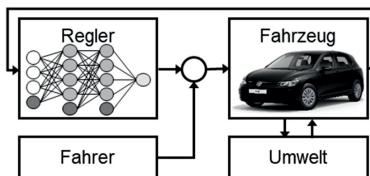


Bild 3: Darstellung eines klassischen Ansatzes zur Realisierung eines FAS

Das Fahrzeug, dargestellt rechts, agiert mit seiner Umwelt. Mithilfe verschiedener fahrzeuginterner Sensorik nimmt das Fahrzeug seine Umgebung wahr. Basierend auf gemessenen und kombinierten Signalen sowie auf weiteren Größen aus der Bahnplanung, wie bereits in Bild 2b gezeigt, ermittelt der Regler eine Stellgröße für die Querführung. Dieses Signal wird mit der Aktion des Fahrers am Lenkrad kombiniert und das Resultat wird entsprechend als Anforderung an die Lenkungsschnittstelle übergeben.

Die Reglerparameter des verwendeten klassischen Regleransatzes werden bislang pro Fahrzeugmodell vor der Fahrzeugauslieferung appliziert. Entsprechend werden weder die vorliegende Kombination aus Motorisierung, Lenkung und Reifenauswahl, noch die individuellen Unterschiede aufgrund von Bauteiltoleranzen berücksichtigt. Um entsprechende Einflüsse zu kompensieren, soll nachfolgend ein möglicher Ansatz vorgestellt werden, mit dem eine adaptive Querführungsregelung realisiert werden kann, die eine fahrzeugspezifische Applikation ermöglicht.



4. Lösungsansatz

Im Rahmen dieses Beitrags soll ein Ansatz zur Realisierung einer Querführungsassistenz durch adaptive Regelungsverfahren betrachtet werden. In diesem Ansatz, dargestellt in Bild 4, wird die gesamte Regelung durch ein künstliches neuronales Netz bzw. durch eine Gruppe von KNN ersetzt.

Bild 4: Darstellung des Ansatzes zur Realisierung einer adaptiven Regelung mit KNN

Die durch unterschiedliche Sensorik ermittelten fahrzeuginternen Signale nutzt das KNN um ein Lenkungskommando zu prädizieren, welches an das Fahrzeug übergeben wird. Der Fahrer ist eine zusätzliche Komponente in diesem System, welche durch Kombination mit der Netzprädiktion für die Fahrzeugquerführung berücksichtigt wird.

Dieser Ansatz bietet den Vorteil, dass die bisherige Regelung als Ausgangslage für die zukünftige fahrzeugindividuelle Querführungsassistenzfunktion imitiert werden kann. Entsprechend kann basierend auf den vorhandenen Trainingsdaten ein Verhalten erlernt werden, welches das Fahrzeug in der Fahrspur hält. Aufgrund der Verwendung eines neuronalen Netzes zur Prädiktion einer Stellgröße an die Lenkungsschnittstelle ist die Absicherung eine herausfordernde Aufgabe. Hier muss eine Strategie entwickelt werden, welche ein sicheres Verhalten des Fahrzeugs auch in Situationen, die dem KNN bislang unbekannt sind, gewährleistet.

Nachfolgend wird das für diesen Lösungsansatz entworfene KNN inklusive der Netzwerkarchitektur und der Trainingsmethodik beschrieben. Im anschließenden Kapitel erfolgt ein Vergleich mit verschiedenen Referenzfahrzeugen.

Die Querregelung durch das KNN wird in das bestehende Modell aus Bild 2b integriert. Als Eingangssignale für das KNN stehen somit neben fahrzeuginternen Messsignalen ebenfalls berechnete modellinterne Signale zur Verfügung.

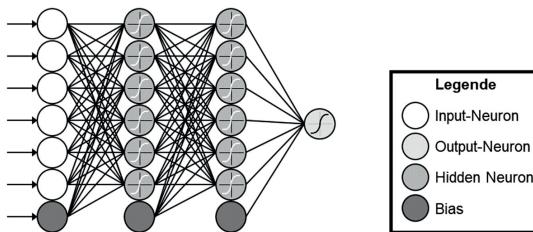


Bild 5: Verwendetes KNN zur Realisierung einer querführenden Fahrerassistenzfunktion angelehnt an

Bei dem verwendeten neuronalen Netz handelt es sich um ein vollverbundenes, vorwärtsgerichtetes KNN mit sechs Eingängen, zwei verdeckten Schichten mit je sechs Neuronen und einem Ausgang, welches in Bild 5 dargestellt ist.

In neuronalen Netzen werden, ausgenommen der Eingabeschicht, in den Neuronen überwiegend nichtlineare Funktionen verwendet, da durch die Kombination dieser beliebig komplexe Funktionen repräsentiert werden können. In allen Neuronen außerhalb der Eingabeschicht wird der Tangens Hyperbolicus (tanh) verwendet, da dieser eine der meist verwendeten Aktivierungsfunktionen mit guter Lernperformance ist. Die Funktion bietet aufgrund der Grenzwerte von ± 1 den Vorteil der Nullzentrierung. Die Eingangs- sowie Ausgangsdaten werden auf diesen Bereich skaliert, um ein optimales Lernen zu ermöglichen.

Als Netzeingänge werden neben der Fahrzeuggeschwindigkeit und -querbeschleunigung vier Größen aus der Bahnplanung verwendet, wobei es sich um den Querabstand zum gewünschten Sollkurs, um die Gierwinkeldifferenz zwischen Fahrzeug- und Spur-markierungsausrichtung, um die Krümmungsdifferenz zwischen dem Sollkurs und den Spurmarkierungen und um eine Größe, welche die Interaktion mit dem Fahrer beschreibt, handelt. Das KNN prädiziert darauf basierend ein Lenkungssignal.

Mithilfe eines Versuchsträgers wurden Messdaten auf vorerst 28 Kilometern bei Geschwindigkeiten zwischen 30 und 100 km/h aufgezeichnet. Bei den vorliegenden Messdaten handelt es sich um 40% auf einer geraden Strecke, 27% Linkskurven und 33% Rechtskurven. Ausgehend von den Systemgrenzen der Fahrerassistenzfunktionen würden die Bereiche zur Skalierung festgelegt und die Daten entsprechend der nachfolgenden Formel für das Training des neuronalen Netzes vorbereitet:

$$x_{\text{skaliert}} = \frac{\text{Zielbereich}_{\text{max}} - \text{Zielbereich}_{\text{min}}}{\text{Systemgrenze}_{\text{max}} - \text{Systemgrenze}_{\text{min}}} \cdot x.$$

Wie zuvor erwähnt werden als Grenzwerte des Zielbereichs aufgrund der tanh-Aktivierungsfunktion ± 1 gewählt.

Das verwendete Netz wurde über 1000 Iterationen mithilfe des Adam-Algorithmus' trainiert. Bei Adam werden die Gewichte anhand des ersten und zweiten Momentes der Gradienten angepasst. Als Schätzungen für diese Momente \hat{m}_t und \hat{v}_t werden der gleitende Mittelwert des einfachen bzw. quadrierten Gradienten verwendet, bei welchen die Parameter β_1 und β_2 den Einfluss vorheriger Zeitschritte regulieren. Für das Training gilt $\beta_1 = 0.9$, $\beta_2 = 0.99$ und Lernrate $\alpha = 0.1$. Die Updates für die Gewichte θ des künstlichen neuronalen Netzes im Zeitschritt t werden mithilfe der folgenden Formel

$$\Delta\theta_t = \alpha \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon}$$

aus berechnet, wobei

$$\hat{m}_t = \frac{m_t}{(1 - \beta_1^t)} = \frac{\beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot \nabla_\theta f_t(\theta_{t-1})}{(1 - \beta_1^t)},$$

$$\hat{v}_t = \frac{v_t}{(1 - \beta_2^t)} = \frac{\beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot (\nabla_\theta f_t(\theta_{t-1}))^2}{(1 - \beta_2^t)},$$

gilt.

Zur Vermeidung von Overfitting und somit zur Verbesserung der Generalisierungsfähigkeit des KNN wird eine L2-Regularisierung ergänzt, bei welcher zur Kostenfunktion die durch den Parameter λ skalierte Summe der quadrierten Gewichte addiert wird. Der entsprechende Regularisierungsfaktor wird auf $\lambda = 0.0001$ gesetzt. Das Training findet in Minibatches von je 256 Datenpaaren statt.

5. Ergebnisse

Im Rahmen erster Fahrversuche soll evaluiert werden, ob eine querführende Fahrerassistenzfunktion durch ein neuronales Netz übernommen werden kann.

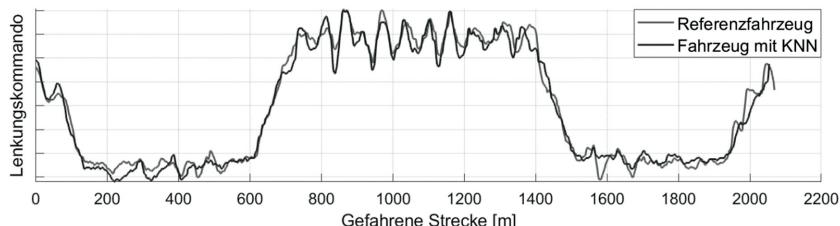


Bild 6: Lenkungskommandoevergleich des Fahrzeugs mit KNN und des Referenzfahrzeugs

Bild 6 stellt vergleichend das prädizierte Lenkungssignal aus dem KNN und das Lenkungssignal des Referenzfahrzeugs gegenüber. Bei der gefahrenen Strecke handelt es sich um einen ca. 2.1 km langen Autobahnabschnitt der A39, welcher aus einer Rechts-Links-Rechts-Kurvenkombination besteht. Aus der Abbildung ist ersichtlich, dass mithilfe des KNN ein ähnliches Lenkungssignal wie vom Referenzfahrzeug prädiziert wird. Durch Hyperparametertuning und Erweiterung der Trainingsdaten sowie durch Variation der Trainingsdauer ist eine weitere Reduzierung der Abweichung denkbar.

Um einen Eindruck bezüglich des tatsächlichen Fahrverhaltens durch das KNN zu erhalten, wird nachfolgend auf die Fahrspur des Fahrzeugs und entsprechend die entstehende Spurmittabenweichung im Vergleich zu verschiedenen Fahrzeugen mit klassischen FAS eingegangen. In untersucht Oschlies unter anderem verschiedene Referenzfahrzeuge in Bezug auf ihre Reglerapplikation im Rahmen eines querführenden Fahrerassistenzsystems. Da es sich hierbei um dieselbe Strecke wie die im Rahmen der hier vorgestellten Fahrversuche handelt, sollen diese Ergebnisse ebenfalls als Referenzen dienen.

Bild 7 zeigt erste Ergebnisse eines Fahrzeugs, dessen Querführung durch ein neuronales Netz erfolgt. Das Fahrzeug befährt die Teststrecke mit einer nahezu konstanten Geschwindigkeit von 100km/h. Visualisiert wird die Spurmittabenweichung. Die dunkle durchgezogene Kurve zeigt hierbei die Fahrt mit dem KNN-FAS, die helle durchgezogene Kurve eine vergleichbare Fahrt mit einem konventionellen FAS, welches kurvenschneidend appliziert ist. Die Minimal- und Maximalwerte der drei Fahrzeuge aus sind zusätzlich als Konstanten in gestrichelten Linien eingezeichnet.

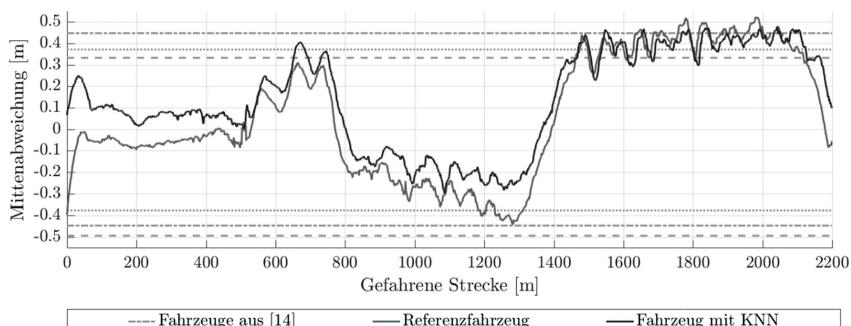


Bild 1: Vergleich der Mittenabweichung des Fahrzeugs mit KNN und verschiedener Referenzfahrzeuge

Aus der Abbildung wird deutlich, dass eine querführende Fahrerassistenzfunktion durch ein KNN eine ähnlich gute Spurführungsgüte wie verschiedene Fahrzeuge mit konventionellen Fahrerassistenzsystemen erreichen kann.

Die stärkere Repräsentation von Rechtskurven in den Trainingsdaten kann als möglicher Grund für die ähnlicheren Spurmittabenweichungen des Fahrzeugs mit KNN und dem Referenzfahrzeug genannt werden. Vor allem in den Linkskurven wählt das Fahrzeug mit KNN eine Fahrspur, welche näher an der Spurmitte liegt als die des Referenzfahrzeugs. Hieraus kann jedoch nicht gefolgert werden, dass das vorgestellte trainierte Netz auf allen Strecken eine geringere Mittenabweichung als das Referenzfahrzeug erzielt.

6. Zusammenfassung und Ausblick

Für die Realisierung eines optimalen Kundenerlebnisses durch eine Fahrerassistenzfunktion ist es notwendig, eine Applikation vorzunehmen, die sich optimalerweise auch an Veränderungen über die Fahrzeuglebenszeit anpasst. Basierend auf einer Literaturrecherche wurde ein Ansatz zur adaptiven Regelung mit maschinellem Lernen vorgestellt, um eine fahrzeugspezifische, querführende Fahrerassistenzfunktion zu realisieren. Bei ersten Fahrversuchen zeigte sich, dass mithilfe eines neuronalen Netzes das Lenkungskommando für ein Fahrzeug prädi-ziert und infolgedessen ein Fahrzeugverhalten erlernt werden kann, welches eine vergleichbar gute Spurführungsgüte wie verschiedene Referenzfahrzeuge erreicht. Basierend auf diesem initialen antrainierten Fahrzeugverhalten soll zukünftig das Lernen im Fahrzeug und entspre- chend eine Anpassung der Fahrerassistenzfunktion realisiert werden, wozu Methoden des Reinforcement Learnings Anwendung finden. In Bezug auf die Serienentwicklung einer Fahrerassistenzfunktion ist weiterhin zu betrachten, wie eine Absicherung des KNNs möglich ist. Neben dem vorgestellten Ansatz kann eine adaptive Regelung auch durch die Kombination einer klassischen Reglerstruktur in Kombination von KNN zur Prädiktion der zugehörigen Pa-rameter realisiert werden. Dieser Ansatz soll in Folgearbeiten betrachtet werden, da er den Vorteil bietet, dass die Parameter einfacher abgesichert werden können.

Literatur

- [1] D. Schramm und S. Schweig, „Fahrerassistenzsysteme - ein Überblick,“ in *Altersgerechte Fahrerassistenzsysteme: Technische, psychologische und betriebswirtschaftliche Aspekte*, Wiesbaden, Springer Fachmedien Wiesbaden, 2020, p. 39–53.
- [2] K. International, *2020 Autonomous Vehicles Readiness Index*.
- [3] D. Verkehrssicherheitsrat, „Autofahrer in Deutschland setzen auf Sicherheit,“ [Online]. Available: <https://bester-beifahrer.de/pressemitteilungen/autofahrer-in-deutschland-setzen-auf-sicherheit/>.
- [4] H. Winner, S. Hakuli, F. Lotz und C. Singer, Hrsg., *Handbuch Fahrerassistenzsysteme: Grundlagen, Komponenten und Systeme für aktive Sicherheit und Komfort*, 3., überarb. und erg. Aufl. Hrsg., Wiesbaden: Springer Vieweg, 2015.
- [5] K. Reif, „Software,“ in *Bosch Autoelektrik und Autoelektronik: Bordnetze, Sensoren und elektronische Systeme*, Wiesbaden, Vieweg+Teubner, 2011, p. 206–229.
- [6] H. Unbehauen, *Regelungstechnik III: Identifikation, Adaption, Optimierung*, 6., verbesserte Auflage Hrsg., Wiesbaden: Vieweg+Teubner Verlag, 2000.
- [7] N. Ahmadian, A. Khosravi und P. Sarhadi, „Integrated model reference adaptive control to coordinate active front steering and direct yaw moment control,“ *ISA transactions*, Bd. 106, p. 85–96, 2020.
- [8] Haykin, *Neural networks and learning machines: A comprehensive foundation*, 3. Hrsg., New York: Pearson Education, 2009.
- [9] D. A. Pomerleau, „ALVINN: An Autonomous Land Vehicle in a Neural Network,“ *Proc. Adv. Neural Inf. Process. Syst.*, p. 305–313, 1989.
- [10] D. Pomerleau, „Neural Network Vision for Robotic Driving,“ in *Intelligent Unmanned Ground Vehicles*, Bd. 388, Boston, MA, Springer US, 1997, p. 53–72.
- [11] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang, X. Zhang, J. Zhao und K. Zieba, *End to End Learning for Self-Driving Cars*.
- [12] S. Hecker, D. Dai und L. van Gool, *End-to-End Learning of Driving Models with Surround-View Cameras and Route Planners*.
- [13] D. P. Kingma und J. L. Ba, „Adam: A method for stochastic optimization,“ *arXiv preprint arXiv:1412.6980v9*, 2017.
- [14] H. Oschlies, *Komfortorientierte Regelung für die automatisierte Fahrzeugquerföhrung*, Bd. 136, Wiesbaden: Springer Fachmedien Wiesbaden, 2019.

Augmentation von Kameradaten mit Generative Adversarial Networks (GANs) zur Absicherung automatisierter Fahrfunktionen

Augmentation of camera data via Generative Adversarial Networks (GANs) for the validation of automated driving functions

Philipp Rigoll, Patrick Petersen, Lennart Ries, Jacob Langner,
Eric Sax, FZI Forschungszentrum Informatik, Karlsruhe

Zusammenfassung

Für die Entwicklung und Absicherung automatisierter Fahrfunktionen sind Datensätze notwendig, die den in der Realität vorkommenden Variantenreichtum von Verkehrs- und Umweltsituationen abzudecken. Beim manuellen Einfahren dieser Daten existieren jedoch Umweltparameter, wie beispielsweise das Wetter, die sich nicht oder nur schwer beeinflussen lassen. Im Folgenden werden ausschließlich Bild-Datensätze betrachtet. Ein Beispiel ist dabei ein Datenpunkt, also ein Bild, im Datensatz. Umweltparameter sind zum Beispiel neben dem Wetter, u. a. andere Verkehrsteilnehmer und Verkehrssituationen. Um dennoch eine effiziente Abdeckung dieses Parameterraums für Kamerabilder zu erreichen, müssen Alternativen zum Einfahren der Daten im Erprobungsfahrten in Betracht gezogen werden. Die Simulation ist eine Möglichkeit, die Abdeckung zu vergrößern. Bei der Simulation kann man alle Umweltparameter genau vorgeben und erreicht dadurch die gewünschte Variabilität. Jedoch ist das auch mit einem hohen manuellen Aufwand für die Modellierung verbunden. Außerdem beruhen die generierten Daten der Simulation auf Sensormodellen. Diese Modelle sind jedoch nur ein näherungsweises Abbild der Realität. Ein anderer Ansatz ist die Augmentation. Im speziellen wird hier die Augmentation mittels maschinellen Lernverfahren betrachtet. Diese nutzt die bei Erprobungen aufgezeichneten Daten und erzeugt durch Veränderung der aufgezeichneten Daten neue Beispiele. Dabei erfolgt die Veränderung durch maschinelle Lernverfahren, welche den realen Sensoreindruck aus realen Beispielen lernen. Hierdurch ist sichergestellt, dass die neuen Beispiele möglichst ähnlich zum ursprünglichen Datensatz sind. Das Ziel ist eine automatisierte Vergrößerung eines Datensatzes und eine damit verbundene vergrößerte Abdeckung des dadurch entstehenden neuen Datensatzes.

Abstract

The development and validation of automated driving functions require large datasets to cover variation in traffic and environmental conditions. However, some environmental parameters, such as specific weather and traffic conditions, as well as road users, are difficult to record manually. To achieve efficient coverage of this parameter space of camera images, alternatives to actual driving data should be considered. Simulation is a solution to achieve an increase in coverage. With simulation, all environmental parameters can be specified and thus the desired variability can be achieved. However, this is also associated with a high level of manual modeling effort. In addition, the data generated by the simulation is based on sensor models. However, these models are only a rough representation of reality. Another approach is augmentation. In particular, increasing data variance via machine learning algorithms is considered in this paper. These algorithms use real driving data and generate new samples by changing the recorded data. Modification is carried out using machine learning algorithms that learn the sensor impression from actual data. This ensures that the new samples are as similar as possible to the original dataset. The goal is automated data replication and the associated expanded coverage of the resulting new dataset.

1. Umfeld des Konzepts

Die Daten-Augmentation ist eine Methode aus dem Training Neuronaler Netze. Man versteht darunter die Vergrößerung eines Trainingsdatensatzes durch die Veränderung einzelner Beispiele. Die Veränderungen sind ursprünglich klassische Bildmanipulationen, wie Rotationen, Spiegelungen, Verzerrungen, Weichzeichnungen und ähnliches von Trainingsbildern. Die dadurch entstehenden vergrößerten Datensätze verbessern die Generalisierung der damit trainierten Neuronalen Netzen [1]. Über die klassischen Bildmanipulationen hinausgehende Augmentationen können mit Hilfe von maschinellen Lernverfahren realisiert werden. Hier wurde bereits gezeigt, dass sich die Performance von Neuronalen Netzen durch den Einsatz von Augmentation verbessert [2,3]. Außerdem wurde mit Augmentation bereits Overfitting, also das Auswendiglernen von Trainingsdatensätzen, verhindert [2,4,5]. Auch bei Datensätzen, deren Anzahl an Klassenbeispielen nicht ausgeglichen ist, konnte mit Augmentation mittels maschinellen Lernverfahren eine Verbesserung erreicht werden [2,6].

2. Konzept

Die Methode der Augmentation mittels maschinellen Lernverfahren soll nun dahingehend angepasst werden, dass es zur Vergrößerung der Abdeckung bei der Entwicklung und beim Testen von automatisierten Fahrfunktionen angewendet werden kann. Hierfür wird ein Augmentationsprozess vorgestellt, der sich in die Entwicklung von automatisierten Fahrfunktionen einbetten lässt.

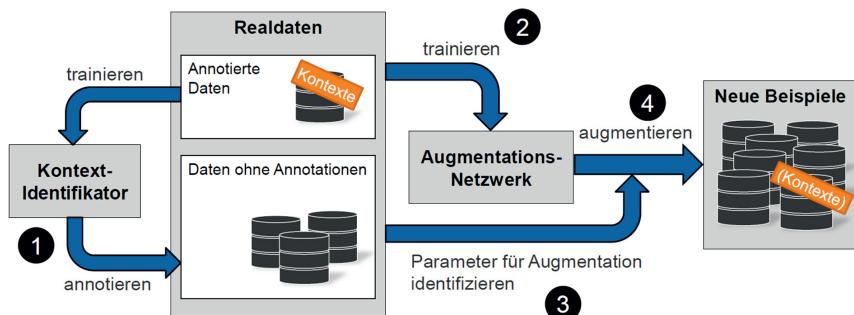


Bild 1: Schema des Augmentationsprozesses mit maschinellen Lernverfahren in der Entwicklung hochautomatisierter Fahrfunktionen

Der Augmentationsprozess besteht aus vier einzelnen Schritten (siehe Abb. 1). Ausgegangen wird dabei von einem Bild-Datensatz, der teilweise annotiert ist. Annotation basiert auf manuellen und damit teuren Aufwänden, weshalb diese oft nicht für den ganzen Datensatz vorhanden sind. So sind beispielsweise beim A2D2-Datensatz [7] weniger als 10 % der Bilder semantisch segmentiert, beim BDD100K Datensatz [8] haben auch nur 10 % der Videos eine Semantische Segmentierung. Daher wird für die bei der Augmentation notwendigen Annotationsen ein Kontext-Identifikator trainiert. Mit diesem ist man anschließend in der Lage, auch für den restlichen Datensatz Annotationsen zu erzeugen.

Ausgehend von dem vollständig annotierten Datensatz wird im zweiten Schritt ein Augmentationsnetzwerk trainiert. Dieses Netzwerk führt die eigentliche Augmentation durch, indem Beispiele aus dem ursprünglichen Datensatz verändert werden oder indem vollständig neue Beispiele erzeugt werden. Je nach Augmentationsmethode ist für die Augmentation weiteres Wissen über zusätzliche Parameter notwendig. Diese Parameter werden im dritten Schritt aus den Daten und Annotationen identifiziert.

Ist das Augmentationsnetz trainiert und sind die entsprechenden zusätzlichen Parameter identifiziert, so wird im letzten Schritt die eigentliche Augmentation durchgeführt. Der ursprüngliche Datensatz und dessen Abdeckung von Verkehrsszenarien wird hierdurch vergrößert. Abhängig von der genutzten Augmentationsmethode sind die neuen Beispiele bereits annotiert oder müssen in einer weiteren Schleife mit dem Kontext-Identifikator annotiert werden. Ein Vorteil dieses Vorgehens liegt darin, dass der Prozess nach der initialen Annotation voll automatisch ablaufen kann.

3. Augmentationsnetzwerke

Zur Generierung neuer Bilder haben sich in den letzten Jahren Generative Adversarial Networks (GANs) bewährt [9,10,11]. Die ursprüngliche GAN-Architektur [11] ist in der Lage Beispiele zu generieren, die so aussehen, als würden sie aus der gleichen Grundverteilung wie die entsprechenden Trainingsbilder stammen. Sie besteht dabei aus zwei Teilnetzen (siehe Abb. 2), die gegeneinander ein Nullsummenspiel betreiben.

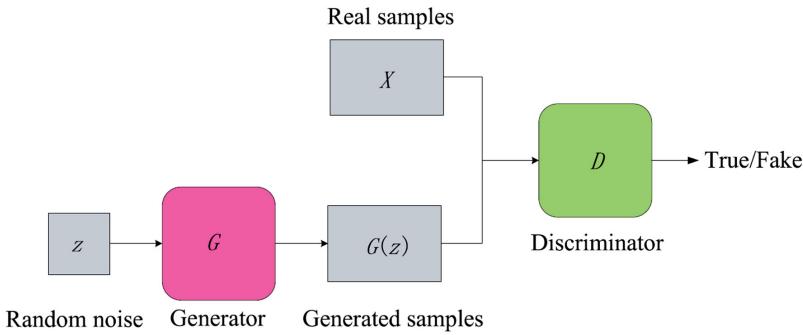


Bild 2: Schema eines Standard Generative Adversarial Networks [9]

Der erste Teil der GAN-Architektur wird Generator genannt. Dieser erzeugt aus einem zufällig generierten Vektor ein neues Beispielbild. Die Aufgabe des zweiten Teilnetzes, dem Diskriminatator, ist es während des Trainings zu entscheiden, ob ein Bild tatsächlich aus dem Trainingsdatensatz stammt oder es ein vom Generator erzeugtes Bild ist. Die generierten Bilder nähern sich immer besser an die Trainingsdaten an, da die Teilnetze dahin trainiert werden, dass sie sich gegenseitig ausspielen. Nach dem Training ist dann nur noch der Generator notwendig, um neue Beispiele zu erzeugen. Neben der ursprünglichen GAN-Architektur existieren

vielfältige Verbesserungen und Abwandlungen [9,10]. Aus diesen lassen sich zahlreiche Augmentationsmethoden konzeptionieren und im vorgestellten Augmentationsprozess umsetzen.

4. Prototypische Umsetzung

Zur Verdeutlichung des vorgestellten Augmentationsprozesses wird das Vorgehen an einem Beispiel dargestellt. Ziel der beispielhaften Augmentationsmethode soll die Wetteränderung der Verkehrsszenarios sein. Dabei soll die Verkehrssituation, also die Position und die vorhandenen Verkehrsteilnehmer, unverändert bleiben.

Kontext-Identifikator: Netz zur Wettererkennung

Für die spätere Augmentation müssen im Ursprungsdatensatz Annotationen für jede betrachtete Wetterklasse vorhanden sein. Mit diesen bereits annotierten Bildern kann dann ein Wetter-Klassifikationsnetzwerk trainiert werden, um auch die nicht annotierten Bilder mit der entsprechenden Wetterklasse zu annotieren. Ein mögliches Klassifikationsnetz für diese Aufgabe ist WeatherNet [12].

Augmentationsnetzwerk trainieren: Bild-zu-Bild

Die eigentliche Augmentation wird durch ein Neuronales Netz erreicht, das ein Bild als Eingabe bekommt und dieses verändert. Hier gibt es eine Vielzahl an GAN-Variationen, die diese Problemstellung angehen. Für zwei Wetterklassen könnte beispielsweise ein CycleGAN [13] verwendet werden. Für jede Kombination von zwei Wetterklassen muss hier ein eigenes Netz trainiert werden. Eine Alternative, die auch mehrere Bild-Klassen unterstützt, ist StarGAN v2 [14]. Nach der Wahl eines geeigneten Augmentationsnetzes muss dieses mit dem vollständig annotierten Ursprungsdatensatz trainiert werden.

Zusätzliche Parameter für die Augmentation

Für die Augmentation ist die Information über die Wetterklasse des Eingangsbildes notwendig, um das entsprechende Augmentationsnetz zu wählen bzw. entsprechend zu parametrisieren. Dabei bieten sich auch Analysen über die Verteilung der Wetterklassen im Ursprungsdatensatz an. So können Lücken im Datensatz, wie zum Beispiel unterrepräsentierte Wetterklassen, identifiziert und gezielt durch Augmentation geschlossen werden.

Augmentation durchführen

Bei der Durchführung der Augmentation werden Bilder aus dem Ursprungsdatensatz zusammen mit der Information über die jeweilige Klasse durch das Augmentationsnetz geleitet. Beispielhaft wurde der vorgestellte Augmentationsprozess mit dem Oxford-Datensatz [13] durchgeführt. Dieser Datensatz ist mit Wetterklassen annotiert. Für die Durchführung wurden die Wetterklassen Sonne, Schnee und Regen ausgewählt. Als Augmentationsnetzwerk wurde ein StarGAN v2 [14] trainiert. Exemplarische Augmentationen sind in Abbildung 3 zu sehen. Der Inferenzschritt des Prozesses, also nach dem Training, ist an einem Beispiel in Abbildung 4 dargestellt.

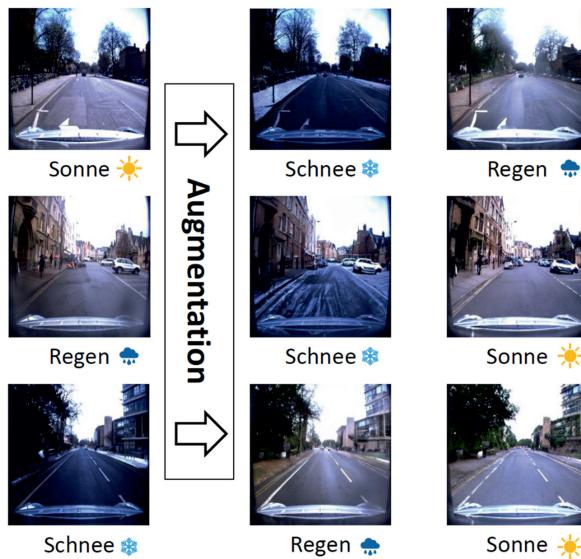


Bild 3: Exemplarische Augmentationen des Oxford-Datensatz [15] mit einem StyleGAN v2 [14]

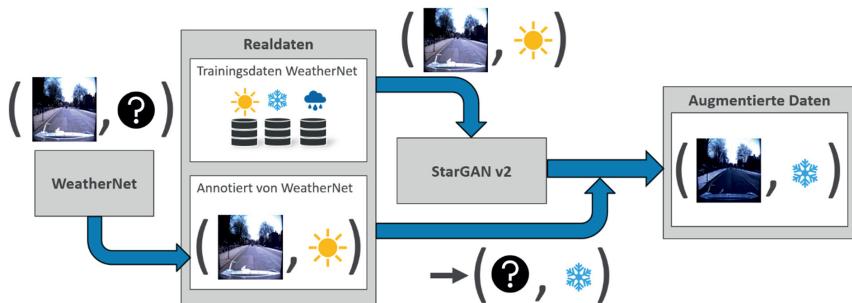


Bild 4: Prozess nach dem Training der Netze an einem Beispiel aus dem Oxford-Datensatz [15]

5. Zusammenfassung und Ausblick

Wir haben einen Prozess vorgestellt, der Daten-Augmentation mit maschinellen Lernverfahren für die Absicherung automatisierbarer Fahrfunktionen nutzbar macht. Mit diesem Augmentationsprozess ist es möglich, Bild-Datensätze automatisiert zu vergrößern. Der Prozess wurde beispielhaft mit der Augmentation von Wetter durchgeführt.

In Zukunft soll der Prozess um eine Bewertung der Augmentationen erweitert werden. Außerdem soll für die vorhandenen Augmentationen eine Systematik erarbeitet werden.

Danksagung

Die Arbeiten an dieser Veröffentlichung wurden durch das Bundesministerium für Wirtschaft und Energie innerhalb des Forschungsprojekts „KI Data Tooling – Methoden und Werkzeuge für das Generieren und Veredeln von Trainings-, Validierungs- und Absicherungsdaten für KI-Funktionen autonomer Fahrzeuge“ gefördert. Die Autoren bedanken sich bei den Projektpartnern für die erfolgreiche Zusammenarbeit.

Wir möchten uns bei Artur Leonhardt für seine Unterstützung bedanken.

Literatur

- [1] Goodfellow, I. (2016): Deep Learning.
- [2] Shorten, C. und Khoshgoftaar, T. M. (2019): A survey on Image Data Augmentation for Deep Learning, *J. Big Data*, Bd. 6, Nr. 1.
- [3] Wong, S. C., Gatt, A., Stamatescu, V. und McDonnell, M. D. (2016): Understanding Data Augmentation for Classification: When to Warp?, 2016 Int. Conf. Digit. Image Comput. Tech. Appl. DICTA 2016.
- [4] Cireşan, D. C., Meier, U., Gambardella, L. M. und Schmidhuber, J. (2010): Deep, big, simple neural nets for handwritten digit recognition, *Neural Comput.*, Bd. 22, Nr. 12, pp. 3207–3220.
- [5] Chawla, N. V., Bowyer, K. W., Hall, L. O. und Kegelmeyer, W. P. (2002): SMOTE: synthetic minority over-sampling technique, *J. Artif. Intell. Res.*, Bd. 16, pp. 321–357.
- [6] Antoniou, A., Storkey, A. und Edwards, H. (2018): Data Augmentation Generative Adversarial Networks, pp. 1–14.
- [7] Geyer, J. et. al. (2020): A2D2: Audi Autonomous Driving Dataset, arXiv:2004.06320 [cs, eess].
- [8] Yu, F. et. al. (2020): BDD100K: A Diverse Driving Dataset for Heterogeneous Multitask Learning, arXiv:1805.04687 [cs].
- [9] Wang, L., Chen, W., Yang, W., Bi, F. und Yu, F. R. (2020): A State-of-the-Art Review on Image Synthesis With Generative Adversarial Networks, *IEEE Access*, Bd. 8, S. 63514–63537, 2020.
- [10] Shorten, C. und Khoshgoftaar, T. M. (2019): A survey on Image Data Augmentation for Deep Learning“, *J Big Data*, Bd. 6, Nr. 1, S. 60.
- [11] Goodfellow, I. et. al. (2014): Generative Adversarial Nets.
- [12] Ibrahim, M., Haworth, J., und Cheng, T. (2019): WeatherNet: Recognising Weather and Visual Conditions from Street-Level Images Using Deep Residual Learning, *IJGI*, Bd. 8, Nr. 12, S. 549.
- [13] Zhu, J.-Y., Park, T., Isola, P. und Efros, A. A. (2017): Unpaired Image-to-Image Translation using Cycle-Consistent Adversarial Networks, arXiv:1703.10593 [cs].
- [14] Choi, Y., Uh, Y., Yoo, J. und Ha, J.-W. (2020): StarGAN v2: Diverse Image Synthesis for Multiple Domains, arXiv:1912.01865 [cs].
- [15] Maddern, W., Pascoe, G., Linegar, C. und Newman, P. (2017): 1 year, 1000 km: The Oxford RobotCar dataset, *The International Journal of Robotics Research*, Bd. 36, Nr. 1, S. 3–15.

Kalibrierung von Neuronalen Netzen für Detektionsmodelle

Fabian Küppers, Jan Kronenberger, Anselm Haselhoff,
Hochschule Ruhr West, Bottrop;
Jonas Schneider,
Elektronische Fahrwerksysteme (EFS), Gaimersheim

Zusammenfassung

Moderne Detektionssysteme auf Basis von Neuronalen Netzen bieten zahlreiche Einsatzmöglichkeiten, bspw. für automatisierte Fahrfunktionen und Fahrerassistenzsysteme. Insbesondere für sicherheitskritische Anwendungen ist aber eine statistisch belastbare Aussage über die Unsicherheit eines Systems von entscheidender Bedeutung. In der Vergangenheit hat sich allerdings gezeigt, dass insbesondere Detektionsmodelle auf Basis Neuronaler Netze eine unzuverlässige Selbsteinschätzung der Unsicherheiten aufweisen, die in vielen Fällen nicht mit der tatsächlich beobachteten Genauigkeit des Modells übereinstimmt. Im Rahmen unserer Forschungsarbeit haben wir Methoden zur Kalibrierung der Unsicherheiten auf die Anwendbarkeit im Bereich der Objektdetektion hin untersucht und erweitert. Mit unseren Methoden ist es möglich, die Fehlkalibrierung von Detektions- und Segmentierungsmodellen **positionsabhängig** festzustellen und zu korrigieren. Somit können diese Methoden zu verlässlicheren Informationen für Fahrerassistenzsystemen beitragen.

Abstract

Neural networks can be used for various tasks, e.g., in the field of automated driving functions and driver assistance systems. However, especially for safety-critical applications, a statistically reliable self-assessment about the uncertainty of a system is of crucial importance. In the past, it has been shown that detection models based on neural networks exhibit an unreliable self-assessment of uncertainties in their predictions. This uncertainty does not correspond to the observed accuracy of the model in many cases. As part of our research, we investigated and extended methods for uncertainty calibration in the field of object detection and segmentation. Using our methods, it is possible to detect and correct the miscalibration of detection and segmentation models in a **position-dependent** manner. Thus, these methods can contribute to more reliable information for automated driving functions and driver assistance systems.

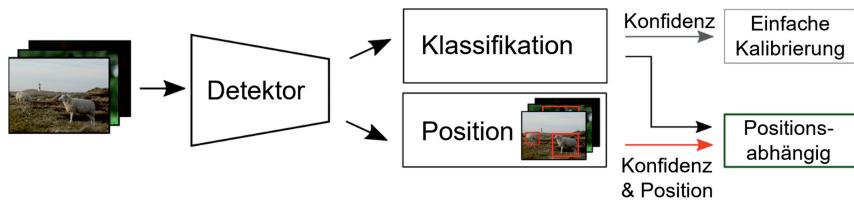


Bild 1: Im Gegensatz zu klassischen neuronalen Netzen schätzt ein Detektionsmodell neben der Klassenzugehörigkeit auch die Position von Objekten. Unser Konzept erweitert Verfahren zur Konfidenz-Kalibrierung, sodass diese Positionsinformationen bei der Kalibrierung mitberücksichtigen werden können (unten), während herkömmliche Verfahren nur auf Basis der Konfidenz kalibrieren (oben) (vgl. [2], Abb. 2).

1. Einleitung

Neuronale Netze können im Bereich automatisierter Fahrfunktionen für verschiedene Zwecke eingesetzt werden, so z.B. zur Umfeldwahrnehmung oder Routenplanung. Bei jeder Entscheidung eines Netzwerks wird neben der eigentlichen Vorhersage ein Score angegeben, der die Selbsteinschätzung des Netzwerks über die entsprechende Vorhersage repräsentiert und als Wahrscheinlichkeit der Korrektheit interpretiert werden kann. Im Idealfall stimmt diese Wahrscheinlichkeit mit der zu beobachtenden Genauigkeit des Modells überein. Ein Beispiel: Ein Fahrerassistenzsystem verwendet zur Erkennung von Personen ein neuronales Netz, welches im Laufe einer Fahrt 100 Objekte detektiert hat, wobei zu jeder dieser Detektionen eine Wahrscheinlichkeit der Korrektheit von 80% prädiziert wurde. Idealerweise kann nun eine Genauigkeit von ebenfalls 80% beobachtet werden.

Jüngste Forschungsergebnisse haben jedoch gezeigt, dass diese Selbsteinschätzungen in vielen Fällen nicht mit der tatsächlich beobachteten Genauigkeit des Modells übereinstimmen [1, 2, 3]. In der Regel wird die Wahrscheinlichkeit der Korrektheit zu hoch eingeschätzt. Diese fehlkalibrierten Modelle stellen jedoch ein potenzielles Sicherheitsrisiko dar, da eine zu selbstsichere Einschätzung den weiteren Verarbeitungs- und Entscheidungsprozess des Systems beeinträchtigen und verzerrn kann.

In der Vergangenheit haben sich zahlreiche Studien mit der Feststellung und der Behebung einer Fehlkalibrierung von neuronalen Netzen im Kontext der Bildklassifikation beschäftigt [1, 4, 5]. Im Rahmen unserer Forschungsarbeit haben wir diese Methoden auf die Anwendbarkeit im Bereich der Objektdetektion hin untersucht und erweitert [2]. Mit unseren Methoden

ist es möglich, die Fehlkalibrierung von Detektionsmodellen positionsabhängig festzustellen und zu korrigieren. Zu diesem Zweck verwenden wir für unsere Kalibrierungsalgorithmen neben der prädizierten Konfidenz die Positionsschätzung des Neuronalen Netzes. Dies ermöglicht eine auf Detektionsmodelle angepasste und flexiblere Messung der Fehlkalibrierung sowie Durchführung der Kalibrierung selbst [2]. Das Verfahren ist in Abbildung 1 schematisch dargestellt. Weiterhin haben wir die Anwendbarkeit unserer Methoden auf Segmentierungsmodelle (Instanzen & semantisch) erweitert, sodass die Fehlkalibrierung in Abhängigkeit der Konfidenzwerte für jeden Pixel einer Segmentierungsmaske festgestellt und behoben werden kann [6]. Dies ermöglicht realistische und insbesondere in den Randbereichen eines Segments feiner aufgelöste Abstufungen der Unsicherheiten. Die von uns entwickelten Methoden zur Feststellung und Behebung einer Fehlkalibrierung von Detektions- oder Segmentierungsmodellen können somit zu verlässlicheren Informationen für automatisierte Fahrerfunktionen und Fahrerassistenzsystemen beitragen [2, 6] und werden nachfolgend im Detail vorgestellt.

2. Konfidenz-Kalibrierung: Definition

Ein Neuronales Netz im Bereich der Klassifikation schätzt zu einer Eingabe $X \in \mathcal{X}$ eine Klasse $\hat{Y} \in \mathcal{Y} = \{1, \dots, K\}$ mit K verschiedenen Klassen. Zu jeder Prädiktion wird zudem ein Score $\hat{P} \in [0,1]$ bestimmt, der als Konfidenz bzw. als die geschätzte Wahrscheinlichkeit der Korrektheit der Prädiktion aufgefasst werden kann. Sei weiterhin $Y \in \mathcal{Y}$ die bekannten Annotations zu X , so gilt ein Klassifikationsmodell als **perfekt kalibriert**, wenn

$$Pr(\hat{Y} = Y | \hat{P} = p) = p, \quad \forall p \in [0,1]$$

erfüllt ist [1]. Anders ausgedrückt bedeutet dies, dass die Wahrscheinlichkeit für die Korrektheit der Prädiktion gegeben einer Konfidenz p gleich dieser Konfidenz sein muss. Dies muss für alle möglichen Konfidenzwerte erfüllt sein. Ein Beispiel: Ein Neuronales Netz klassifiziert 100 Bilder mit einer Konfidenz von jeweils 80%. Der Definition für die Kalibrierung folgend wird nun erwartet, dass 80 von diesen 100 Bildern richtig klassifiziert wurden. Diese Aussage muss für alle möglichen Konfidenzwerte p gelten.

Im Gegensatz zu einem Klassifikationsmodell kann ein Detektionsmodell für eine Eingabe X mehrere Objekte \hat{Y} mit entsprechender Konfidenz \hat{P} im Bild bestimmen. Des Weiteren werden Position sowie Größe der erkannten Objekte prädiziert, welche nachfolgend mit $\hat{R} \in \mathcal{R}^J$ bezeichnet werden, wobei J die Größe der Box-Kodierung angibt (Position, Breite, Höhe, etc.). Bei der Kalibrierung im Bereich der Klassifikation wird die geschätzte Konfidenz

mit der beobachteten Genauigkeit (Accuracy) des Modells verglichen. Im Kontext der Objektdetektion kann die Accuracy jedoch häufig nicht bestimmt werden, da für die Berechnung der Accuracy die Informationen über so genannte *true negatives* fehlen (Bildregionen, die korrekt als Hintergrund identifiziert wurden). Daher wird für Detektionsmodelle die Precision als Maß zur Kalibrierung verwendet (*wie viele der präzisierten Objekte wurden richtig erkannt?*). Demnach ist ein Detektionsmodell **perfekt kalibriert**, wenn

$$Pr(M = 1 | \hat{P} = p, \hat{Y} = y, \hat{R} = r) = p, \quad \forall p \in [0,1], y \in \mathcal{Y}, r \in \mathcal{R}^J$$

erfüllt ist [2]. Der Ausdruck $Pr(M = 1)$ bezeichnet hier die Wahrscheinlichkeit, dass ein präzisiertes Objekt ein annotiertes Objekt derselben Klasse mit einer bestimmten *intersection over union* (IoU) überschneidet (Überlappung zwischen präzisiertem und tatsächlichem Objekt).

Diese Definition lässt sich darüber hinaus auch auf Segmentierungsmodelle übertragen [6]. Bei der Instanzsegmentierung wird zusätzlich zu der präzisierten Position und Größe eines Objekts dessen Kontur bestimmt. Dabei wird jedem Pixel innerhalb der Prädiktion eine Konfidenz zugeordnet, welche als Vertrauensmaß interpretiert wird, ob dieses Pixel Teil der Segmentierungsmaske ist. Im Fall der Instanzsegmentierung kann für jede Maske die Genauigkeit bestimmt werden, da für jedes Pixel eines Objekts eine Konfidenz bestimmt wird und somit Informationen auch für die *true negatives* zur Verfügung stehen. Ziel der Kalibrierung ist in diesem Fall, dass die Konfidenz für die einzelnen Pixel innerhalb der Segmentierungsmasken der tatsächlich beobachteten Genauigkeit entspricht [6].

Im Gegensatz zur Instanzsegmentierung werden bei der semantischen Segmentierung keine einzelnen Objekte präzisiert, sondern die Klassenzugehörigkeiten ganzer Segmente. Somit wird für jedes Pixel im Bild eine Konfidenz bestimmt, die die Schätzung des Segmentierungsmodells über die Klassenzugehörigkeit des Pixels widerspiegelt. Ähnlich zur Instanzsegmentierung kann auch bei der semantischen Segmentierung die Genauigkeit der Klassifikationen über alle Pixel bestimmt werden. Die von dem Segmentierungsmodell geschätzte Konfidenz soll somit der beobachteten Genauigkeit entsprechen [6].

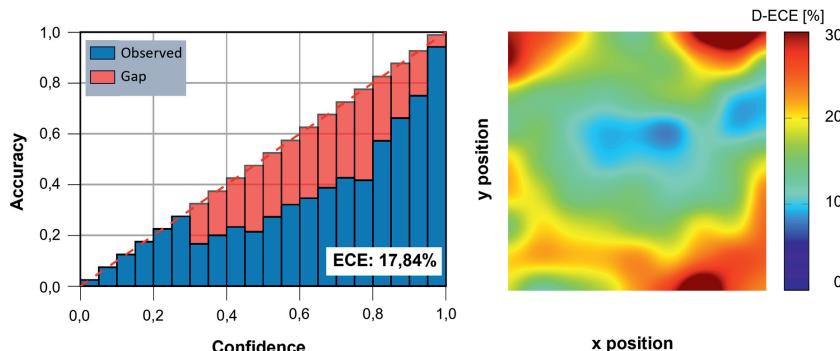


Bild 2: Reliability Diagramm für ein Faster R-CNN Detektionsmodell auf MS COCO (links).

Die Konfidenz ist in 20 gleich große Bins aufgeteilt. In jedem Bin wird die durchschnittliche Accuracy berechnet, um den Kalibrierungsfehler in Abhängigkeit der Konfidenz zu bestimmen. Der ECE ist der gewichtete Mittelwert aus den Abweichungen aller Bins. Auf der rechten Seite ist ein Reliability Diagramm für ein Mask R-CNN Detektionsmodell (ohne Segmentierungsmasken) auf MS COCO Daten dargestellt, welches die Fehlkalibrierung des Detektors in Abhängigkeit von der Objektposition zeigt (vgl. [6]).

3. Messung der Fehlkalibrierung

Weicht ein Klassifikations-, Detektions- oder Segmentierungsmodell von den o.g. Definitionen ab, wird es als fehlkalibriert bezeichnet. Für Klassifikationsprobleme kann die Fehlkalibrierung mittels des *Expected Calibration Errors* (ECE) gemessen werden [1, 5]. Das Maß für die Fehlkalibrierung ist definiert als

$$\text{ECE} = \mathbb{E}_{\hat{p}} [|Pr(\hat{Y} = Y | \hat{p} = p) - p|], \quad \forall p \in [0,1].$$

Da es sich bei \hat{p} um eine kontinuierliche Zufallsvariable handelt, benötigen wir eine Approximation des ECE. Konkret wird der ECE daher über ein Binning-Verfahren angenähert, wobei der Wahrscheinlichkeitsraum $[0,1]$ in B gleich große Intervalle (oder Bins) eingeteilt wird. In Abhängigkeit von der Konfidenz wird jedes Sample in ein Bin eingruppiert, sodass der ECE mittels

$$\text{ECE} \approx \sum_{m=1}^B \frac{1}{N_m} |\text{acc}(m) - \text{conf}(m)|$$

angenähert wird, wobei $\text{acc}(m)$ die durchschnittliche Accuracy, $\text{conf}(m)$ die durchschnittliche Konfidenz und N_m die Anzahl an Samples in einem Bin m darstellen [1, 5]. Die Berechnung des ECEs ist schematisch in Abbildung 2 dargestellt.

Analog dazu erfolgt die Berechnung des Kalibrierungsfehlers für Detektionsmodelle über ein multidimensionales Binning über den Wahrscheinlichkeitsraum $[0,1]$ sowie den möglichen Objektpositionen und -größen über \mathcal{R}^J . Diese Metrik wird als *Detection Expected Calibration Error* (D-ECE) bezeichnet und ist definiert als

$$\text{ECE} = \mathbb{E}_{\hat{P}, \hat{Y}, \hat{R}} [|Pr(M = 1 | \hat{P} = p, \hat{Y} = y, \hat{R} = r) - p|]$$

und wird über das multidimensionale Binning mit

$$\text{D-ECE} \approx \sum_{m=1}^B \frac{1}{N_m} |\text{prec}(m) - \text{conf}(m)|$$

angenähert [2]. Über dieses Binning kann der Kalibrierungsfehler in Abhängigkeit von der Position ermittelt werden.

Der D-ECE kann auch für Segmentierungsmodelle eingesetzt werden. Hier läuft das Binning über beliebige Merkmale der einzelnen Pixel (bspw. relative Position im Bild, Distanz zur nächsten Segmentgrenze). Auch hierbei werden alle prädizierten Konfidenzen der einzelnen Pixel nach einem Binning-Schema in Intervalle eingruppiert. Innerhalb dieser Bins kann für Segmentierungsmasken anschließend die Genauigkeit (Accuracy) in Abhängigkeit der Konfidenz sowie der Pixel-Position bestimmt werden [6].

4. Methoden zur Kalibrierung

Bei bekanntem Kalibrierungsfehler eines Modells können verschiedene Methoden eingesetzt werden, um eine nachträgliche Rekalibrierung der geschätzten Unsicherheiten vornehmen zu können. Es wird dabei zwischen Binning- und Skalierungsverfahren unterschieden. Ähnlich zu der Berechnung des ECE werden bei den Binning-Verfahren (bspw. Histogram Binning [7]) alle Prädiktionen anhand der Konfidenz in verschiedene Bins eingeteilt. Für jedes Bin kann anschließend die Genauigkeit bestimmt werden. Somit ist es möglich, die Genauigkeit in Abhängigkeit der Konfidenz zu ermitteln. Des Weiteren kann das Binning dazu verwendet werden, neuen Prädiktionen auf Basis der Konfidenz eine neue, kalibrierte Wahrscheinlichkeit zuzuordnen. Bei Objektdetektionsmodellen konnte bereits festgestellt werden, dass die Fehlkalibrierung nicht nur von der Konfidenz, sondern auch von der Position und Größe der Objekte abhängig ist [2]. Aus diesem Grund wird für Detektionsmodelle ein multidimensionales Histogram Binning verwendet, welches die Samples nach der Konfidenz, der

Position und der Größe in verschiedene Bins einteilt [2]. Anschließend kann für jedes Bin die Precision bestimmt und als kalibriertes Maß für die Konfidenz weiterverwendet werden.

Im Gegensatz zu den Binning-Verfahren verfolgen die Skalierungsverfahren eine andere Strategie. Ein neuronales Netz berechnet für eine Eingabe x_i einen Logit z_i , welcher durch eine Aktivierungsfunktion $\phi(\cdot)$ (bspw. Sigmoid-Funktion) zu einer Konfidenz umgerechnet wird, sodass $\hat{p}_i = \phi(z_i)$. Bei der Kalibrierung anhand eines Skalierungsverfahrens (bspw. Logistic Calibration [8], Beta Calibration [4]) werden die Logits des Modells so verändert, dass die Konfidenz nach der Aktivierungsfunktion eher der beobachteten Genauigkeit entspricht. Bei der logistischen Kalibrierung (oder auch Platt Scaling) [8] wird unter der Annahme normalverteilter Konfidenzen für jede Klasse eine kalibrierte Konfidenz \hat{q}_i durch den Skalierungsparameter a und einen Bias c nach

$$\hat{q}_i = \phi(a \cdot z_i + c)$$

berechnet. Bei der Beta Calibration [4] werden stattdessen Beta-Verteilungen für die Konfidenzen im Intervall [0,1] angenommen, sodass die Kalibrierung anhand der Parameter a, b und Bias c nach

$$\hat{q}_i = \phi(a \cdot \log(z_i) - b \cdot \log(1 - z_i) + c)$$

berechnet wird. Ähnlich zu dem multidimensionalen Histogram Binning können auch die Skalierungsverfahren um Positions- und Größeninformationen erweitert werden. Zu diesem Zweck werden für die Verteilungen der Daten multivariate Normal- oder Beta-Verteilungen [2] eingesetzt. Die Datenverteilungen können beispielsweise so modelliert werden, dass alle Informationen als statistisch unabhängig betrachtet werden. Dies hat den Vorteil eines wesentlich einfacheren und schneller zu trainierendem Kalibrierungsverfahrens. Im Gegensatz dazu können Konfidenz-, Positions- und Größeninformationen aber auch als statistisch voneinander abhängige Variablen interpretiert werden. Diese Verfahren sind komplexer zu trainieren, es ist aber möglich, Korrelationen zwischen den einzelnen Größen zu erfassen [2].

5. Experimente

Alle hier vorgestellten Ergebnisse sind Teil der Experimente aus dem Kapitel „Confidence Calibration for Object Detection“ des Buchs „Deep Neural Networks and Data for Automated Driving - Robustness, Uncertainty Quantification, and Insights Towards Safety and Segmentation“ [6]. Dort werden die Untersuchungen und Auswertungen auf weitere Netzwerke erweitert. Zur Evaluierung der zuvor beschriebenen Kalibrierungsmethoden verwenden wir zum einen die Objekt-Prädiktionen (ohne Segmentierungsmaske) eines Mask R-CNN Netzwerks [9], welches für den Cityscapes Datensatz trainiert wurde [10]. Zum Training der Kalibrie-

Tabelle 1: Kalibrierungsergebnisse zur Objektdetektion unter Verwendung der Konfidenz sowie Positions- und Größeninformationen für das Training der Kalibrierungsmethoden und die Berechnung des Kalibrierungsfehlers (D-ECE) [6].

	D-ECE [%]	Brier	NLL	AUPRC
Uncalibrated	10,168	0,146	0,500	0,951
Multidimensional Histogram Binning	5,241	0,151	0,535	0,854
Dependent Logistic Calibration	4,551	0,134	0,440	0,946
Dependent Beta Calibration	6,117	0,134	0,413	0,925

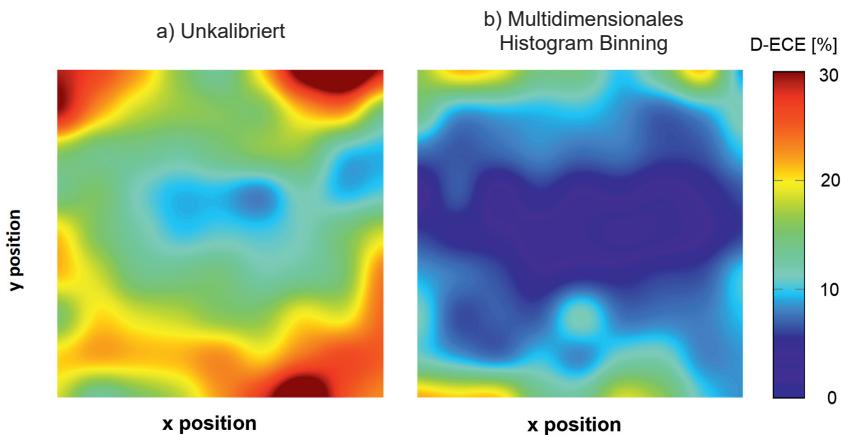


Bild 3: Positionsabhängige Fehlkalibrierung eines Objektdetektors vor (a) und nach (b) der Kalibrierung mit multidimensionalem Histogram Binning (vgl. [6]). Der Kalibrierungsfehler kann in Abhängigkeit von der Position erfolgreich reduziert werden.

rungsmethoden werden die „Münster & Lindau“ Sequenzen aus dem Validierungsdatensatz verwendet, während die „Frankfurt“ Sequenzen zur Bewertung der Methoden zum Einsatz kommen. Zur Evaluierung der Kalibrierungsmethoden für die Instanzensegmentierung verwenden wir ein Mask R-CNN Netzwerk, welches für den MS COCO Datensatz [12] trainiert wurde. Hierbei teilen wir die Validierungsdaten des Datensatzes in zwei gleich große Teile und verwenden den einen Teil zum Training und den zweiten Teil zur Auswertung der Kalibrierungsmethoden. Eine Bewertung erfolgt anhand des D-ECE [2], dem Brier Score und der Negative Log Likelihood (NLL). Je geringer diese Metriken ausfallen, desto geringer ist der

Tabelle 2: Kalibrierungsergebnisse zur Instanzensegmentierung unter Verwendung der Konfidenz sowie der Positionsinformationen der jeweiligen Pixel innerhalb der Detektionen für das Training der Kalibrierungsmethoden und die Berechnung des Kalibrierungsfehlers [6].

	D-ECE [%]	Brier	NLL	AUPRC
Uncalibrated	23,409	0,222	0,940	0,663
Multidimensional Histogram Binning	13,640	0,171	0,776	0,760

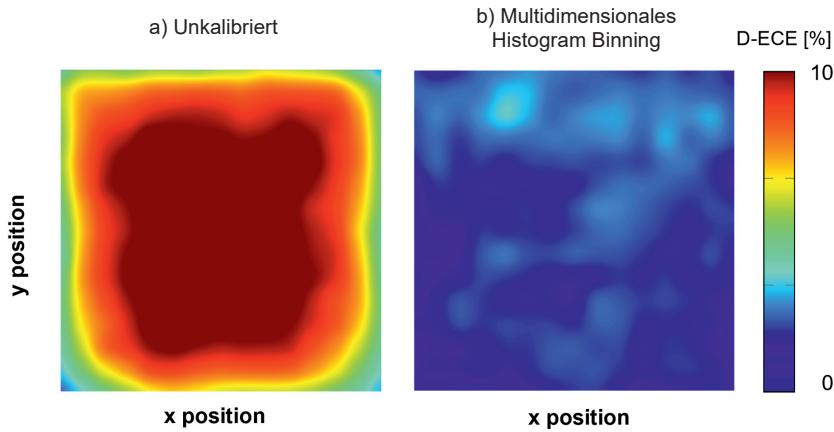


Bild 4: Positionsabhängige Fehlkalibrierung eines Instanzensegmentierungsmodells vor (a) und nach (b) der Kalibrierung mit multidimensionalem Histogram Binning (vgl. [6]). Der Kalibrierungsfehler kann in Abhängigkeit von der Position erfolgreich reduziert werden.

Kalibrierungsfehler. Final wird die Performance der Netzwerke mit dem AUPRC Score (*Area under Precision/Recall Curve*) sowohl für die Objektdetektion als auch für die Instanzensegmentierung bewertet. Diese Metrik wird zur Bewertung der Detektionsperformance verwendet. Je höher der Score, desto besser fallen die Prädiktionen aus. Sowohl zur Messung der Fehlkalibrierung als auch zum Training der Kalibrierungsmethoden werden alle Informationen des Objektdetektors herangezogen. Dies schließt die Konfidenz, die Position sowie die Größe der prädizierten Objekte mit ein. Gleiches gilt für das Training und die Bewertung der Methoden zur Segmentierung – hier werden die Konfidenz, die Position sowie der kürzeste Abstand jedes Pixels zur Konturengrenze verwendet. Zur Bewertung der semantischen

Tabelle 3: Kalibrierungsergebnisse zur semantischen Segmentierung unter Verwendung der Konfidenz sowie der Positionsinformationen der jeweiligen Pixel für das Training der Kalibrierungsmethoden und die Berechnung des Kalibrierungsfehlers [6].

	D-ECE [%]	Brier	NLL	mIoU
Uncalibrated	0,186	0,060	0,139	0,623
Multidimensional Histogram Binning	0,119	0,062	0,189	0,589

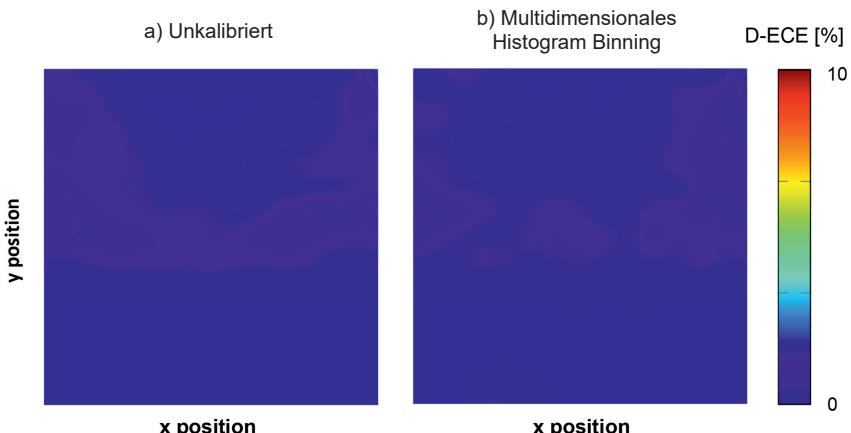


Bild 5: Positionsabhängige Fehlkalibrierung eines semantischen Segmentierungsmodells vor (a) und nach (b) der Kalibrierung mit multidimensionalem Histogram Binning (vgl. [6]). Das Modell weist von vornherein einen sehr geringen Kalibrierungsfehler auf, daher können die Kalibrierungsmethoden in diesem Fall keine wesentliche Veränderung mehr erzielen.

Segmentierung verwenden wir denselben Datensatz und als Segmentierungsnetzwerk ein Deeplabv3+ Modell [11]. Die Ergebnisse zu den Untersuchungen einer Objekterkennung werden in Tabelle 1 und Abbildung 3 dargestellt. Anhand der Ergebnisse ist erkennbar, dass mit jeder verwendeten Kalibrierungsmethode der Kalibrierungsfehler reduziert werden konnte. In diesem Fall erzielen die Skalierungsmethoden die geringste Fehlkalibrierung. Der Kalibrierungsfehler kann in Abhängigkeit von der Position über alle Bildbereiche hinweg reduziert werden. Im Gegensatz zu den herkömmlichen Kalibrierungsverfahren lässt sich die positionabhängige Kalibrierung nicht als monoton steigende Funktion abbilden. Aus diesem Grund verändern diese Verfahren die Reihenfolge der Prädiktionen basierend auf deren

Konfidenz. Dies hat wiederum einen geringen, negativen Einfluss auf den AUPRC-Score. Im Vergleich zu der Verbesserung in der Kalibrierung ist dieser Effekt jedoch vernachlässigbar gering. Somit kann die positionsabhängige Kalibrierung von Detektionsmodellen einen wertvollen Beitrag insbesondere für sicherheitskritische Systeme leisten, da sie zu statistisch besser kalibrierten Konfidenzwerten führen. Ähnlich zur Objektdetektion kann auch bei Segmentierungsmodellen ein Kalibrierungsfehler beobachtet werden. Da in diesem Fall jeder Pixel als eigene Eingabe zu der Trainings- und Kalibrierungsmenge gezählt wird und somit eine große Datenmenge entsteht, beschränken sich die Untersuchungen für die Segmentierungsmodelle auf das multidimensionale Histogram Binning, da das Training der Skalierungsverfahren mit zunehmender Datenmenge zu rechenintensiv wird. Die Kalibrierungsergebnisse für das Instanzensegmentierungsmodell sind in Tabelle 2 sowie Abbildung 4 dargestellt. Auch in diesem Fall weist das Modell eine Fehlkalibrierung auf, die mithilfe des Histogram Binnings reduziert werden kann. Darüber hinaus ist anhand des AUPRC Scores erkennbar, dass neben der eigentlichen Kalibrierung auch die Qualität der Segmentierungsmasken durch positionsabhängige Kalibrierung verbessert wird. Im Gegensatz dazu weisen die Modelle in der semantischen Segmentierung (Tabelle 3 und Abbildung 5) eine nur sehr geringe Fehlkalibrierung auf, welche durch Kalibrierungs-methoden nicht wesentlich verbessert werden kann. Ein Grund für den großen Unterschied zwischen den Kalibrierungseigenschaften zwischen Instanz- und semantischer Segmentierung könnte der Unterschied in dem Training der einzelnen Modelle sein. Im Gegensatz zur semantischen Segmentierung muss ein Instanzensegmentierungsmodell zunächst die Position des Objekts korrekt schätzen, um qualitativ gute Ergebnisse bei der Maskeninferenz erzielen zu können. Dies ist bei der semantischen Segmentierung nicht erforderlich [6].

6. Zusammenfassung

Für sicherheitskritische Systeme sind statistisch belastbare Aussagen über die von einem neuronalen Netz selbst eingeschätzten Unsicherheiten über seine Prädiktionen von entscheidender Bedeutung. Im Rahmen dieser Arbeit haben wir daher die statistische Fehlkalibrierung von Objektdetektions- und Segmentierungsnetzwerken untersucht und ausgewertet. Zu diesem Zweck haben wir die Definition der statistischen Kalibrierung von einfachen Klassifikationsnetzwerken auf die Objektdetektion und die Segmentierung erweitert und ein Maß zur positions- und großenabhängigen Messung der Fehlkalibrierung vorgestellt. Darüber hinaus haben wir bestehende Methoden aus der Unsicherheitskalibrierung von neuronalen Netzen aus dem Bereich der Klassifikation auf die Bereiche der Objektdetektion und Segmentie-

rung hin so erweitert, dass eine positions- und größenabhängige Unsicherheitskalibrierung möglich ist.

Während bei Modellen der semantischen Segmentierung in unseren Experimenten nur ein minimaler Kalibrierungsfehler festgestellt werden konnte, wurde bei Detektions- und Instanzensegmentierungsmodellen eine positionsabhängige Fehlkalibrierung festgestellt, die anhand der vorgestellten Kalibrierungsmethoden erfolgreich reduziert werden konnte. Somit ist es möglich, statistisch belastbare Aussagen über die Unsicherheiten von Detektions- und Segmentierungsmodelle zu erhalten. Dies ist eine Grundlage für den Einsatz solcher Modelle insbesondere in sicherheitskritischen Anwendungen.

Danksagung

Die Arbeiten an dieser Veröffentlichung wurden durch das Bundesministerium für Wirtschaft und Energie innerhalb des Forschungsprojekts „Methoden und Maßnahmen zur Absicherung von KI basierten Wahrnehmungsfunktionen für das automatisierte Fahren (KI-Absicherung)“ sowie aus Mitteln des Europäischen Fonds für regionale Entwicklung (EFRE) innerhalb des Forschungsprojekts „Centrum für Automatisierte Mobilität“ gefördert. Die Autoren bedanken sich bei allen Projektpartnern für die erfolgreiche Zusammenarbeit.

Literaturverzeichnis

- [1] Guo, C., Pleiss, G., Sun, Y., & Weinberger, K. Q. (2017, July). On calibration of modern neural networks. In International Conference on Machine Learning (pp. 1321-1330). PMLR.
- [2] Küppers, F., Kronenberger, J., Shantia, A., & Haselhoff, A. (2020). Multivariate confidence calibration for object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (pp. 326-327).
- [3] Schwaiger, F., et al. (2021). From Black-box to White-box: Examining Confidence Calibration under different Conditions. arXiv preprint arXiv:2101.02971.
- [4] Kull, M., Silva Filho, T., & Flach, P. (2017, April). Beta calibration: a well-founded and easily implemented improvement on logistic calibration for binary classifiers. In Artificial Intelligence and Statistics (pp. 623-631). PMLR.
- [5] Naeini, M. P., Cooper, G., & Hauskrecht, M. (2015, February). Obtaining well calibrated probabilities using bayesian binning. In Twenty-Ninth AAAI Conference on AI
- [6] Küppers, F., Kronenberger, J., Schneider, J., & Haselhoff, A. (2021, in press): Confidence calibration for object detection. In Fingerscheid, T., Gottschalk, H. & Houben, S. (eds): Deep Neural Networks and Data for Automated Driving - Robustness, Uncertainty Quantification, and Insights Towards Safety and Segmentation. Springer Nature Switzerland, Cham, CH
- [7] Zadrozny, B., & Elkan, C. (2001, June). Obtaining calibrated probability estimates from decision trees and naive bayesian classifiers. In Icmi (Vol. 1, pp. 609-616).
- [8] Platt, J. (1999). Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods. Advances in large margin classifiers, 10(3), 61-74.
- [9] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 2961-2969).
- [10] Cordts, M. et al. (2016). The cityscapes dataset for semantic urban scene understanding. In Proceedings of the IEEE CVPR (pp. 3213-3223).
- [11] Chen, L. C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2017). Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs. IEEE transactions on pattern analysis and machine intelligence, 40(4)
- [12] Lin, T. Y., et al. (2014, September). Microsoft coco: Common objects in context. In European conference on computer vision (pp. 740-755). Springer, Cham.

Projekt COPE – Collective Perception zur Vermeidung von Kollisionen und gefährlichen Situationen mittels V2X

Vorgehensweise und Lösungskonzepte für die Kollisionsvermeidung mit ungeschützten Verkehrsteilnehmern

Dr. Andreas Kuhn, Markus Lauterbacher, Elvira Thonhofer, Karin Glader, ANDATA, Hallein, Österreich;
Patrick Luley, JOANNEUM RESEARCH, Graz, Österreich;
Dr. Thomas Novak, SWARCO, Perchtoldsdorf, Österreich;
Dr. Walter Aigner, HiTec, Wien, Österreich;
Prof. Dr. Wolfgang Schildorfer, Logistikum FH Steyr, Steyr, Österreich

Zusammenfassung

Durch die Vernetzung von Fahrzeugen untereinander und mit der Infrastruktur ergeben sich völlig neue Möglichkeiten in der Fahrzeug- und Verkehrssicherheit. Im Projekt COPE (Collective Perception) wird ermittelt, wie man diese Vernetzung für eine verbesserte Vermeidung von Kollisionen und gefährlichen Situationen mit ungeschützten Verkehrsteilnehmern an ampelgeregelten Kreuzungen nutzen kann. Der vorliegende Artikel stellt die Vorgehensweise im Projekt vor und motiviert diese durch die spezifischen Herausforderungen der gegebenen Problemstellung.

Abstract

The introduction of vehicles and infrastructure connectivity enables new possibilities in vehicle and road safety. The project COPE (Collective Perception) strives to evaluate the potential from connecting vehicles with each other and with the infrastructure for the improved avoidance of collisions and dangerous situations with vulnerable road users at urban intersections with smart traffic lights. This article presents the approach in the project and motivates it by the specific challenges of the given problem.

Einleitung und Kontext

Nach der Reduktion der Kollisionsfolgen durch Insassenrückhaltesysteme (vor allem durch Insassenrückhaltesysteme wie Airbags) und der Vermeidung von Kollisionen (mit autonomen Notbremsen und Warnsystemen) steht gerade der nächste technologische Schub durch die Einführung der Fahrzeug- und Infrastruktur-Vernetzung an. Dadurch soll neben der weiteren

Verbesserung der Kollisionsvermeidung vor allem auf die Vermeidung von gefährlichen Situationen am Weg zur Vision Zero abgezielt werden.

Technisch gesehen geht es darum, dass Aktivierung der autonomen Notbremsen und Warnsysteme nicht mehr bloß eine unilaterale Aktion durch bloße Auswertung und Interpretation der Sensorsignale und eines Fahrzeugs darstellen. Mit den erweiterten Kommunikationsmöglichkeiten durch die Vernetzung der Fahrzeuge untereinander sowie mit der Infrastruktur sollen einerseits Verkehrsteilnehmer ihre Konfigurationen und Intentionen teilen, um so weitere Verbesserungen in der unmittelbaren Kollisionsvermeidung zu ermöglichen. Andererseits sollen Fahrzeuge und die Infrastruktur auch ihre Wahrnehmungen der Bewegungen von den selbst sensierten Umgebungsobjekten teilen können, sodass die gemeinsame Wahrnehmung (Collective Perception) mehrerer Fahrzeuge in Kombination mit der Wahrnehmung und lokalen Erfahrung der Infrastruktur eine „zusätzliche Erweiterung des Sichtfeldes“ für die vorausschauende Vermeidung von Gefahren ergibt. Diese multilaterale Kommunikation erhöht allerdings die Komplexität der Aufgabenstellung beträchtlich, sodass spezielle Methoden notwendig sind. Dies gilt insbesondere, wenn die Lösung allgemeingültig für beliebige Fahrzeuge (PKW, LKW) in Kombination mit ungeschützten Verkehrsteilnehmern (Radfahrer, Fußgänger, Moped- und Motorradfahrer) und der Infrastruktur in beliebigen Szenarien und Konstellationen erfolgen soll.

Herausforderungen bei der kooperativen Kollisionsvermeidung

Die Themenstellung der vernetzten Kollisionsvermeidungssysteme beinhaltet eine Reihe von spezifischen Herausforderungen und Anforderungen, welche in der Funktionsentwicklung und System-Absicherung berücksichtigt und antizipiert werden müssen.

- Kunden-Akzeptanz bzgl. Erkennungsraten und Falsch-Auslösungen

Warn- und Informationssysteme benötigen einen Zeithorizont von mindestens 1 bis 2 Sekunden vor einer potenziellen Kollision. In diesen Zeiträumen können sich Situationen sehr unterschiedlich weiterentwickeln bzw. ist die Schätzung der Time-To-Collision mit großen Unsicherheiten behaftet, sodass es zu vielen Falschauslösungen kommen kann, wenn hier keine ausreichend genauen und performanten Modelle zur Bewegungs- und Verhaltensprognose sowie zur Interpretation der Situation eingesetzt werden. Die Anzahl der Falschauslösungen beeinflusst idR die Kunden-Akzeptanz deutlich. Ohne Verbesserung des Informationsgehalts der Sensorsignale und Car2X-Nachrichten ergibt sich eine Reduktion der Falschauslösungen nur mit einer einhergehender Reduktion der Erkennungsraten, sodass es zu entsprechenden Abstrichen in der Wirkung und im Nutzen der Funktion kommt.

- Interaktive Verhaltensmuster der Verkehrsteilnehmer

Der Einsatz entsprechender Warn- und Informationssysteme kann das interaktive Verhalten der Verkehrsteilnehmer derart ändern, sodass es zu überraschenden Rückkopplungseffekten kommen kann. Derartige Effekte sind entsprechend zu antizipieren und mit Adoptionsfähigkeit und lernfähiger Ausprägung der Funktionen zu berücksichtigen, indem diese sich einfach und schnell an sich ändernde Verhaltensmuster der anderen Verkehrsteilnehmer anpassen lassen (>Adoptions- und Lernfähigkeit).

- Repräsentativität und Relevanz der Szenarien

Die Szenarien zur Abstimmung und Auslegung der Funktionen und Prognose-Modelle müssen relevant und repräsentativ für die tatsächlichen Einsatz- und Betriebsbereiche der Fahrzeuge sein. Beispielsweise ist darauf zu achten, dass man sich nicht nur auf eine geringe Menge von dezidierten Kreuzungs- und Straßen-Konfigurationen und Topologien spezialisiert. Hier ist auf eine ausreichende Allgemeingültigkeit bzgl. Verhalten der FahrerInnen, Verhalten der anderen Verkehrsteilnehmer, Straßen-Konfigurationen, Bedingungen & Topologien, Verkehrsregeln, Umwelt- und Umgebungsbedingungen sowie Verkehrssituationen zu achten.

- Komplexität der verteilten Algorithmen

Bisher sind die Fahrzeugsysteme ausschließlich unilateral ausgebildet. D.h., dass diese nur auf Basis eigener aufgenommener Informationen und Daten unabhängig entscheiden. Mit den neuartigen, vernetzten Funktionen ergeben sich verteilte Algorithmen, welche sich gegenseitig beeinflussen und mit welchen erst durch die koordinierten und kooperativen Entscheidungen verbesserte Wirkung erreicht wird. Verteilte Algorithmen ergeben aber nicht unbedingt eine gesamthafte Verbesserung, wenn diese nicht gemeinsam und als gesamtheitliches System-of-Systems abgestimmt sind. In der Konsequenz ist damit dann auch die organisatorische Komplexität zu berücksichtigen, wenn unterschiedliche Teile der verteilten Algorithmen von unterschiedlichen Organisationseinheiten umgesetzt werden. Dazu werden klare Schnittstellen benötigt, welche allgemein für beliebige Verkehrsteilnehmer gelten müssen und deren Fähigkeiten und dynamischen Möglichkeiten und Intentionen in einer generischen Form berücksichtigen. Um kooperieren zu können benötigt man eine klare Vorstellung davon, was die anderen können und wollen.

Lösungskomponenten und Vorgehensweise

- Probabilistische Ansätze und Stochastische Simulation zur Bewegungsprognose

Erste Konsequenz aus den oben beschriebenen Herausforderungen ist die durchgängige Einführung einer probabilistischen Sichtweise und die Akzeptanz und Berücksichtigung, dass es

sich hier um eine stochastische Problemstellung handelt, welche entsprechende Lösungsmethoden und Vorgehensweisen benötigt.

Die Schätzung von Kollisionsrisiken benötigt die Prognose der Bewegungen der verschiedenen Verkehrsteilnehmer. Diese Bewegungen sind in der Regel nicht vollständig deterministisch, sondern hängen von einer Reihe von Einflussfaktoren ab, wie etwa dem Verhalten und der Fähigkeiten der Fahrer, dem Zustand und Fähigkeiten des Fahrzeugs, den Umgebungsbedingungen wie etwa dem Fahrbahnzustand und den geometrischen, topologischen Verhältnissen der Fahrbahn, dem Wetter, dem Verhalten der anderen Verkehrsteilnehmer, der subjektiven Interpretation der Verkehrsregeln, Verkehrszeichen und Signale, usw. All die hier angeführten Größen sind stochastischer Natur.

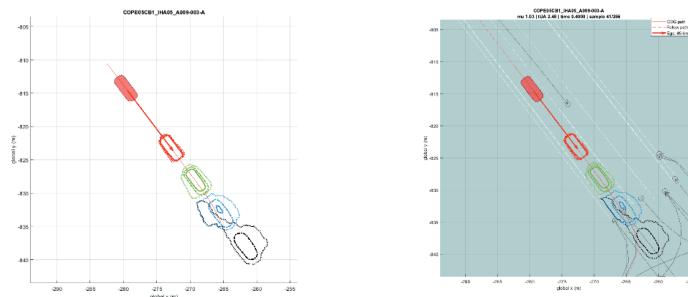


Bild 1: Aufenthalts wahrscheinlichkeiten eines fahrenden Fahrzeugs nach 1 / 1,5 / 2 / 2,5 Sekunden in Rot / Blau / Grün / Schwarz in Abhängigkeit des Fahrerverhalten und der Fahrbahn beschaffenheit (links) sowie unter zusätzlicher Berücksichtigung des Straßenverlaufs (rechts)

Mathematisch gesehen handelt es sich dabei um bedingte Wahrscheinlichkeiten gemäß dem Satz von Bayes und den dazugehörigen weiterführenden Ansätzen (z.B. Bayes-Netze, graphische Modelle, Monte-Carlo-Markov-Ketten, u.Ä.). In Bild 1 links sind dabei beispielsweise die Aufenthalts wahrscheinlichkeiten eines fahrenden Fahrzeugs in Abhängigkeit von der Fahr intention des Fahrers bei normaler Vorausfahrt („Normal Drive Straight“-Annahme), dem Fahrbahnzustand und dem Fahrzustand des Fahrzeugs nach verschiedenen Zeitpunkten in der Zukunft eingezeichnet. Wenn man den Verlauf und die Geometrie der Fahrbahn mitberücksichtigt, dann ergeben sich unter der Annahme des Fahrverhaltens des gegebenen Fahrers, dass diese sich entsprechend seiner Fähigkeiten und Intentionen innerhalb der Fahrspur

bewegt, leicht unterschiedliche Aufenthaltswahrscheinlichkeiten, wie man der Abbildung rechts in Bild 1 entnehmen kann.

Bei Kurvenfahrten mit anderen Fahrern, anderen Fahrzeugen, angepassten Verhaltensmustern bei Gegenverkehr, etc. sind die Verteilungen entsprechend angepasst. Die Wahrscheinlichkeitsverteilungen dafür kann man aus Naturalistic Driving Studien, aus Flottendaten oder aus Field Operational Tests beziehen. Die Berechnung der bedingten Wahrscheinlichkeiten der jeweiligen Aktionsannahmen erfolgt dann mittels Monte-Carlo-Simulationen auf Basis der empirischen Verteilungen.

Wenn man die Aufenthaltswahrscheinlichkeiten mehrerer Verkehrsteilnehmer nimmt, kann man damit die Kollisionsgefahr abschätzen. In Bild 2 auf der linken Seite sind die Aufenthaltswahrscheinlichkeiten in eines links abbiegenden PKWs und eines geradeaus fahrenden, entgegenkommenden Motorrads auf einer Kreuzung eingezeichnet. Zum Zeitpunkt 1,5 Sekunden später ergibt sich hier keine Kollisionsgefahr gemäß dem Umstand, dass sich die Aufenthaltsflächen der beiden Verkehrsteilnehmer nicht überlappen und ausreichend Abstand aufweisen. Im Bild 2 auf der rechten Seite gibt es hingegen eine Überlappungsfläche, sodass eine Kollisionsgefahr besteht. Wenn die Kollision verhindert werden soll, muss hier eine alternative Aktion gewählt werden, welche von der ursprünglichen Bewegungsannahme abweicht.

Je nach Fähigkeiten und Aktionsmöglichkeiten der Verkehrsteilnehmer benötigt man unterschiedliche Prognose-Horizonte für die Entscheidung der Alternativ-Aktionen. „Sportliche“ Verkehrsteilnehmer benötigen kürzere Zeiten aufgrund höherer Beschleunigungsfähigkeiten als „komfortabel“ eingestellt Verkehrsteilnehmer.

Die Performance zur Identifikation einer Kollisionsgefahr und die Ableitung einer passenden Alternativ-Aktion zur Vermeidung der Kollision hängt damit im Wesentlichen von der Leistungsfähigkeit der Prognose der Bewegungen und der Bewegungsmöglichkeiten der anderen Verkehrsteilnehmer sowie der eigenen Bewegungsintention ab. Diese können deutlich gesteigert werden, wenn

- Verkehrsteilnehmer ihre Intentionen, Fähigkeiten und Reserven (also der Operational Design Domain) teilen,
- Bewegungsintentionen von Verkehrsteilnehmern geteilt werden, welche nicht vom eigenen Fahrzeugen gesehen werden,
- der lokale Kontext (wie etwa Layout einer Kreuzung, Sichtverhältnisse, Verkehrssignale, lokale Regelungen, usw.) in die Prognose mit einbezogen werden können.

All diese Informationen lassen sich mittels Vernetzung der Verkehrsteilnehmer übertragen. Durch die Verbesserung der Gefahren-Klassifikation reduzieren sich die Falsch- und Nicht-Erkennungsraten und steigt damit einhergehend Kunden-Akzeptanz und Nutzen deutlich.

Letztendlich ergibt sich durch die so beschriebenen Bewegungsprognosen unter den Bedingungen verschiedener Aktionen (wie normale Weiterfahrt, moderate Bremsung nach einer Warnung, intensive Bremsung durch autonome Notbremse mit jeweils dedizierten Beschleunigungsprofilen) eine Identifikation der besten Aktionen für die jeweilige Situation. Diese können entweder für eine direkte Manöverplanung oder für eine explizite Aktionsspezifikation in der Entwicklung und Absicherung verwendet werden.

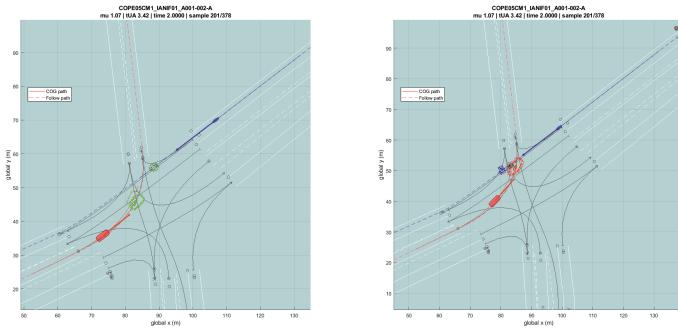


Bild 2: Aufenthaltswahrscheinlichkeiten zweier Verkehrsteilnehmer (hier links abbiegender PKW und geradeaus fahrendes Motorrad auf einer Kreuzung) ohne Überlappung links und mit Überlappung der Aufenthaltswahrscheinlichkeiten recht.

- Szenario-Management und beispielbasierte Anforderungsrepräsentation

Im vorigen Abschnitt wurde gezeigt, wie man Situationen mit Kollisionsgefahr erkennt und daraus alternative Aktionen zur Kollisionsvermeidung ableitet. Je nachdem, mit welchen Bewegungsprofilen (sportlich mit hohen Beschleunigungsvermögen oder komfortabel mit geringeren Beschleunigungen) ergeben sich dabei jeweils unterschiedliche Aktionszeiten als Anforderungen an Kollisionsvermeidungssysteme wie autonomes Bremsen/Ausweichen oder Informieren und Warnen der Fahrer. Zur Entwicklung und Implementierung eines robusten Systems reicht allerdings nicht die Betrachtung weniger, einzelner Fälle. Stattdessen sind eine Vielzahl von derartigen Szenarien zu untersuchen und als Auslegungsgrundlage für Kollisionsvermeidungssysteme mit unterschiedlichen Variationen aller relevanten Konstellationen heranzuziehen.

Das grundsätzliche Schema dafür ist in Bild 3 abgebildet. Szenarien, wie das oben beschriebene, werden in großer Menge von 100 Tausenden bis mehreren Millionen automatisiert erstellt, indem Variationen der Anfangs-/Randbedingungen sowie der Konfigurationsparameter

von Aktionsannahmen der Fahrzeuge und Verkehrsteilnehmer, der Verkehrsregelung, der Verkehrssituation, des Verhaltens der Verkehrsteilnehmer, des Verhaltens des Fahrers und der anderen Verkehrsteilnehmer, der Infrastruktur und Umgebungsbedingungen sowie von lokalen Verkehrsregeln ausgeführt werden.

Die Szenarien und Situationen werden mittels Monte-Carlo-Simulationen derart variiert, dass sich idealerweise eine Menge von Szenarien ergibt, welche repräsentativ für das operative Einsatzgebiet des zu entwickelnden Systems sind. Zusätzlich werden noch weitere relevante Fälle in den Szenarien-Katalog aufgenommen, welche sich aus den verschiedenen System-Anforderungen ergeben. Dazu gehören etwa Gesetzesanforderungen und Verbrauchertests, wie der EuroNCAP und ähnliche.

Zusätzlich zu den erzeugten Aktionsspezifikationen werden für alle Szenarien noch die dazugehörigen Sensor-Signale und Nachrichten simuliert, wie sie das Fahrzeug sehen und hören würde. Damit ergibt sich eine beispielbasierte Repräsentation der funktionalen Anforderungen, welche klar und umfassend in daten-basierter Weise beschreibt,

- in welchen Szenarien und Situationen (= Wirkfeld bzw. Operational Design Domain des Systems)
- welche Aktionen wann (= Outputs der Regelungsalgorithmen)
- auf Basis welcher Sensor-Signale, Nachrichten-Informationen und Kontext-Information aus lokalen (HD-)Karten (= Inputs der Regelungsalgorithmen)

angesteuert werden sollen.

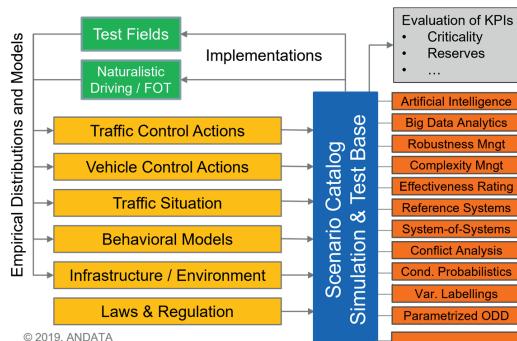


Bild 3: Grundlegendes Schema des Szenario-Managements

- Signal und Informationsbewertung mittels Maschine-Learning

In den vorigen Abschnitt wurde beschrieben, wie eine umfassende Szenarien-Datenbasis zur beispielbasierten Repräsentation der funktionalen Anforderungen an die Regelungsalgorithmen zur Kollisionsvermeidung von vernetzten Systemen aufgebaut wird. Hier soll nun gezeigt werden, wie man damit die besten Nachrichten zur Vermeidung der Kollisionen bestimmt. Dazu trainiert man Machine-Learning-Modelle mit den gewünschten Aktionen aus dem Szenarien-Pool als Modell-Outputs auf Basis ausgewählter Sets von Sensorsignalen, Kartendaten und dazugehörigen Car2X-Nachrichten als Modell-Input. Je nach Brauchbarkeit der Sensoren und Nachrichten ergeben sich dann Plots, wie exemplarisch in Bild 4 dargestellt. Sensoren und Nachrichten, welche einen positiven Beitrag für eine sichere Entscheidung zur Kollisionsvermeidung der jeweiligen Aktionen liefern, resultieren in Modellen mit besseren Erkennungsraten/Falschauslösungsralten, welche in der Grafik weiter links oben gelegen sind. Damit kann man klar entscheiden, ob eine bestimmte Nachricht eine positive Wirkung zur Kollisionsvermeidung beiträgt. Um eine Referenzierbarkeit der Aussagen garantieren zu können, sollte man dabei allerdings Machine-Learning-Verfahren verwenden, welche „universelle Approximatoen“ darstellen, z.B. etwa künstliche neuronale Netze mit ausreichender Anzahl an Neuronen. Damit kann die Suche nach den geeigneten Informationen aus Sensoren, Nachrichten und Karten auch automatisiert werden, sodass sich systematisch herausfinden lässt welche Nachrichteninhalte am besten für die Entscheidung zur Kollisionsvermeidung geeignet sind.

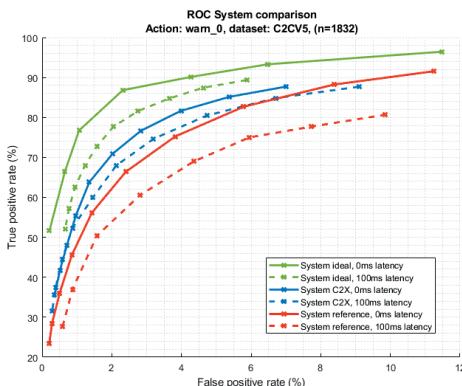


Bild 4: Erreichbare Erkennungsraten und Falschauslösungsralten für ausgewählte Sets von Sensoren und Nachrichten-Informationen.

Wenn mit den Maschine-Learning-Modellen keine ausreichenden Erkennungsraten und minimale Falschauslösungsrate erzielt werden können, liegt es in der Regel an sogenannten Anforderungskonflikten. D.h. dass in Szenarien, welche unterschiedliche Aktionen erfordern, aber die Informationen aus Sensoren, Nachrichten, Karten äquivalent sind. D.h. bei denselben Inputs in die Algorithmik werden unterschiedliche Outputs gefordert. Um derartige Konflikte aufzulösen, benötigt es eine der folgenden Maßnahmen:

- Vorhalt zusätzlicher Informationen (also z.B. Hinzunahme weitere Nachrichteninhalte, Kartendaten oder Sensorsignalen), welche die Szenarien unterscheidbar machen
- Änderung der Aktionsanforderungen (also z.B. später entscheiden, wenn sich ein Unterschied in den Informationen einstellt)
- Funktionseinschränkung durch Reduktion des Operational Design Domains (Wegnahme von Szenarien aus den Anforderungen)

Dies wiederholt man so lange, bis ausreichende und akzeptable Systemwirkung und Performance erzielt werden kann.

- Effektivitätsbewertung und Validierung

Als zusätzlichen Prozessschritt führt man noch eine Effektivitätsbewertung zur gesamthaften Validierung und Absicherung der Systemwirkung durch, so wie es in [4] beschrieben ist. In der Ausprägung von [4] lässt sich damit vorteilhaft auch noch identifizieren, welche System-Komponenten (Aktion, Algorithmik, Sensorik, Elektronik) die meiste Wirkung auf die System-Performance hat, um so die Entwicklungsaufwendungen gezielt zu steuern. Für die Nachrichten ergeben sich so Aussagen, mit welchen Nachrichten-Typen sich am meisten System-Wirkung erzeugen lässt. Methodische Details sind den Referenzen zu entnehmen.

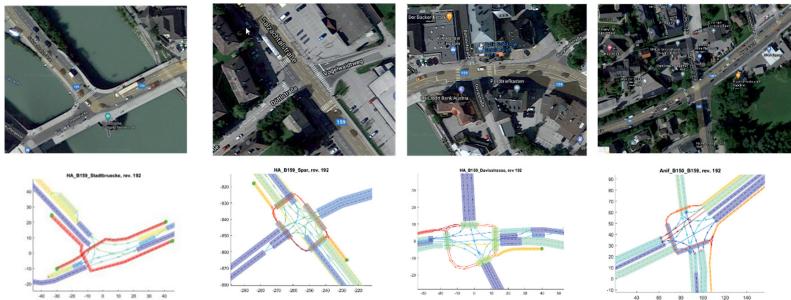


Bild 5: Szenarien aus Testumgebung der Projekte Connecting Austria und WienZWA in Hallein und Anif im Land Salzburg.

Anwendung im Projekt COPE- Collective Perception

Die im vorigen Abschnitt beschriebene Methodik wird durchgängig im Projekt COPE (Collective Perception) in direkter Nachfolge zu den Projekten Connecting Austria und WienZWA ausgeführt. Dabei wurden einerseits die Kreuzungen aus diesen Projekten als Szenarien modelliert und in einem detaillierten und ausgiebig validierten digitalen Zwilling abgebildet (vgl. Bild 5). Diese Gebiete umfassen teilweise Verkehrserfassungsanlagen, welche den vollen Zyklus aus Bild 3 erlauben. Zusätzlich sind diese Gebiete auch umfassend mit Fahrzeugen mit Naturalistic Driving-Ausstattung befahren worden, sodass für die Umgebung realistische Verhaltensmuster in das Szenario-Management einfließen. Mit der Ergänzung von umfassenden generellen Szenarien ergibt sich somit eine fundierte Aussage, welche vorgesehenen C-ITS-Nachrichten der anstehenden Day2-UseCases die beste Wirkung zur Vermeidung von Kollisionen mit ungeschützten Verkehrsteilnehmern aufweisen und welche idealerweise eine breite Implementierung erfahren sollten.

Danksagung

Das Projekt COPE wird von der österreichischen Forschungsförderungsgesellschaft und dem Bundesministerium für Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie gefördert sowie von einer Reihe weiterer Lol-Partner aktiv unterstützt. Weitere Details dazu findet man auf der Seite www.project-cope.eu.

- [1] Neubohn, A., Weiss, C., Keck, F., and Kuhn, A., KISS - a universal approach to the development and design of occupant restraint systems. In Proceedings of the 19th International Technical Conference on the Enhanced Safety of Vehicles, Washington, USA, pages 1-11, 05-0107, June 2005.
- [2] Hons. C., Neubohn, A., Weiss, C., Keck, F., Kuhn, A., KISS - Weiterentwicklung und Umsetzung eines methodischen Ansatzes zur Systemauslegung in der Fahrzeugsicherheit, VDI-Kongress Insassenschutz 2007
- [3] Keck F., Kuhn, A., Sigl. S., Altenbuchner, M., Palau, T., Roth, F., Stoll, J., Zobel, R., Kohsieck, A., Zander, A., Prüf- und Evaluierungsverfahren für den vorrausschauenden Fußgängerschutz im Spannungsfeld zwischen Simulation und Erprobung, VDI-Berichte 2010, Issue-Nr.: 2106
- [4] Sigl, S.; Golewski, T.; Miehling, T.; Kuhn, A., About Development Processes and Accompanying Performance Evaluations of Integral Automotive Safety Systems, In Proceedings of FISITA 2014 World Automotive Congress, Maastricht, F2014-AST-084, June 2014
- [5] Kuhn, A.; Carmona, J.; Novak, T.; Aigner, W.; Schildorfer, W.; Patz, D.; Test Fields and Advanced Accompanying Methods as Necessity for the Validation of Automated Driving, Proceedings of 7th Transport Research Arena TRA 2018, April 16-19, 2018, Vienna, Austria
- [6] Kuhn, A.; Carmona, J.; Schirrer, A.; Jakubek, S.; Hoch, T.; Pollhammer, K.; Hirschmann-Kraschl, K.; Schildorfer, W., Scenariobased approach in Connecting Austria for the development and validation of connect, cooperative, (semi-)automated driving in the case of intersections, Proceedings of 8th Transport Research Arena TRA 2020, April 27-30, 2020, Helsinki, Finland, Paper nr. # 870

UNICARagil – Disruptive Modular Architectures for Agile Automated Vehicle Concepts

T. Woopen, M.Sc., R. van Kempen, M.Sc.,

B. Lampe, M.Sc., Univ.-Prof. Dr.-Ing. L. Eckstein,

Institute for Automotive Engineering (ika) – RWTH Aachen University

Zusammenfassung

Vernetzte und automatisierte Mobilität sind essenzielle Bestandteile aktueller Megatrends im Bereich der Automobilentwicklung. Nach wie vor stehen Unternehmen, Start-Ups und Forschungseinrichtungen vor großen Herausforderungen bei der Entwicklung und Absicherung neuartiger Assistenz- und automatisierter Fahrfunktionen. Es wird deutlich, dass die Automatisierung ein Neudenken bestehender Architekturkonzepte im Automobil mit sich bringt. Das vom BMBF geförderte Projekt UNICARagil adressiert diese Architekturkonzepte auf Hardware- und Softwareebene. Es stellt dafür unter anderem eine neuartige Automotive-taugliche dienstorientierte Softwarearchitektur vor, welche künftige Herausforderungen in Bezug auf die Update- und Upgradefähigkeit von Software im Automobil adressiert.

Abstract

Connected and automated mobility are essential components of current megatrends in the field of automotive development. Companies, start-ups, and research institutions still face major challenges in development and safety assurance of novel assistance and automated driving functions. It is becoming clear that automation entails rethinking existing architectural concepts in the automobile. The BMBF-funded project UNICARagil addresses these architecture concepts regarding hardware and software. It presents, among other things, a novel automotive-ready service-oriented software architecture that addresses future challenges with respect to the updatability and upgradeability of software in the automobile.

1. Einleitung und Motivation

Vernetztes und automatisiertes Fahren wird die Zukunft der Mobilität nachhaltig beeinflussen. Neben der Elektromobilität und der „Shared Mobility“ bilden Vernetzung und Automatisierung die aktuellen Megatrends in der Mobilitätsforschung und -entwicklung [1]. Dennoch bestehen

weiterhin Herausforderungen bei der Implementation und Integration automatisierter Fahrfunktionen der SAE Level 3 [2] und höher. Bereits 2017 kündigte Audi mit dem A8 das erste Fahrzeug an, welches für eine Automatisierung nach SAE Level 3 entwickelt wurde [3]. Auch wenn Deutschland bereits 2017 das Straßenverkehrsgesetz für Level-3-Systeme öffnete, wurden einheitliche gesetzliche Anforderungen für die Zulassung von Level-3-Systemen („Automated Lane Keeping Systems“ kurz: ALKS) jedoch erst im Januar 2021 mit der UN-R157 [4] in Kraft gesetzt. Im Dezember 2021 erfüllt Mercedes-Benz als erstes Automobilunternehmen weltweit die Anforderungen und bietet die aktuelle S-Klasse inklusive eines SAE Level-3 Systems an [5]. Die Vorausschau und Ankündigungen zur weiteren flächendeckenden Einführung dieser Systeme sind jedoch deutlich konservativer geworden als noch vor einigen Jahren. So gab BMW bekannt, Level-3 Systeme bis 2025 auf den breiten Markt zu bringen [6].

Mit dem Gesetz zum autonomen Fahren schaffte Deutschland 2021 erneut die ersten Rahmenbedingungen für den Einsatz fahrerloser automatisierter Fahrzeuge des SAE Level 4 [7]. Diese Entwicklungen zeigen einerseits den politischen Willen, automatisiertes Fahren auf die Straße zu bringen, andererseits jedoch auch, dass die Komplexität der Systeme für eine Markteinführung nach wie vor hoch ist. Um Systeme für den breiten Markt tauglich zu machen, werden vollständig neue Software- und Hardwarearchitekturen benötigt, die in aktuellen Serienfahrzeugen kaum bis gar nicht vorhanden sind. Damit sind aktuelle Fahrzeuge nicht hinreichend update- und upgradefähig. In Anbetracht der kurzen Produktentwicklungszyklen von Software und Hardware zur Datenverarbeitung, welche spätestens mit dem automatisierten Fahren auch in Automobilen eine zentrale Rolle spielen, ist ein Wandel erforderlich. Langfristig sehen einige Automobilkonzerne daher eine Transformation von klassischen „Automobilern“ zu Softwarekonzernen als unumgänglich an.

2. Warum neue Architekturen für fahrerlose und automatisierte Fahrzeuge?

Die Einführung von automatisierten Fahrfunktionen höherer Automatisierungslevel bedingt nahezu automatisch den Einzug von hochkomplexen Softwaresystemen in das Fahrzeug. Die Beherrschung dieser Systeme ist jedoch sowohl aus Entwicklungs- allerdings auch aus Absicherungs- und Sicherheitsperspektive enorm anspruchsvoll. Zur Bewältigung der Komplexität sollte das System „automatisiertes Fahrzeug“ deshalb auf verschiedene Architektsichten abstrahiert werden.

Das vom Bundesministerium für Bildung und Forschung geförderte Konsortialprojekt UNI-CARagil [8] adressiert verschiedene neuartige Architekturkonzepte im Kontext der Entwicklung automatisierter und fahrerloser Fahrzeugkonzepte. Im Projekt forschen acht deutsche

Universitäten gemeinsam mit acht Industriepartnern in einem interdisziplinären Kooperationsprojekt. Dabei wird bewusst nicht auf existierenden Fahrzeugarchitekturen aufgesetzt. Stattdessen wird auf einer „grünen Wiese“ begonnen, um nicht auf proprietäre und historisch gewachsene Architekturen aufsetzen zu müssen. Dieser Ansatz ermöglicht, neutral aufzeigen zu können, wie Architekturen für neuartige automatisierte und fahrerlose Fahrzeuge in Zukunft gestaltet werden können.

Im Projekt werden vier fahrerlose Fahrzeuge von Grund auf neu aufgebaut und mit einer automatisierten und vernetzten Fahrfunktion auf Basis der ebenfalls entwickelten dienstesorientierten Softwarearchitektur ausgestattet. Abschnitt 5 zeigt eine kompakte Übersicht der bisherigen Projektergebnisse in Bezug auf den Fahrzeugaufbau.

Bild 1 zeigt eine zu Beginn des Projekts entstandene Übersichtsgrafik zur Darstellung der verschiedenen betrachteten Architekturebenen.

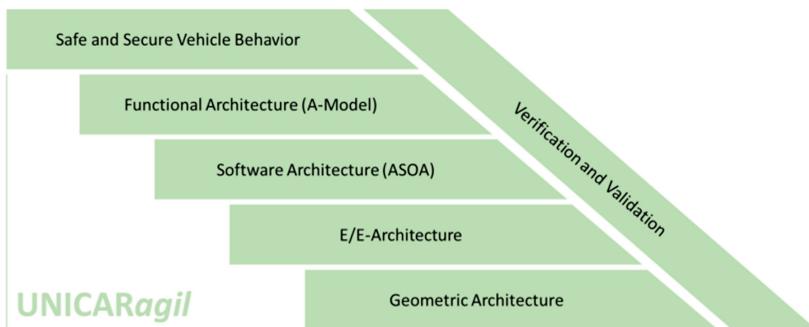


Bild 1: Verschiedene Architektsichten im Projekt UNICARagil [8]

Die Basis eines jeden Fahrzeugs bildet die geometrische Architektur. Sie beschreibt die tatsächliche Konstruktion des Fahrzeugs, definiert aber auch Sicherheitsziele und wichtige geometrische Randbedingungen wie Gewicht, Abmaß, Bauraum & Positionierung oder Sensorpositionierung. Die darüber liegende E/E-Architektur bildet ein Kernelement der Forschungen in UNICARagil. Sie beschreibt die Art und Weise, in der verschiedene Steuergeräte entworfen, entwickelt, verbaut und miteinander vernetzt sind. Wichtige Fragestellungen beziehen sich hierbei vor allem auf die Anzahl, Art und Aufgaben der verschiedenen Steuergeräte. Die wiederum darüber liegende Software-Architektur bildet die verschiedenen Software-Komponenten ab, welche auf den zuvor genannten Steuergeräten ausgeführt werden. In UNICARagil

wird erstmalig eine diensteorientierte Softwarearchitektur implementiert, welche bekannte Vorteile aus der, in der Informatik lange etablierten, softwaretechnischen Diensteorientierung für das Automobil anwendbar macht. Einzelne Softwaredienste bilden dabei nicht zwangsläufig einzelne Funktionen eines automatisierten Fahrzeugs ab. Daher wird in der übergeordneten Architekturebene eine gesonderte abstrahierte funktionale Architektur beschrieben. Diese beschreibt die Art und Weise, in der abstrakte Funktionalitäten eines automatisierten Fahrzeugs interagieren. Die oberste und daher auch am stärksten abstrahierte Architektsicht beschreibt das Fahrzeugverhalten und setzt damit Randbedingungen für eine sichere Bewegung des automatisierten Fahrzeugs im Straßenverkehr. Sicherheits- und Absicherungsbetrachtungen werden in UNICARagil über alle Ebenen hinweg koordiniert.

3. Von der Funktion zur Software-Architektur

Wird eine automatisierte und vernetzte Fahrfunktion beschrieben, so lässt sich diese in verschiedene abstrakte funktionale Blöcke dekomponieren. Die Darstellung der Interaktion dieser verschiedenen Funktionsbestandteile wird in der bereits in Abschnitt 2 vorgestellten funktionale Architektur repräsentiert. Die funktionale Dekomposition kann dabei in nahezu beliebigem Detailgrad erfolgen. Eine bekannte und stark abstrahierte funktionale Architektur kann bspw. schon die Aufteilung in „Sense“, „Plan“ und „Act“ darstellen. Im Projekt UNICARagil wurde jedoch ein detaillierterer Ansatz gewählt und weiterentwickelt. Zum Einsatz kommt hier das in Bild 2 dargestellte „A-Modell“.

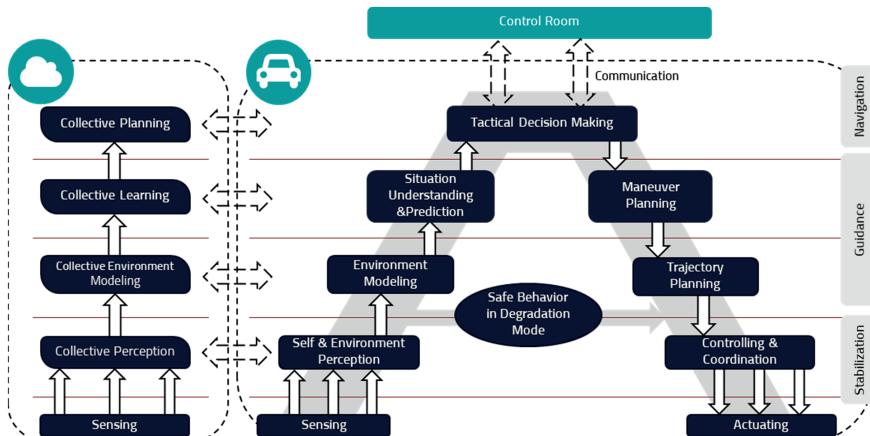


Bild 2: Visualisierung der funktionalen Architektur automatisierter und vernetzter Fahrzeuge im sog. A-Modell.

In diesem werden funktionale Blöcke in einer Weise abstrahiert, die eine Dekomposition der Funktionalität in Bezug auf das Fahrerverhaltensmodell nach Donges & Rasmussen [9] ermöglicht. Die Anordnung der Funktionalitäten im Fahrzeug wird dabei in Form des Buchstabens „A“ vorgenommen. Beginnend mit der Sensordatenerfassung werden die Daten verarbeitet, fusioniert, Objekte erkannt und prädiziert, bis schließlich auf oberster Ebene taktische Entscheidungsfindungen stattfinden können. Von dort an werden die einzelnen Schritte wieder konkretisiert, über die Pfad- und Trajektorienplanung, die Regelung sowie die Ansteuerung der Aktorik. Durch die im Projekt umgesetzte Funktion des sicheren Anhaltens [10] kann dabei ein beliebiger Teil des A-Modells im Falle von Fehlern überbrückt werden und das Fahrzeug auch bei funktionaler Degradation stets in einen risikominimalen Zustand überführt werden. Das zunächst fahrzeugzentrierte A-Modell wird in UNICARagil um zusätzliche essenzielle Komponenten erweitert. Zum einen wird eine Leitwarte, im Gesetz zum autonomen Fahren [7] als „technische Aufsicht“ bezeichnet, angebunden. Die Leitwarte ist dabei direkt an die Ebene der taktischen Entscheidungsfindung angebunden, da sie nicht aktiv in die Fahrt eingreift, wohl aber im Falle einer für das Fahrzeug selbst unlösbarer Situation aus dem Stand heraus taktische Befehle an das Fahrzeug übermitteln kann, welche in Folge eigenständig vom Fahrzeug umgesetzt werden. Nur im Ausnahmefall kann die Leitwarte das Fahrzeug per Teleoperation aus dem Stand übernehmen und direkt manövrieren. Weiterhin besteht auf verschiedenen funktionalen Ebenen eine Anbindung an eine Cloud, welche die Fahrzeugautomatisierung mit zusätzlichen Informationen unterstützen kann. Jedes automatisierte Fahrzeug wird im Straßenverkehr zunächst auch ohne Unterstützung aus einer Cloud oder der Infrastruktur mit einer Mindestfunktionalität operabel sein müssen, da eine sichere Verbindung aktuell nicht permanent sichergestellt werden kann. Auch in UNICARagil liefert die Cloud daher primär ergänzende Informationen, die durch die Fahrzeugautomatisierung genutzt werden können, um die Fahrt sicherer, effizienter und komfortabler zu gestalten. Die Cloud spiegelt dabei im übertragenen Sinne den Perzeptionsast der Fahrzeugautomatisierung. Auch außerhalb eines Fahrzeugs können Informationen mit Sensoren – entweder stationär in vernetzter Infrastruktur oder instationär in Form von UAVs oder anderer Fahrzeuge – gewonnen werden und nach einer Verarbeitung durch die Cloud den Fahrzeugen zur Verfügung gestellt werden. In der Cloud können Daten mittels Methoden der kollektiven Wahrnehmung zu einem kollektiven Umfeldmodell kombiniert werden. Im Vergleich zum Fahrzeug können hier jedoch deutlich mehr Informationen gleichzeitig verarbeitet werden, da diese aus verschiedensten Quellen stammen können und mehr Rechenleistung zu Verfügung steht. Die Cloud bietet außerdem das Potential, große Datenmengen zu sammeln und somit aus den Daten vieler Verkehrsteilnehmer zu

lernen. Das in einem kollektiven Gedächtnis vorliegende Flottenwissen kann so genutzt werden, um Funktionalitäten der Fahrzeuge off-board zu optimieren und Funktionsupdates zu generieren, oder um live Vorschläge zur optimalen Trajektorienplanung geben zu können [11]. Die Update- und Upgradefähigkeit von funktionalen Bausteinen wird dabei in zukünftigen Fahrzeugsystemen eine zentrale Rolle einnehmen. Durch die stärkere Fokussierung auf Software und die daraus resultierenden kürzeren Entwicklungszyklen, stehen funktionale und vor allem sicherheitsbezogene Updates bereits während der Nutzungszeit und vermutlich in regelmäßigen Zyklen zur Verfügung. Entgegen bisheriger Produktlebenszyklen, in denen mit dem SOP („Start of Production“) ein fertiges Produkt verkauft wird, welches nur in wenigen Ausnahmefällen aktualisiert wird, werden automatisierte Fahrzeuge wie Software auf einem Smartphone oder einem PC aktualisiert werden müssen. Bisherige starre Softwarearchitekturen im Kraftfahrzeug erlauben dies nur bedingt. Aus diesem Grund wurde auch die Softwarearchitektur in UNICARagil neu betrachtet. Ziel ist dabei die Etablierung einer – in der Informatik bereits weit verbreiteten – dienstorientierten Architektur, welche die eigentliche Funktion in einzelne Software-Dienste unterteilt: Die *Automotive Service Oriented Software Architecture* (ASOA). Ein Software-Dienst stellt dabei die kleinste Einheit dar, welche unter Einhaltung gegebener Anforderungen eine Menge von Garantien erfüllt [12]. Bild 3 zeigt die Verbindung zwischen Services/Diensten in abstrahierter Form.

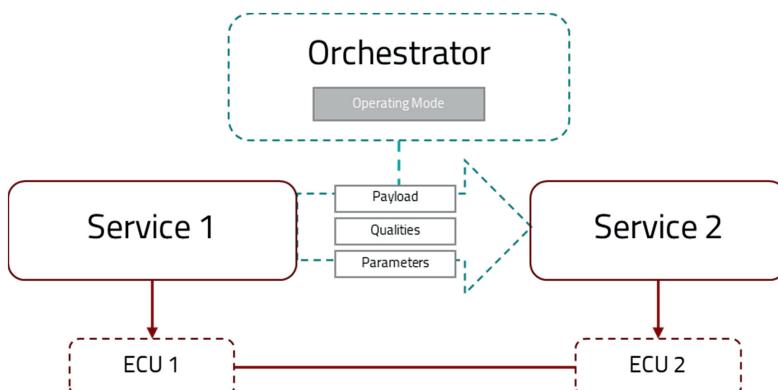


Bild 3: Schematische und vereinfachte Darstellung der Interaktion innerhalb der ASOA

Im Vergleich zu klassischen Architekturen werden Dienste der ASOA nicht bei der Auslegung des Systems starr auf Steuergeräten implementiert, sondern werden während der Laufzeit des

Systems orchestriert und bleiben so (re-)konfigurierbar. Wichtige Definitionen und eine genaue Systembeschreibung geben Kampmann et al. in [12].

4. Auswirkungen auf die Elektrik und Elektronik

Durch die Einführung einer strikten diensteorientierten Softwarearchitektur ergeben sich weitere Freiheiten in der Architektur des elektrischen und elektronischen Bordnetzes. Auch in der Industrie ist ein Wandel von klassischen Architekturen mit vielen starr verdrahteten eingebetteten Steuergeräten hin zu Domänen-zentrierten Architekturen zu erkennen [13]. In UNICARAgil wurde eine darüber hinausgehende zentralisierte Elektronikarchitektur entwickelt (vgl. Bild 4), welche sich am derzeit performantesten funktionalen System im Straßenverkehr orientiert, dem Menschen.

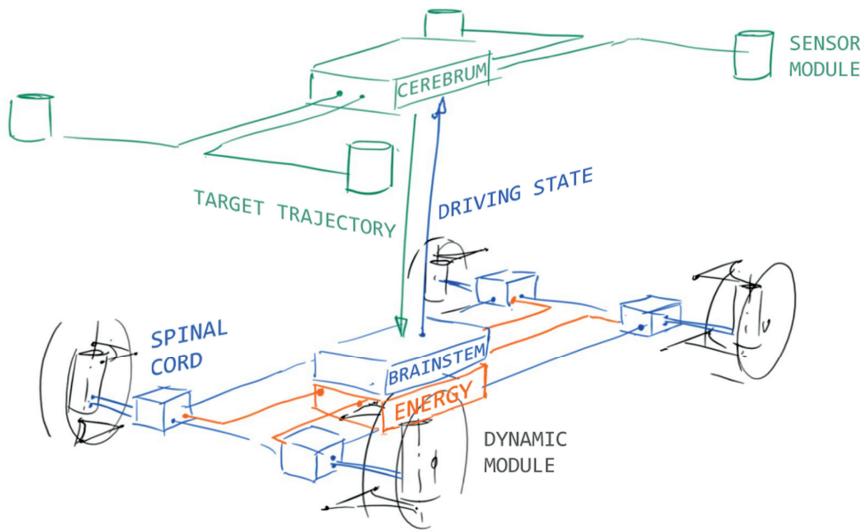


Bild 4: Vereinfachte schematische Darstellung der UNICARAgil E/E-Architektur

Die Architektur wurde dabei so gestaltet, dass diese sowohl in einem Forschungsfahrzeug, an dem zukünftig sowohl Wissenschaftlerinnen und Wissenschaftler als auch Studierende arbeiten werden und daher auf offene Hardwareplattformen wie x64-Linux-Rechner angewiesen sind, als auch für eine Industrialisierung verwendet werden kann. Dabei werden auf Sensormodul- und Großhirn-Ebene (engl. „Cerebrum“) Sensorrohdaten verarbeitet und fusioniert. Die

Fahrzeugregelung agiert vom Großhirn weitgehend unabhängig im Stammhirn (engl. „Brains-tem“), einem redundant ausgeführten sicherheitsklassifizierbaren und echtzeitfähigen Steuergerät. Die Ansteuerung der Aktorik – in diesem Falle der vier individuell ansteuerbaren Dynamikmodule – wird nachgelagert durch das Rückenmark (engl. „Spinal Cord“) realisiert. Vier speziell entwickelte eingebettete Steuergeräte arbeiten hierfür in einem redundanten Verbund, um die korrekte Ansteuerung der Aktorik zu jedem Zeitpunkt zu realisieren.

Die diensteorientierte Softwarearchitektur ASOA ist dabei auf allen verschiedenen Hardwareplattformen vom „Standard“ x64-PC bis zum Aurix-Mikrocontroller im Rückenmark durchgängig eingesetzt und lauffähig. Sie stellt damit eine der wenigen DDS („*Data Distribution Service*“) und RTPS („*Real-Time Publish-Subscribe*“) Implementierungen für eingebettete Systeme dar [14]. Die Open-Source-Entwicklungsgemeinschaft von ROS2 unterstützt diese Implementierung aktuell experimentell.

5. Entwicklung neuartiger Fahrzeugkonzepte unter Anwendung dieser Architekturen

Die zuvor beschriebenen Architektsichten wurden in UNICARagil konzeptioniert, entwickelt und umgesetzt. Dafür wurden vier Forschungsfahrzeuge aufgebaut, die jeweils auf einer modular in Länge skalierbaren Fahrplattform mit entsprechend in Höhe skalierbaren Aufbaumodulen aufsetzen. Zwei Fahrzeuge werden dabei mit kürzerem Radstand und niedriger Höhe aufgebaut, zwei mit verlängertem Radstand und vergrößerter Höhe. Die im Projekt adressierten Anwendungsfälle zeigen dabei nur einen Bruchteil der möglichen Use-Cases auf und stehen sinnbildlich für alle denkbaren Konzepte von automatisierten und fahrerlosen Fahrzeugen. Bild 5 zeigt die vier Anwendungsfälle im Entwicklungsstadium.



Bild 5: Die UNICARagil Fahrzeuge im Designstadium

Das autoSHUTTLE und autoCARGO bieten dabei jeweils eine Türöffnung von bis zu 2 Metern, um einen effizienten Ein- und Ausstieg sowie eine Containerbeladung zu ermöglichen. Sie sind so konzeptioniert, dass Fahrgäste angenehm aufrecht stehen können. autoELF und autoTAXI weisen aufgrund der niedrigen Einstiegshöhe eine erweiterte Türöffnung auf, um einen ergonomischen Einstieg für die bis zu vier Fahrgäste zu ermöglichen.

Die Fahrzeuge wurden von Grund auf neu entwickelt und sind konsequent ohne den limitierenden Auslegungsfaktor eines Fahrerarbeitsplatzes gestaltet worden. Sie sind daher vollständig auf eine fahrerlose Fahrt ausgelegt, können aber im Testbetrieb oder zum Manövrieren mittels eines zusätzlich anzubringenden Sidesticks gesteuert werden. Die Ansteuerung erfolgt dabei vollständig by-wire, es besteht also kein direkter mechanischer Durchgriff auf Lenkung oder Bremse. Die Fahrzeuggeometrie wurde aus handelsüblichen Aluminiumprofilen entworfen und virtuell abgesichert. In einem iterativen Prozess wurde die geometrische Gestaltung mit den technischen Randbedingungen abgeglichen. So besteht beispielsweise ein direkter Zusammenhang des maximalen Fahrzeuggewichts mit den zulässigen Lenkmomenten der Dynamikmodule und die hohen Lenkwinkel von bis zu 90° haben Auswirkungen auf den verfügbaren Bauraum im Fahrzeuginneren. Weiterhin beeinflusste die Positionierung der Sensormodule, welche die Umgebung optimal wahrnehmen können sollen, das Design. In der gewählten externen Positionierung der Sensorik kann ein Sensormodul – jeweils ausgestattet mit

vier Kameras, davon ein Stereo-Kamerapaar, zwei Radarsensoren und einem 360° Laserscanner – ca. 270° Sichtbereich pro Fahrzeugecke abdecken. Die vier Sensormodule decken somit das Fahrzeugumfeld von 360° redundant ab. Bild 6 zeigt die Entstehung der Plattform inkl. Aufbaumodule und Sensormodul.



Bild 6: Einblick in die Entstehung der Fahrzeuggeometrie am Beispiel des autoTAXI

In den nachfolgenden Schritten wurden die verschiedenen Bordnetze – elektrisch, elektronisch, thermisch – integriert und in Betrieb genommen. Außerdem wurden die Dynamikmodule montiert und vermessen. Schlussendlich konnten alle relevanten Steuergeräte integriert und erstmalig in Betrieb genommen werden. Bild 7 zeigt die autoELF bei ersten manuellen Testfahrten mit verschiedenen radindividuellen Lenkwinkeln, wobei ein spezieller abgesicherter Erprobungsfahrerarbeitsplatz zum Einsatz kommt, der nach Erprobung wieder entfernt wird. Es wird deutlich, dass durch die konsequente Verwendung der Dynamikmodule neuartige Bewegungsformen möglich werden.



Bild 7: links: autoELF bei ersten manuellen Fahrtests mit radindividuellen Lenkwinkeln; rechts: Inbetriebnahme der Umgebungswahrnehmung.

Im weiteren Verlauf wurden erste Wahrnehmungsfunktionen sowie die Fahrzeugregelung in Betrieb genommen und getestet. Insgesamt wurden bis zum Zeitpunkt des Verfassens dieses

Manuskripts (April 2022) alle vier Fahrzeuge erfolgreich aufgebaut und manuell in Betrieb genommen. Die Automatisierung der Fahrzeuge, die Vernetzung mit Leitwarte und Cloud sowie die Integration der Fahrzeuginnenräume werden in der verbleibenden Projektlaufzeit vorgenommen. Bild 8 zeigt abschließend drei der vier Fahrzeuge an einem Testtag an der TU Darmstadt.



Bild 8: autoTAXI, autoELF und autoSHUTTLE mit demontierten Front- und Heckschürzen zum Testen an der Technischen Universität Darmstadt. © Patrick Pintscher

6. Zusammenfassung

Automatisierte und vernetzte Fahrzeuge tragen dazu bei, die Mobilität der Zukunft nachhaltig, effizient, komfortabel und sicher zu gestalten. Die den Fahrzeugen zugrunde liegenden Konzepte werden sich jedoch stark von denen heutiger Fahrzeuge unterscheiden. Der Wegfall der Komponenten zur manuellen Fahrzeugsteuerung in ihrer heutigen Form bietet neuen gestalterischen Freiraum in der geometrischen Architektur. Dafür werden jedoch die Sichtbereiche der Sensoren ein wichtiger Gestaltungsparameter werden. Der zunehmende Einzug von agil entwickelten Softwarebausteinen ins Kraftfahrzeug unterliegt deutlich kürzeren Entwicklungszyklen als sie bisher in der Kraftfahrzeugentwicklung gelebt werden. Außerdem steigt der Bedarf an Rechenleistung mit zunehmender Intelligenz im Fahrzeug an. Nicht zuletzt muss Software daher in kürzeren Zyklen aktualisiert werden. Die dafür notwendige Änderung bzw. Neugestaltung der verschiedenen Architektursichten in einem automatisierten und vernetzten Fahrzeug wurden und werden im Projekt UNICARagil konzeptioniert, erforscht und prototypisch umgesetzt. Die vier entstandenen Fahrzeuge wurden konsequent ohne Einbezug proprietärer Architekturen und strikt auf einen fahrerlosen Betrieb ausgelegt. Aktuell befinden sich

die Fahrzeuge in der Erprobungsphase, in der schrittweise Automatisierungsfunktionalitäten in das Fahrzeug Einzug erhalten.

7. Literatur

- [1] Pfeil, F. (2018): Megatrends und die dritte Revolution der Automobilindustrie: Eine Analyse der Transformation der automobilen Wertschöpfung auf Basis des Diamantmodells.- Research Papers on Marketing Strategy 13/2018, ISBN 978-3-00-059102-0, Würzburg.
- [2] SAE International, „On-Road Automated Vehicles Standards Committee“ (2018): „Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles“ Dokument Nr. J3016_201806 Revision Juni 2018, Kurzdarstellung im Internet veröffentlicht unter:
https://www.sae.org/works/documentHome.do?comtID=TE-VAVS&docID=J3016_201806&inputPage=dOcDeTalIS
- [3] Der neue Audi A8 – hochautomatisiertes Fahren auf Level 3, Pressemitteilung, Audi 2017, online verfügbar unter: <https://www.audi-mediacenter.com/de/per-autopilot-rich-tung-zukunft-die-audi-vision-vom-autonomen-fahren-9305/der-neue-audi-a8-hochauto-matisiertes-fahren-auf-level-3-9307>
- [4] UN-R157: Uniform provisions concerning the approval of vehicles with regard to Automated Lane Keeping Systems
- [5] Erste international gültige Systemgenehmigung für hochautomatisiertes Fahren, Pressemitteilung, Mercedes-Benz, 2021, online verfügbar unter: <https://group.mercedes-benz.com/innovation/produktinnovation/autonomes-fahren/systemgenehmigung-fuer-hochautomatisiertes-fahren.html>
- [6] Was die BMW-Kooperation mit Qualcomm bedeutet, Presseartikel, Golem.de, 2022, online verfügbar unter: <https://www.golem.de/news/autonomes-fahren-was-die-bmw-kooperation-mit-qualcomm-bedeutet-2203-163901.html>
- [7] Bundesministerium für Verkehr und digitale Infrastruktur (2021): Entwurf eines Gesetzes zur Änderung des Straßenverkehrsgesetzes und des Pflichtversicherungsgesetzes–Gesetz zum autonomen Fahren, Bearbeitungsstand: 08.02.2021, online Verfügbar unter: <https://www.bmvi.de/SharedDocs/DE/Artikel/DG/gesetz-zum-autonomen-fahren.html>
- [8] Woopen, T. (2018): UNICARagil - Disruptive Modular Architectures for Agile, Automated Vehicle Concepts. - 27. Aachener Kolloquium Fahrzeug- und Motorentechnik, Aachen, 2018

- [9] Donges, E. (1982): Aspekte der Aktiven Sicherheit bei der Führung von Personenkraftwagen.- Automobil-Industrie 27, S.183–190
- [10] Ackermann, S. (2020): Systemarchitektur und Fahrmanöver zum sicheren Anhalten modularer automatisierter Fahrzeuge. - 13. Uni-DAS e. V. Workshop FAS 2020, Wallting
- [11] Lampe, B. (2019): Collective Driving - Cloud Services for Automated Vehicles in UNICARagil. - 28. Aachener Kolloquium 2019, Aachen
- [12] Kampmann, A. (2019): A Dynamic Service-Oriented Software Architecture for Highly Automated Vehicles. - IEEE ITSC 2019, Auckland, New Zealand
- [13] Reger, L. (2017): The EE Architecture for Autonomous Driving A Domain-based Approach. - ATZ Elektronik worldwide 06 | 2017
- [14] Kampmann A. (2019): A Portable Implementation of the Real-Time Publish-Subscribe Protocol for Microcontrollers in Distributed Robotic Applications. - IEEE ITSC 2019, Auckland, New Zealand

Ein industrieübergreifender Überblick von fehlertoleranten Ansätzen in autonomen Systemen

Cross-Industry Overview of Fault-Tolerant Approaches used in Autonomous Systems

M.Sc. **T. F. Horeis**, Dr.-Ing. **F. Plinke**,

Institut für Qualitäts- und Zuverlässigkeitssmanagement GmbH, Hamburg;

Dipl.-Ing. **T. Kain**, B.Sc, Ao.Univ.-Prof. Dr.techn. **H. Tompits**,

Technische Universität Wien, Österreich;

M.Sc. **R. C. Rinaldo**, DFKI, Bremen;

M.Sc. **J. Heinrich**,

Institut für Qualitäts- und Zuverlässigkeitssmanagement GmbH, Wuppertal

Kurzfassung

In zahlreichen Industrien kann ein stetiger Anstieg automatisierter und autonomer Systeme bzw. deren Entwicklung beobachtet werden. Beispielsweise wird im Kernkraftsektor als auch im Bahnsektor die Adaption und Integration autonomer Systeme vorangetrieben. Dabei sind viele Aufgaben die von einem autonomen System ausgeführt werden sicherheitskritisch. Zu nennen sind hier z.B. das Überwachen der Reaktorkerne eines Kernkraftwerks oder der Betrieb von autonomen U-Bahnen. Aufgrund der hohen Sicherheitskritikalität der Aufgaben müssen dementsprechend Maßnahmen implementiert werden, um einen zuverlässigen und sicheren Betrieb bzw. eine zuverlässige und sichere Durchführung der Aufgaben zu gewährleisten. Die implementierten Maßnahmen müssen sicherstellen, dass zufällige und systematische auftretende Fehler behandelt werden können. Diese Maßnahmen, welche die Konsequenzen solcher Fehler limitieren und entsprechend behandeln werden als „fehlertolerante Ansätze“ („fault-tolerant approaches“) bezeichnet. Dabei werden unterschiedliche solcher fehlertoleranten Ansätze innerhalb der verschiedenen Industrien verwendet um den anwendungsspezifischen Grad an Zuverlässigkeit und Sicherheit zu erreichen. Nichtdestotrotz können trotz unterschiedlicher Automatisierungsgrade, Eigenschaften und Herausforderungen innerhalb der unterschiedlichen Anwendungen einige Gemeinsamkeiten identifiziert werden. Im Rahmen dieser Arbeit wird eine Literaturrecherche bzgl. fehlertoleranter Ansätze autonomer Systeme in der Luft-, Raum- und Bahnfahrt, sowie der Automobil- und Kernkraftbranche durchgeführt. Ziel

ist es dabei, den Wissenstransfer zwischen den einzelnen Industrien hinsichtlich fehlertoleranter Ansätze zu stärken und zu ermöglichen.

Abstract

A constant increase in autonomous systems can be found across various industries. For instance, the nuclear power industry or the railway industry have already started adopting autonomous systems. Many tasks performed by these systems are safety critical, e.g., monitoring reactor cores of a nuclear power plant or operating automated subway lanes. Consequently, it is essential to implement measures that guarantee a reliable and safe operation of these systems. Therefore, it is required to handle occurring faults. Measures that aim to limit the consequences of faults are referred to as *fault-tolerant approaches*. Different industries apply different fault-tolerant approaches to achieve an acceptable level of reliability and safety for their automated systems. Even though deviations regarding the level of automation, characteristics, and challenges between the different fields exist, some commonalities can be identified. In this paper, we perform a literature study of fault-tolerant approaches for autonomous systems which are used in the aviation, aerospace, automotive, railway, and nuclear power industry. In doing so, we aim to encourage the knowledge transfer concerning fault-tolerance approaches between these industries.

1. Introduction

According to the McKinsey Center for Future Mobility [45], a total number of 9100 patents for autonomous vehicles sensors and ADAS (Advanced Driver Assistance System) components have been published in the automotive industry between the years 2010 and 2019. For none of the other examined categories, including battery technology and infotainment, such a high number of patents was filed. Besides the automotive branch, also other industries are thriving towards more automation. For example, two autonomous railway systems started their operation in 2021. The Chinese city Ningbo introduced a fully autonomous subway lane, and in the city of Dubai, a fully automated railway system went into operation [30]. This emphasizes the rapid and increasing development of autonomous systems that can be seen across various industries.

However, the introduction of autonomous systems also induces various challenges. A main challenge is maintaining the dependability, i.e., the reliability, availability, and safety, of autonomous systems that perform safety-critical tasks, e.g., driving a vehicle and flying an airplane

or spacecraft, as faults cannot be handled by a human operator [40]. Consequently, fault-tolerant approaches are required, which aim to limit the consequences of occurring faults.

In this paper, we present a survey of fault-tolerant approaches for autonomous systems, which are used in several different industries. In particular, we study the following four transportation industries: the aviation industry, the aerospace industry, the automotive industry, and the railway industry. In recent years the level of automation in these industries has been rising, and consequently also the need for fault-tolerant approaches increased. Additionally, we analyze fault-tolerant approaches applied in the nuclear power industry, as the safety and reliability requirements of nuclear power plants are particularly high.

Our study shows that all these industries require fault-tolerant approaches. However, as the individual industries differ in their characteristics and the resulting challenges, different fault-tolerance approaches are required. For instance, the complexity of the autonomous systems, the environment of the systems, the potential number of individuals affected by a system failure, and economic factors influence the applied fault-tolerance concepts. Our aim is that the present paper inspires the exchange of knowledge between the discussed industries.

The paper is organized as follows: Section 2 illustrates why fault-tolerant approaches are required to maintain the dependability of autonomous systems. Section 3 presents fault-tolerant approaches for autonomous systems used in the aviation, aerospace, automotive, railway, and nuclear power industry. The paper concludes with Section 4, containing a brief summary and outlook.

2. Dependable Autonomous Systems

Autonomous systems have to be *dependable* as they are controlled by computer systems and not by human operators [40]. Dependability is defined as the capability of a system to avoid failures that are more frequent, more severe, or the resulting outages are longer than is acceptable for the users of such a system [2].

2.1 Characteristics of Dependability

According to Aviziens et al. [2], dependability comprises several attributes, including *reliability*, *availability*, and *safety*, which are defined as follows:

- reliability is the ability of a system to continue performing a designed function;
- availability is the ability of a system to be ready to perform a designed function when requested; and
- safety is the absence of catastrophic events that endanger the integrity of the users and the environment of the system.

These parameters are threatened by *failures*, *errors*, and *faults*, which can be defined as follows [2, 29, 32, 46]: A failure is the inability of a system to perform the designed function. Failures are caused by errors, which are defined as discrepancies between the measured, computed, or received value and the intended value. Errors, in turn, are consequences of faults which are defects of the system. Faults can be grouped into various categories including, for instance, hardware faults (e.g., cross talk, short circuit, defect memory section), software faults (e.g., use incorrect data type, division by zero, wrong interpretation of values), and manufacturing faults (e.g., incorrect mounting, damaged wire, miscalibrated sensor). Note that a fault does not necessarily have to lead to an error. For instance, a defective section in a memory might never be used. Consequently, this fault has no impact.

2.2 Measures to Maintain Dependability

Aviziens et al. [2] define the following four means to maintain dependability: *fault prevention*, *fault removal*, *fault forecasting*, and *fault tolerance*. The goal of fault prevention is to avoid introducing faults in the first place. For instance, programming languages that support strong typing can prevent functions from returning wrong data types. Fault removal methods aim to eliminate faults during development and use. Both fault prevention and fault removal are categorized as *fault-avoidance measures*, i.e., methods that avoid the occurrence of faults. On the other hand, fault tolerance and fault forecasting are classified as means that accept the occurrence of faults and are, therefore, referred to as *fault-acceptance means*. The goal of fault forecasting is to predict future faults and their consequences. Fault-tolerance means aim to prevent failures in case faults occur.

As autonomous systems are complex systems consisting of various hardware and software components, numerous faults may occur. Therefore, we must assume that fault prevention and fault-removal means cannot eliminate all faults [52]. Faults that fault-avoidance measures have not addressed have to be handled by fault-acceptance means. Fault-forecasting measures can eliminate predictable faults. However, unpredictable faults that have not been eliminated by fault-avoidance means, e.g., undetected software bugs, are not handled by fault-

forecasting measures. Only fault-tolerance strategies can handle such faults. Consequently, fault-tolerance measures have to be implemented in autonomous systems to avoid that faults that are neither handled by fault prevention, fault removal, nor fault forecasting measures do not cause failures.

A fault-tolerant strategy can improve the reliability, availability, and safety of a system:

- Reliability: A fault-tolerant system may handle occurring faults so that they do not cause failures. Hence, the system can continue performing the designed function.
- Availability: A fault-tolerant system may handle faults that occur before the operation of the system. Thus, the system is ready to perform a designed function when requested.
- Safety: A fault-tolerant system may handle occurring faults so that they do not cause faults that in return cause catastrophic events that endanger the integrity of the users and the environment of the system.

The foundation of a fault-tolerant system is *redundancy* [46], i.e., auxiliary components are introduced that perform the same or a similar function as the existing components [29]. Redundancy can be defined on different levels of the system, including, for example, on the software and hardware level [10].

Besides the introduction of redundancies into a system, fault-tolerance strategies require further elements including, for instance, *detection*, *diagnosis*, *isolation*, and *reconfiguration* measures [2, 27]. Detection means are required to identify errors and hence initiate the subsequent steps of the fault tolerance strategy. Diagnosis measures aim to identify the cause of an error. Isolation actions are performed to prevent the further propagation of an error. Finally, reconfiguration procedures ensure that a correct functioning component takes over the task of the faulty component.

3. Fault-Tolerant Approaches for Autonomous Systems in Different Industries

In several industries, a drive towards autonomous systems can be observed. Many of these autonomous systems execute safety-critical tasks. Consequently, fault-tolerant approaches have to be implemented to maintain a safe and reliable operation.

In what follows, we examine fault-tolerant approaches of autonomous systems used in the aviation, aerospace, automotive, railway, and nuclear power industry.

3.1 Aviation Industry

Since the tremendous growth of commercial aviation following the years after the end of the Second World War, the distance covered by air travelers increased from 8 billion kilometers in 1952 [60] to over 8,000 billion kilometers in 2018 [31]. Already from the early beginnings of the aviation industry onwards, the automation of various functions has been a key factor in increasing aircraft safety and reliability. For instance, in 1912, the first autopilot was designed by the Sperry Corporation [42]. This autopilot was able to keep the aircraft straight and steady without any interventions from the pilot. In 1947, the first entirely automatic flight across the Atlantic Ocean was performed by an application programmed on punched cards [16]. Shortly after, autopilots became a standard aircraft functionality. Since then, various improvements to increase or support the automation of aircraft have been developed. For instance, additional components such as automatic navigation and automatic tracking capabilities have been added to support the pilot [16].

The main incentive for automation in aviation is the reduction of human errors to increase safety [14, 50]. Further benefits of automation in aviation are the decrease of operational costs, workload reduction, and an improved job satisfaction. Therefore, a continuous drive towards a higher level of automation in the aviation industry can be perceived. The implemented automation approaches have to fulfill a high level of safety and reliability as faults can lead to a large number of fatalities [47].

To improve the safety of airplanes, fault-tolerant systems based on duo-duplex, triple-triple, or quadruplex architectures are employed [13, 23]. A duo-duplex architecture, as illustrated in Figure 1, consists of two lanes, whereby each lane includes two identical computing nodes, which differ from the computing nodes of the other lane. Each of the four computing nodes executes a distinct software implementation. Although the implementations are not identical, they provide the same functionality. A voting mechanism executed by each lane checks whether the output of the two implementations is identical. In case diverging outputs are detected, the voter instructs the switch to use the output determined by the redundant lane. Consequently, duo-duplex systems, which are, for example, found in airplanes manufactured by Airbus [57], can tolerate one failure.

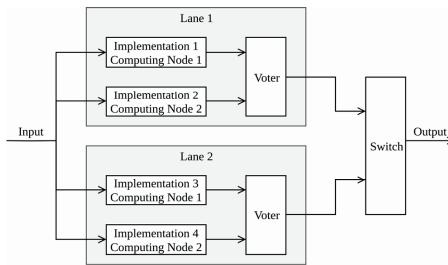


Fig. 1: Duo-duplex architecture.

An architecture that can tolerate more failures is the triple-triple redundancy architecture, which is used, for example, by Boeing [63]. A triple-triple system, as shown in Figure 2, consists of three homogeneous lanes, whereby each lane includes three distinct computing nodes, which execute distinct software implementations. Furthermore, each lane implements a 2oo3 (“two out of three”) voting logic, which compares the output of the three distinct implementations. Only if all three outputs diverge, the voter instructs the switch to use the output of another lane. Otherwise, the dominating output value is considered to be valid. Hence, triple-triple can tolerate two-lane failures as well as one-implementation failures of the remaining lane.

Also, the quadruplex architecture can tolerate up to two lane failures [11]. Systems based on this architecture, as illustrated in Figure 3, implement four homogeneous lanes, whereby each lane consists of a voter, a switch, as well as a computing node, which executes the software implementation.

The voters implement a 3oo4 (“three out of four”) voting logic, i.e., if three of the four lanes provide the same output, this output is considered correct. After the first fault, the lane which provided the faulty output will be isolated, and the voting logic degrades to 2oo3. Therefore, another fault can be tolerated.

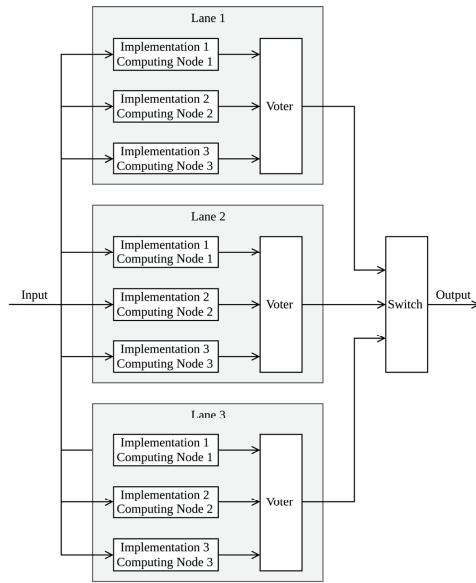


Fig. 2: Triple-triple architecture.

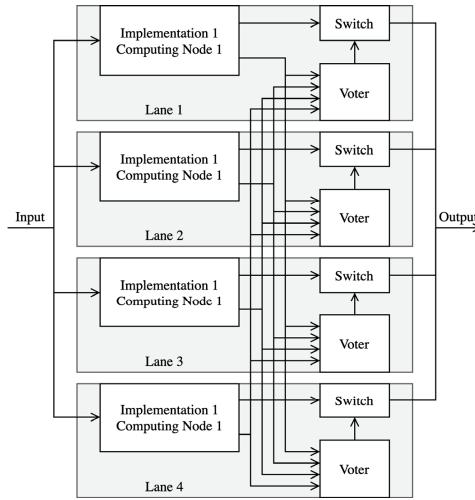


Fig. 3: Quadruplex architecture.

3.2 Aerospace Industry

Since the launch of the first artificial satellite, Sputnik 1, in 1957 [55], about 12,070 satellites have been placed in the Earth orbit [17]. Recently, an increase in the launch of commercial missions can be observed, making spacecraft engineering a topic of growing interest also for non-research organizations.

In general, space missions require high levels of automation. These high automation levels are, for instance, a consequence of the fact that space is inhabitable and that physical contact with the spacecraft becomes almost impossible after the launch. For example, the Rosetta asteroid lander *Philae* [9] traveled ten years in space to reach its destination and perform a complex, pre-programmed landing maneuver, followed by another two years of operation. However, also in manned space missions, the demand for automated and autonomous systems is increasing, ranging from general tasks such as controlling the water quality [24] and managing errors to robots performing dedicated tasks such as scientific experiments [38] as well as performing repairs and maintenance tasks [8].

In order to ensure the safety and reliability of space missions, several standards and guidelines have been published by the European Cooperation for Space Standardization (ECSS). For instance, the ECSS-Q-ST-40C Rev.1 standard [18] defines fault tolerance to be the basic safety requirement to control hazards. Therefore, the system design shall include an on-board redundancy management for safety-critical functions and mechanisms for fault detection, fault isolation, and switching to redundant units, which are also referred to as FDIR approaches. Furthermore, it is required to provide the ground station with necessary information concerning fault detection, isolation, tolerance, and the current redundancy status.

Consequently, FDIR concepts are essential to increase safety, reliability, and availability. The computational complexity of the applied FDIR systems is, in general, rather limited due to the limited computational power of the spacecraft [49]. Therefore, advanced methods such as AI-supported FDIR methods are hardly applicable, especially in smaller, low-power satellites. However, research regarding AI-supported FDIR approaches has increased in the last few years. For instance, Jaekel and Scholz [34] present an AI-supported FDIR architecture for future spacecrafts.

In classical spacecrafts (we focus on satellites here), two main strategies for model-based FDIR methods are applied according to Olive [49]: the *half-satellite FDIR strategy* and the *hierarchical FDIR strategy*.

The half-satellite FDIR strategy defines that the satellite is fully reconfigured after the fault detection, and all units are switched to redundant ones without performing any fault isolation.

Afterward, the ground station has to identify the cause of the anomaly and correct it. On the other hand, the hierarchical FDIR strategy specifies that the fault is recovered on the lowest possible layer after the fault is detected and isolated. In order to permit a graduated reaction, four levels of faults, which are illustrated in Figure 4, are distinguished.

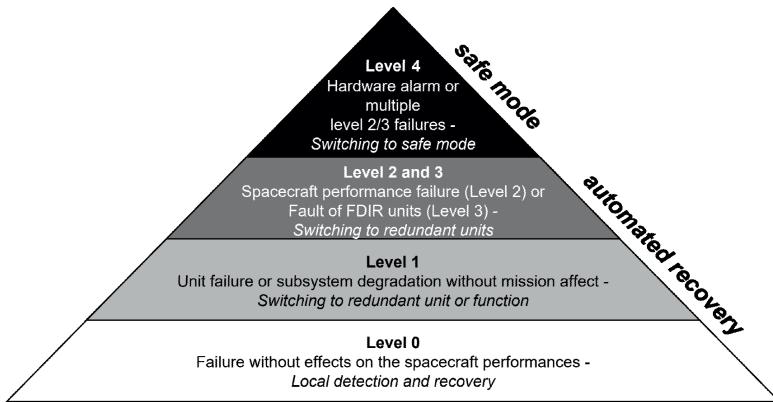


Fig. 4: The levels of the hierarchical FDIR strategy.

Level 0 comprises faults that have no impact on the performance of the satellite, e.g., a single bit-flip. The detection and automated recovery are performed locally in the affected unit.

Level 1 concerns faults requiring switching to a redundant unit, leading to a temporary degraded mode that does not affect the mission goal. In this case, the detection must be performed outside the affected unit, and the concerned subsystem performs the recovery.

Levels 2 and 3 share the same kind of detection and recovery scheme, whereas Level 2 strictly concerns faults that could not have been covered by previous levels, and Level 3 concerns faults of the FDIR units. Faults of Level 2 and 3 are recovered by switching to redundant modules.

Level 4 is the most critical level, which concerns multiple faults on Level 2 and 3 or hardware alarms. Such faults are a consequence of a critical breakdown. The recovery of Level 4 faults is performed by the ground and requires an interruption of the mission.

As redundancy is crucial for a successful FDIR process, choosing the right amount of redundancy is essential. Suich and Patterson [56] discuss considerations between financial costs and reliability benefits regarding redundancy. Although redundancy is indispensable, it can

also decrease the safety of the spacecraft by, for example, adding additional faults or masking design flaws [48].

3.3 Automotive Industry

According to the McKinsey Center for Future Mobility [45], autonomous vehicles will cover more than 60 percent of the kilometers passengers will travel in 2040. Furthermore, it is estimated that 2 trillion dollars of market revenue will be generated from mobility services and sales of autonomous vehicles by that year. These numbers show that the automotive market is thriving towards autonomous vehicles in the future [37].

The introduction of autonomous vehicles can induce an improved traffic flow [20], fewer emissions [44], and provide mobility to the elderly or disabled individuals [19]. However, the aim of automating the task of driving also leads to challenges. Besides social [12] and legal hurdles [3], a significant technical challenge is the development of economic software and hardware concepts that satisfy the high safety and reliability requirements [62]. This challenge is the subject of many research activities. For instance, the project SafeAdapt (Safe Adaptive Software for Fully Electric Vehicles) proposes a cost-effective fault-tolerant electrical and electronic architecture (so-called *E/E architecture*) based on adaptation [54]. The adaptive behavior, which is distributed over all subcomponent cores of the system, is implemented by the so-called *safe adaption platform core* (SAPC). The main features of it include:

- the deactivation and isolation of failed system components to prevent error propagation,
- the activation of applications,
- the reconfiguration of the distribution of resources, and
- the reallocation of applications to different computational resources.

These features can be used to implement various mechanisms to improve the reliability and safety of the system. For instance, advanced strategies like graceful degradation can be implemented to maintain a limited operation of the vehicle in case of occurring faults.

To implement a graceful degradation behavior, information about the priorities of the application is provided to SAPC. This information can be used by SAPC to execute a prioritized deactivation, activation, reconfiguration, and reallocation based on a predefined policy. A policy can, e.g., specify that low prioritized applications are deactivated to free resources in order to run high prioritized applications.

Kohn et al. [36] also argue that in autonomous vehicles a system-wide fault-tolerance is required. In their work, the authors focus on fault-tolerance approaches in multi-core systems. For instance, they present a concept based on lockstep microcontroller units (MCUs). As illustrated in Figure 5a, the idea of this approach is that a function is executed on two lockstep MCUs. Therefore, each channel runs the function redundantly and thus can detect errors. Furthermore, each channel can monitor the other channel and take over the operation in case a failure is detected. Kohn et al. illustrated that this fault-tolerant concept could be used to design an autonomous driving electronic control unit (ECU) as well as an electronic power steering ECU, whereby the latter uses the two inputs of the autonomous driving ECU. Besides the lockstep MCU-based concept, Kohn et al. also present an alternative fault-tolerant approach, which is illustrated in Figure 5b. The idea of this concept is to use only one single multi-core CPU, whereby one core executes the main algorithm and the other core runs a backup algorithm. Both cores continuously monitor the other core to detect faults. In case a core fails, the other one takes over the operation.

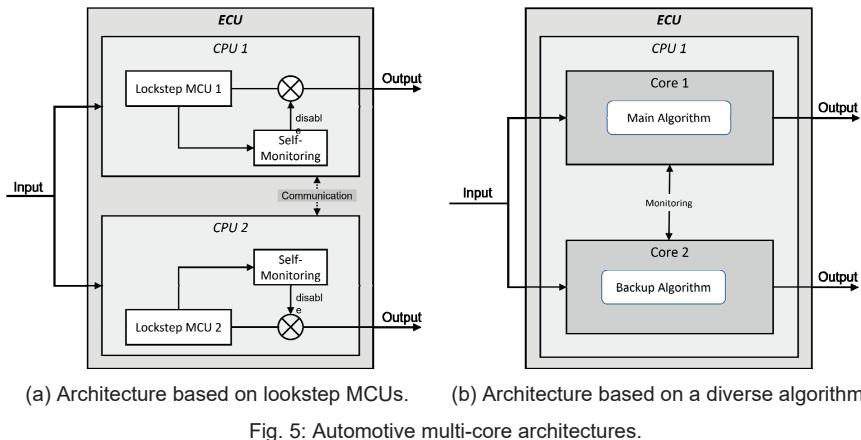


Fig. 5: Automotive multi-core architectures.

3.4 Railway Industry

Railway traffic is an essential conveyance for passengers as well as for cargo. The German railway company Deutsche Bahn, for instance, transported more than 4,500 million passengers and 250 million tons of cargo in 2018 [15]. In order to compete with other means of transportation, the railway industry is currently in a modernization phase [51]. This includes, for example, the introduction and improvement of highly automated on-board control and wayside

functions [25]. For instance, France, Korea, Singapore, and the United Arab Emirates, among others, already operate automated subway lines [59]. However, it has to be noted that the level of automation of subway lines is relatively high in comparison to non-suburban railways since the environment is relatively simple and controlled due to, e.g., the limited interference with dynamic objects, such as crossing road vehicles [1].

In the railway industry, the unavailability of a train or rail track can lead to the downtime of various trains due to their interdependence. Due to this interdependence, a single failure can affect a large number of people. As a consequence, a significant challenge of automation in the railway industry is the increased demand for availability while guaranteeing safety [1]. Consequently, new fault-tolerance approaches are required.

For example, Wang et al. [58] present a concept for an efficient and safe train-control system, i.e., a system that controls, e.g., the track intervals, the train routes, and the train speed. The proposed train control system is a distributed system consisting of subsystems that are located in trains and on the wayside, whereby the safety of the system is enhanced by applying parallel monitoring. For instance, the system that determines the maximum permitted speed of a train is implemented by a unit located on-board the train as well as by the wayside system. The train compares the two independent results, and if they are equal, the determined maximum speed is further processed. Due to the heterogeneously computed values, common-mode faults are excluded.

Furthermore, Lopez et al. [39] propose an approach to increase the reliability of a railway signaling system, i.e., a system that transmits commands and information between trains and the control center. Their approach makes use of the migration of railway signaling systems from a circuit-switched to a packet-switched network. The authors suggest implementing an approach that combines spatial and temporal redundancy in order to increase reliability. Note that old approaches, which are based on circuit-switched networks, employ spatial redundancy, whereby spatial redundancy is obtained by transmitting information through different network paths. Temporal redundancy, which the packet-switched network enables, is realized by delivering the same information multiple times with an offset in time. Sending the packets repeatedly with a time delay can improve the resilience against burst errors, i.e., a sequence of incorrect bits transmitted via a communication channel.

3.5 Nuclear Power Industry

Nuclear power plants play an important role in the global energy infrastructure. According to the International Atomic Energy Agency (IAEA), nuclear power plants contribute to about 10 percent of the total global electricity generation [28].

Operating nuclear power plants efficiently is a complex task [4]. Thus, automation approaches are applied to control and run several parts of the systems. For instance, the startup of the reactors [5] and the in-core fuel management [53] are tasks that can be optimized using automation.

Besides increasing the efficiency of nuclear power plants, another major challenge is to ensure a safe operation. Accidents in the past, such as those at the Fukushima Daiichi (2011), Chernobyl (1986), or Three Mile Island (1979) plants, have demonstrated the enormous environmental consequences of radiation leakage, which contaminate affected areas for decades and leaving them uninhabitable [6, 7].

In general, for systems that are essential for the safety of nuclear power plants, redundant backup systems are installed [61]. These systems take over the operation in case the primary system fails. For instance, as a consequence of the Fukushima disaster, all power plants in the United States were equipped, e.g., with additional emergency pumps, generators, and battery banks [21].

However, backup systems can also have unexpected and unwanted side effects. For instance, Kettunen, Reiman, and Wahlström [35] argue that the introduction of redundant systems increases the complexity of systems, which in turn can cause that faults might not get recognized and thus accumulate over time. Furthermore, common-cause failures are more likely to occur.

To handle faults in safety-critical systems as well as in their backup systems, fault detection and diagnosis (FDD) approaches can be employed [41]. FDD methods aim to detect, isolate, and identify the size and time-variant behavior of faults [33]. Application areas of FDD methods in nuclear power plants include instrument calibration monitoring [27], reactor core monitoring [43], and monitoring of loose parts that can damage the reactor coolant system [22]. Instrument calibration monitoring, for instance, can be implemented by using a redundant set of sensors [26]. The average of the measurements of the redundant sensors is used to determine whether the sensor under investigation shows an abnormal derivation. Instead of the data perceived by redundant sensors, correlating measurements in the system can be used to predict the output of a sensor. The advantage of such an analytical approach is that hardware can be reduced.

By introducing FDD methods, the safety of nuclear power plants can be improved as, for instance, the equipment reliability is enhanced, and the exposure of personal to radiation can be reduced [41]. Furthermore, also the efficiency is increased since, for example, the availability and the lifetime can be optimized.

4. Conclusion

In this paper, we presented fault-tolerant approaches for autonomous systems used in the aviation, aerospace, automotive, railway, and nuclear power industry. As can be seen from the reviewed industries, they are all driving towards a higher level of system automation. In recent years, especially in the aerospace and automotive field, there are increasing research activities concerning automation and fault-tolerance. Furthermore, our cross-industry discussion illustrates that the characteristics of the individual industries influence the applied fault-tolerance approaches. In general, we identified two different types of fault-tolerance approaches: (i) fault-tolerance approaches that rely on a dynamic behavior of the system, e.g., the FDIR approach applied in the aerospace domain, and (ii) fault-tolerance approaches that rely on several redundant components. A tendency towards either one of the solutions can be seen by the industry-specific relevance of cost and weight restrictions. For instance, a goal of the automotive industry is to save costs and weight in modern vehicles [36]. Therefore, fault-tolerance approaches which do not require extensive hardware resources are developed in this field, i.e., smart and dynamic fault-tolerance approaches are required. Other industries like the nuclear power industry have less harsh costs and weight restrictions. Consequently, the employed fault-tolerance approaches are often based on additional hardware.

However, despite the different characteristics of the individual industries, in our opinion, a cross-industry knowledge transfer is beneficial. For instance, the railway industry already applies distributed fault-tolerant systems, which are distributed among trains and the infrastructure. It is conceivable that similar approaches may be applicable in the automotive industry.

One can observe that the examined industries use a different and, to some extent, ambiguous taxonomy concerning fault-tolerant autonomous systems. Therefore, a proposal for a cross-industry taxonomy to facilitate the exchange between the individual industries would be desirable. Also, an in-depth comparison of the most commonly used fault-tolerant concepts from the individual industries is subject to future work.

References

- [1] J. Athavale, A. Baldovin, and M. Paulitsch. Trends and Functional Safety Certification Strategies for Advanced Railway Automation Systems. In *Proceedings of the IEEE International Reliability Physics Symposium (IRPS 2020)*. IEEE, 2020.
- [2] A. Avizienis, J.-C. Laprie, B. Randell, and C. Landwehr. Basic Concepts and Taxonomy of Dependable and Secure Computing. *IEEE Transactions on Dependable and Secure Computing*, 01(1):11–33, 2004.

- [3] I. Barabás, A. Todoruț, N. Cordoș, and A. Molea. Current Challenges in Autonomous Driving. In *Proceedings of the 2017 International Congress of Automotive and Transport Engineering - Mobility Engineering and Environment (CAR2017)*, 2017.
- [4] H. Basher and J. S. Neal. Autonomous Control of Nuclear Power Plants. Report, Nuclear Science and Technology Division, 2003.
- [5] R. Berkan, B. Upadhyaya, L. Tsoukalas, R. Kisner, and R. Bywater. Advanced automation concepts for large-scale systems. *IEEE Control Systems Magazine*, 11(6):4–12, 1991.
- [6] J. Bevelacqua. Applicability of health physics lessons learned from the Three Mile Island Unit 2 accident to the Fukushima Daiichi accident. *Journal of Environmental Radioactivity*, 105:6–10, 2012.
- [7] J. J. Bevelacqua. *Basic Health Physics: Problems and Solutions*, volume 1. John Wiley & Sons, 2010.
- [8] A. Beyer, G. Grunwald, M. Heumos, M. Schedl, R. Bayer, W. Bertleff, B. Brunner, R. Burger, J. Butterfaß, R. Gruber, T. Gumpert, F. Hacker, E. Krämer, M. Maier, S. Moser, J. Reill, M. A. R. Garzon, H.-J. Sedlmayr, N. Seitz, M. Stelzer, A. Stemmer, G. T. Manteiga, T. Wimmer, M. Grebenstein, C. Ott, and A. O. Albu-Schäffer. CAESAR: Space Robotics Technology for Assembly, Maintenance, and Repair. In *Proceedings of the 69th International Astronautical Congress (IAC 2018)*, 2018.
- [9] J.-P. Bibring, H. Rosenbauer, H. Boehnhardt, S. Ulamec, J. Biele, S. Espinasse, B. Feuerbacher, P. Gaudon, P. Hemmerich, P. Kletzkine, D. Moura, R. Mugnuolo, G. Nietner, B. Pätz, R. Roll, H. Scheuerle, K. Szegö, K. Wittmann, Philae Project Office, and the Entire Philae Team. The Rosetta Lander (“Philae”) Investigations. *Space Science Reviews*, 128(1):205–220, 2007.
- [10] A. Birolini. *Reliability Engineering: Theory and Practice*. Springer, 2013.
- [11] H. Blair-Smith. Space Shuttle Fault Tolerance: Analog and Digital Teamwork. In *Proceedings of the 28th IEEE/AIAA Digital Avionics Systems Conference (DASC 2009)*, pages 6.B.1–1–6.B.1–11, 2009.
- [12] J.-F. Bonnefon, A. Shariff, and I. Rahwan. The Social Dilemma of Autonomous Vehicles. *Science*, 352(6293):1573–1576, 2016.
- [13] R. Brockhaus, W. Alles, and R. Luckner. *Flugregelung*. Springer, 2011.
- [14] A. Chialastri. *Automation*, Chapter *Automation in Aviation*. IntechOpen, 2012.
- [15] Deutsche Bahn AG. Deutsche Bahn Facts & Figures 2018. https://www.deutschebahn.com/en/facts_figures-6929188, 2018. Published: 2019-03-14, Assessed: 2022-01-26.

- [16] M. Elish and T. Hwang. Praise the Machine! Punish the Human! The Contradictory History of Accountability in Automated Aviation. *SSRN Electronic Journal*, 2015.
- [17] ESA Space Debris Office. ESA's Annual Space Environment Report. Report, ESA ESOC, 2021.
- [18] European Cooperation for Space Standardization. 2017.
- [19] D. J. Fagnant and K. Kockelman. Preparing a Nation for Autonomous Vehicles: Opportunities, Barriers and Policy Recommendations. *Transportation Research Part A: Policy and Practice*, 77:167 – 181, 2015.
- [20] P. Fernandes and U. Nunes. Platooning with IVC-Enabled Autonomous Vehicles: Strategies to Mitigate Communication Delays, Improve Safety and Traffic Flow. *IEEE Transactions on Intelligent Transportation Systems*, 13(1):91–106, 2012.
- [21] K. Ferral. Redundant Backup Systems are Keeping Nuclear Plants Safe. <https://www.govtech.com/em/safety/redundant-backup-systems-keeping-nuclear-plants-safe.html>, 2014. Published: 2014-12-15, Assessed: 2022-01-28.
- [22] S. Figedy and G. Oksa. Modern Methods of Signal Processing in the Loose Part Monitoring System. *Progress in Nuclear Energy*, 46(3-4):253–267, 2005.
- [23] H. Flühr. *Avionik und Flugsicherungstechnik*. Springer, 2012.
- [24] J. Frank, D. Iverson, C. Knight, S. Narasimhan, K. Swanson, M. Scott, M. Windrem, K. Pohlkamp, J. Mauldin, K. McGuire, and H. Moses. Demonstrating Autonomous Mission Operations Onboard the International Space Station. In *Proceedings of the 2015 AIAA SPACE Conference and Exposition*, pages 1–24. American Institute of Aeronautics and Astronautics, 2015.
- [25] A. Ghosh, V. Rana, B. Johnson, and J. Profeta. A Distributed Safety-Critical System for Real-Time Train control. In *Proceedings of the 21st Annual Conference on IEEE Industrial Electronics (IECON'95)*, volume 2, pages 760–767, 1995.
- [26] H. Hashemian. *Maintenance of Process Instrumentation in Nuclear Power Plants*. Springer, 2006.
- [27] J. W. Hines and B. Rasmussen. Online Sensor Calibration Monitoring Uncertainty Estimation. *Nuclear Technology*, 151(3):281–288, 2005.
- [28] IAEA. IAEA Annual Report 2019. Report, International Atomic Energy Agency (IAEA), 2019.
- [29] Institute of Electrical and Electronics Engineers. IEEE Standard Glossary of Software Engineering Terminology. *IEEE Std 610.12-1990*, pages 1–84, 1990.
- [30] M. Intelligence. Smart Railways Market - Growth, Trends, COVID-19 Impact, and Forecasts (2021 - 2026). Technical report, 2021.

- [31] International Civil Aviation Organization. The World of Air Transport in 2018. <https://www.icao.int/annual-report-2018/Pages/the-world-of-air-transport-in-2018.aspx>, 2018. Published: 2019-03-14, Assessed: 2022-01-26.
- [32] International Organization for Standardization. ISO/IEC 2382:2015(en) *Information Technology — Vocabulary*. 2015.
- [33] R. Isermann and P. Ballé. Trends in the Application of Model-based Fault Detection and Diagnosis of Technical Processes. *Control Engineering Practice*, 5(5):709–719, 1997.
- [34] S. Jaekel and B. Scholz. Utilizing Artificial Intelligence to Achieve a Robust Architecture for Future Robotic Spacecraft. In *Proceedings of the 2015 IEEE Aerospace Conference (AeroConf 2015)*. IEEE, 2015.
- [35] J. Kettunen, T. Reiman, and B. Wahlström. Safety Management Challenges and Tensions in the European Nuclear Power Industry. *Scandinavian Journal of Management*, 23(4):424–444, 2007.
- [36] A. Kohn, M. Kasmeyer, R. Schneider, A. Roger, C. Stellwag, and A. Herkersdorf. Fail-Operational in Safety-Related Automotive Multi-Core Systems. In *Proceedings of the 10th IEEE International Symposium on Industrial Embedded Systems (SIES 2015)*, pages 144–147, 2015.
- [37] T. Litman. Autonomous Vehicle Implementation Predictions: Implications for Transport Planning. Technical report, 2019.
- [38] J. Liu, Q. Gao, Z. Liu, and Y. Li. Attitude Control for Astronaut Assisted Robot in the Space Station. *International Journal of Control, Automation and Systems*, 14(4):1082–1095, 2016.
- [39] I. Lopez, M. Aguado, D. Ugarte, A. Mendiola, and M. Higuero. Exploiting Redundancy and Path Diversity for Railway Signalling Resiliency. In *Proceedings of the 2016 IEEE International Conference on Intelligent Rail Transportation (ICIRT 2016)*, pages 432–439, 2016.
- [40] B. Lussier, A. Lampe, R. Chatila, J. Guiochet, F. Ingrand, M.-O. Killijian, and D. Powell. Fault Tolerance in Autonomous Systems: How and How Much? In *Proceedings of the 4th IARP-IEEE/RAS-EURON Joint Workshop on Technical Challenges for Dependable Robots in Human Environments (DRHE 2005)*, 2005.
- [41] J. Ma and J. Jiang. Applications of Fault Detection and Diagnosis Methods in Nuclear Power Plants: A Review. *Progress in Nuclear Energy*, 53(3):255–266, 2011.
- [42] R. Mahony, R. W. Beard, and V. Kumar. *Modeling and Control of Aerial Robots*, pages 1307–1334. Springer, 2016.

- [43] M. Mori, M. Kaino, S. Kanemoto, M. Enomoto, S. Ebata, and S. Tsunoyama. Development of advanced core noise monitoring system for BWRs. *Progress in Nuclear Energy*, 43(1):43–49, 2003.
- [44] W. R. Morrow, J. B. Greenblatt, A. Sturges, S. Saxena, A. Gopal, D. Millstein, N. Shah, and E. A. Gilmore. *Key Factors Influencing Autonomous Vehicles' Energy and Environmental Outcome*, pages 127–135. Springer, 2014.
- [45] T. Möller, A. Padhi, D. Pinner, and A. Tschiesner. The Future of Mobility is at Our Doorstep. <https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/the-future-of-mobility-is-at-our-doorstep>, 2019. Published: 2019-12-19, Assessed: 2022-01-26.
- [46] V. P. Nelson. Fault-Tolerant Computing: Fundamental Concepts. *Computer*, 23(7):19–25, 1990.
- [47] P. Nenninger. *Vernetzung verteilter sicherheitsrelevanter Systeme im Kraftfahrzeug*. Dissertation, Universitätsverlag Karlsruhe, 2007.
- [48] R. P. Ocampo. Limitations of Spacecraft Redundancy: A Case Study Analysis. In *Proceedings of the 44th International Conference on Environmental Systems (ICES 2014)*, 2014.
- [49] X. Olive. FDI(R) for Satellite at Thales Alenia Space: How to Deal with High Availability and Robustness in Space Domain? In *Proceedings of the 2010 Conference on Control and Fault-Tolerant Systems (SysTol'10)*, pages 837–842, 2010.
- [50] L. Pierallini. Automation in the Aviation Industry and Liability Related Aspects. <https://www.expertguides.com/articles/automation-in-the-aviation-industry-and-liability-related-aspects/arraamlx>, 2019. Published: 2019-08-20, Assessed: 2022-02-23.
- [51] V. Profillidis. *Railway Management and Engineering*. Routledge, 2017.
- [52] M. Realpe, B. X. Vintimilla, and L. Vlacic. A Fault Tolerant Perception System for Autonomous Vehicles. In *35th Chinese Control Conference (CCC 2016)*, pages 6531–6536, 2016.
- [53] M. Sadighi, S. Setayeshi, and A. A. Salehi. PWR Fuel Management Optimization Using Neural Networks. *Annals of Nuclear Energy*, 29(1):41–51, 2002.
- [54] P. Schleiss, C. Drabek, and G. Weiss. SafeAdapt Deliverable 3.1 - Concept for Enforcing Safe Adaptation during Runtime. https://www.safeadapt.eu/content/dam/esk/safeadapt/en/documents/SafeAdapt_D3_1_Concept_for_Enforcing_Safe_Adaptation_during_Runtime.pdf, 2014. Published: 2014-12-31, Assessed: 2022-01-26.

- [55] Space News and SPACE.com Staff. Timeline: 50 Years of Spaceflight. <https://www.space.com/4422-timeline-50-years-spaceflight.html>, 2012. Published: 2012-09-28, Assessed: 2022-01-26.
- [56] R. Suich and R. Patterson. How much Redundancy: Some Cost Considerations, Including Examples for Spacecraft Systems. In *AIChE Summer National Meeting Session on Space Power Systems Technology*, number E-5592, 1990.
- [57] P. Traverse, I. Lacaze, and J. Souyris. Airbus Fly-by-Wire: A Total Approach to Dependability. In *Proceedings of the IFIP 18th World Computer Congress, volume 156 of IFIP International Federation for Information Processing*, pages 191–212. Springer, 2004.
- [58] J. Wang, J. Wang, C. Roberts, and L. Chen. Parallel Monitoring for the Next Generation of Train Control Systems. *IEEE Transactions on Intelligent Transportation Systems*, 16(1):330–338, 2015.
- [59] Y. Wang, M. Zhang, J. Ma, and X. Zhou. Survey on Driverless Train Operation for Urban Rail Transit Systems. *Urban Rail Transit*, 2(3):106–113, 2016.
- [60] J. Wilson. *Turbulence Aloft: The Civil Aeronautics Administration Amid Wars and Rumors of Wars*, 1938–1953. U.S. Department of Transportation, 1979.
- [61] World Nuclear Association. Safety of Nuclear Power Reactors. <https://world-nuclear.org/information-library/safety-and-security/safety-of-plants/safety-of-nuclear-power-reactors.aspx>, 2022. Published: 2022-03-01, Assessed: 2022-03-11.
- [62] I. Yaqoob, L. U. Khan, S. M. A. Kazmi, M. Imran, N. Guizani, and C. S. Hong. Autonomous Driving Cars in Smart Cities: Recent Advances, Requirements, and Challenges. *IEEE Network*, 34(1):174–181, 2020.
- [63] Y. Yeh. Triple-Triple Redundant 777 Primary Flight Computer. In *Proceedings of the 1996 IEEE Aerospace Applications Conference (AeroConf 1996)*, pages 293–307, 1996.

Abbildung rechtlicher Anforderungen auf die Architektur komplexer Fahrerassistenzsysteme und autonomer Fahrzeuge

Mapping of Legal Requirements to the Architecture of Complex Driver Assistance Systems and Autonomous Vehicles

Günter Becker, Alexios Camarinopoulos, Aikaterini Papasileka, Evangelia Karamanoli,

RISA Sicherheitsanalysen GmbH, Berlin;

Maria Rill, Daniel Vonderau, Bo Liu, Victor Pazmino Betancourt,

Jürgen Becker,

FZI Forschungszentrum Informatik, Karlsruhe

Zusammenfassung

Der Entwurf sicherheitsrelevanter Systeme im Automotive Bereich erfordert neue Methoden und Werkzeuge, die eine durchgängige und nachvollziehbare Entwicklung erlauben. Um der Regulierungsdynamik im Automotive Sektor für automatisierte Fahrfunktionen, angemessen zu begegnen, ist es erforderlich, entsprechende Gesetzesänderung sowohl retrospektiv nachzuvollziehen als auch die bei Gesetzesänderungen eintretenden Folgen in der funktionalen Architektur abzubilden und nachvollziehbar zu machen, da sie Auswirkungen auf Entwicklungsprozesse hat.

Es wird daher ein Informationssystem modellhaft dargestellt, welches u.a. eine Zuordnung zwischen Bestandteilen des funktionalen Entwurfs und der Architektur und den dabei eingeflossenen Regelwerken und Gesetzen und deren Änderungen verwaltet.

Abstract

The design of safety-relevant systems in the automotive sector requires new methods and tools that allow consistent and comprehensible development. In order to adequately address the regulatory dynamics in the automotive sector for automated driving functions, it is necessary to understand the corresponding changes in the law retrospectively and to map the consequences of changes in the law in the functional architecture and make them comprehensible, as they have an impact on development processes.

An information system is therefore presented as a model which, among other things, manages an assignment between components of the functional design and the architecture and the regulations and laws that have flowed in and their changes.

Einführung

Mit Einführung von Fahrerassistenzsystemen zunehmender Komplexität bis hin zum vollständig autonom fahrenden Kraftfahrzeug, nehmen die Anforderungen an die Realisierung der Elektrik/Elektronik (Sensoren, Aktoren, IT) in einem Ausmaß zu, dass dies Konsequenzen für den ganzen Lebenszyklus hat. So entstehen neben neuen Daten und auch die dafür notwendigen Zugriffsmöglichkeiten, wie beispielsweise für die Kontrolle und Wartung über Fernzugriffe.¹

Zuverlässige und sichere Kommunikation ist dabei eine der wichtigsten Schlüsseltechnologien, um insbesondere vollständig autonom fahrende Fahrzeuge in der Breite praxistauglich zu machen. Neben der zunehmenden Komplexität der Technik nimmt auch die Menge der einzuuhaltenden technischen und rechtlichen Regelwerke zu. Zusätzlich zu den technischen Vorgaben, die für die Fahrzeugzulassung zu erfüllen sind und somit die Beschaffenheit des Fahrzeugs betreffen, müssen Hersteller autonomer Fahrzeuge nun auch rechtliche Anforderungen hinsichtlich des Verhaltens des Fahrzeugs im Straßenverkehr beim Einsatz autonomer Fahrfunktionen umsetzen. Dies betrifft insbesondere die Fahrassistenzsysteme. Zu den rechtlichen Anforderungen zählen insbesondere auch Anforderungen an Nachweise sowohl für die funktionale Sicherheit (Ausfallsicherheit/Safety), als auch für die Manipulationssicherheit und den Datenschutz (Security). So besteht beispielsweise für den Datenschutz die rechtliche Verpflichtung diesen bereits zur Entwicklungszeit (Art 25 DSGVO Privacy by Design) zu berücksichtigen. Dies hat aufgrund der Verankerung in Art 32. DSGVO auch Auswirkungen auf das Systemdesign hinsichtlich der Sicherheit der Daten und des Systems. Für einige dieser Anforderungen fehlen jedoch entsprechende sektorspezifische Regelungen, was es erforderlich macht, auf allgemeine Vorgaben zurückzugreifen.² Dabei ist sowohl der internationale als auch der nationale Normgeber bestrebt, entsprechende Regelungen dem tatsächlich vorherrschenden Innovationsstand anzupassen, um die Rahmenbedingungen für die autonome Mobilität zu schaffen.³

¹ Ernst, R. (2018): Automated Driving: The Cyber-Physical Perspective. - Computer, 9/2018: 76-79.

² Oppermann, Stender-Vorwachs: Autonomes Fahren - Technische Grundlagen, Rechtsprobleme, Rechtsfolgen, 2020, München, C.H. Beck

³ Wagner, M, Das neue Mobilitätsrecht - Der Rechtsrahmen zum automatisierten und vernetzten Fahren, S. 80, Nomos, 2021.

Um dieser Regulierungsdynamik angemessen zu begegnen, ist es erforderlich, entsprechende Änderungen sowohl retrospektiv nachzuvollziehen als auch die bei Gesetzesänderungen eintretenden Folgen in der funktionalen Architektur abzubilden und nachvollziehbar zu machen. Ein Informationssystem, welches u.a. eine Zuordnung zwischen Bestandteilen des funktionalen Entwurfs und der Architektur und den dabei eingeflossenen Regelwerken und Gesetzen und deren Änderungen verwaltet, kann daher nicht nur hilfreich, sondern sogar zwingend erforderlich sein. Der Entwurf sicherheitsrelevanter Systeme im Automotive Bereich erfordert neue Methoden und Werkzeuge, die eine durchgängige und nachvollziehbare Entwicklung erlauben. Ein sicheres Gesamtsystem beinhaltet deshalb nicht nur die Absicherung der zu übertragenden Daten und die Absicherung der Kommunikationssignale auf physikalischer Ebene und der externen Zugriffe, sondern im Besonderen auch die Verwaltung der Einhaltung der rechtlichen Aspekte. Auf detaillierter Ebene wird daher erfasst, welche rechtliche Anforderung wo zu berücksichtigen war und in welche Systeme und Komponenten eingeflossen ist.

Im Anschluss wird dieser Ansatz in den Kontext weiterer Ergebnisse des DEFEEnD Projekts gestellt, und es wird in einem Ausblick auf weitere Forschungsthemen verwiesen.

Modellbasierte Entwicklungsansätze: DEFEEnD

DEFEEnD ist ein vom BMBF unter der Rubrik KMU – Innovativ gefördertes Forschungsprojekt. Zentrale Anliegen im Projekt sind die Erforschung und Entwicklung von Methoden und Werkzeugen für das sichere hochautomatisierte und autonome Fahren.

Im Fokus stehen dabei modellbasierte Vorgehensweisen, die den Entwicklungsprozess von der Anforderungsanalyse über die Umsetzung bis hin zu Tests und Validierung unterstützen. Zielkonflikte, die aus sich widersprechenden Anforderungen an Betriebs-, Ausfall- und Angriffssicherheit resultieren, können auf diese Weise in frühen Planungs- und Entwicklungsphasen erkannt und behandelt werden. Bewährte, jedoch oftmals voneinander isoliert angewendete Methoden, Prozesse und Normen aus dem Bereich der Betriebs- und Angriffssicherheit werden integriert und harmonisiert.

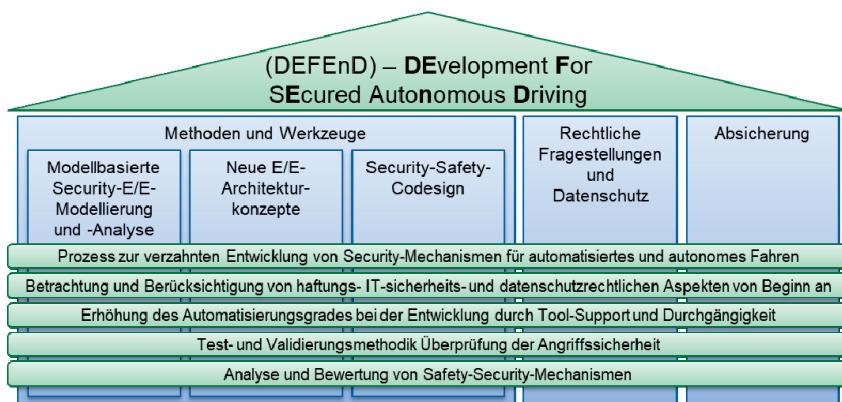


Bild 1: Aufbau des DEFEnD Projekts

Die umfassende Berücksichtigung technischer, sicherheitsrelevanter und rechtlicher Aspekte ermöglicht zudem eine angemessene Risikobewertung. Vom einzelnen Bauteil bis hin zur komplexen Software müssen vielschichtige Sicherheitsanforderungen in künftigen Fahrzeuggenerationen berücksichtigt werden. Ein innovativer Ansatz ist dabei der modellbasierte Entwicklungsansatz. Dieser ermöglicht es, Entwicklungs- und Verifizierungsprozesse zu optimieren, die Wahrscheinlichkeit von Schwachstellen und Fehlern in der Fahrzeugsoftware zu verringern und somit Systemqualität, Zuverlässigkeit, Schutz und Sicherheit zu erhöhen.

Struktur der rechtlichen Grundlagen

Die Struktur der rechtlichen Rahmenbedingungen zum automatisierten Fahren, wird ganz wesentlich durch den internationalen Rahmen zum Straßenverkehrsrecht geprägt. So gibt das Wiener Übereinkommen über den Straßenverkehr von 1968 (WÜ)⁴ die völkerrechtlichen Rahmenbedingungen für die Gestaltung des nationalen Verkehrsrechts vor. Da da WÜ als völkerrechtlicher Vertrag keine unmittelbare Wirkung in den Vertragsstaaten hat⁵, bedarf es zur Weiterentwicklung und Ausarbeitung von technischen Regularien eines weiteren Gremiums – die Arbeitsgruppen der UNECE (United Nations Economic Commission for Europe). Die technischen Regularien, die diese Gruppen (sog. Working Groups) aufstellen (im folgen-

⁴ Wiener Übereinkommen vom 8. November 1968 über den Straßenverkehr (BGBl. 1977 II S. 809, 811).

⁵ Lutz, NJW 2015, 119 (122).

den UNECE-Regeln genannt) enthalten detaillierte Vorgaben zu bestimmten Bauteilen oder Fahrzeugfunktionen und orientieren sich jeweils am aktuellen Stand der Technik.

Konkretisierende UNECE-Regeln, die automatisierte und autonome Fahrzeuge betreffen, reichen beispielhaft von der UN Regulation No. 79⁶ (Lenkanlagen) über die UN Regulation No. 155⁷ (Cyber Security & Cyber Security Management System und No 156⁸ (Software Update & Software Update System) bis hin zur UN Regulation No. 157⁹ (Automated Lane Keeping Systems).

Für Fahrzeuggenehmigungen ist zudem die EU-Rahmengesetzgebung maßgeblich (EU-Typengenehmigung), die sich unmittelbar auf die nationalen Vorschriften aus dem Zulassungsrecht auswirkt und die wiederum auf die technischen Anforderungen der UNECE-Regeln verweist.

Die Komplexität des Zusammenspiels der verschiedenen Normenebenen und der sich hieraus ergebenden Änderungen, die Auswirkungen für die Entwicklung von autonomen Fahrzeugen haben, lässt sich an folgendem Beispiel veranschaulichen: Auf internationaler Ebene werden Regeln zum Cybersecuritymanagement und Softwareupdates erlassen. Hierunter fallen auch Regelungen zum sogenannten Data Storage System for Autonomous Driving (kurz DSSAD). Der nationale Gesetzgeber in Deutschland erließ dieses Jahr eine Rechtsverordnung zum Fahrmodusspeicher¹⁰. Nun stellt sich die Frage, wie diese Regelungen zusammenhängen, welche vorrangig gelten und wie sie umzusetzen sind. Ob somit eine Änderung auf internationaler Ebene zu Änderungen oder Anpassungen auf nationalem Level führt, ist für den Rechtslaien daher kaum noch zu beantworten. Hierfür ist historische Verwaltung rechtlicher Anforderungen notwendig.

Historische Verwaltung rechtlicher Anforderungen (Change Impact Management)

Gesetze und andere verbindliche Vorschriften sind in unterschiedlichem Maße volatil. Sie ändern sich, werden als obsolet außer Kraft gesetzt, und es kommen neue hinzu. Die im Rahmen des Projekts entwickelte Datenstruktur muss hinreichend flexibel sein, um diesen

⁶ <https://unece.org/sites/default/files/2021-02/R079r4am3e.pdf>.

⁷ <https://unece.org/sites/default/files/2021-03/R155e.pdf>.

⁸ <https://unece.org/sites/default/files/2021-03/R156e.pdf>.

⁹ <https://unece.org/sites/default/files/2021-03/R157e.pdf>.

¹⁰ Verordnung zur Regelung des Betriebs von Kraftfahrzeugen mit automatisierter und autonomer Fahrfunktion und zur Änderung straßenverkehrsrechtlicher Vorschriften; https://www.bmvi.de/SharedDocs/DE/Anlage/K/presse/008-verordnung-automatisierte-autonomefahrfunktion.pdf?__blob=publicationFile.

Herausforderungen Rechnung zu tragen. Die Nachvollziehbarkeit von Architekturentscheidungen, die auf rechtlichen Anforderungen basieren sowie das Tracking von Änderungen von rechtlichen Anforderungen mit Auswirkung auf die funktionale Systemdefinition werden daher durch ein sog. Change-Impact-Management ermöglicht.

Rechtliche Anforderungen sollten bereits im Rahmen der Anforderungsanalyse im Entwurfsprozess erhoben werden. Diese können als funktionale und nichtfunktionale Anforderungen in den Systementwurf einfließen. In der Automobilbranche wird standardmäßig das Werkzeug DOORS von IBM zur Erfassung und Dokumentation von Anforderungen im Allgemeinen verwendet. Hierbei können auch die rechtlichen Anforderungen, die sich beispielsweise aus Normen (bspw. DIN) ergeben, berücksichtigt werden. Jedoch müssen bei Änderungen im Rahmen des Change Impact Managements auf diese Anforderungen zugegriffen werden können. Eine Nachvollziehbarkeit bzgl. des Ursprungs der funktionalen Anforderung muss hier gegeben sein. Eine Software, die die Auswirkungen von Änderungen von rechtlichen Anforderungen aufzeigt bzw. die die Erfassung dieser Änderungen auf feingranularer Ebene ermöglicht ist bisher nicht bekannt.

Im Folgenden wird diese Struktur EDV-technisch nachgebildet. Die einzelnen gesetzlichen Anforderungen, die man in ihrer Gesamtheit als Lastenheft für den technischen Entwurf interpretieren kann, korrelieren nun mit einem technischen Pflichtenheft, mit dem die Anforderungen umgesetzt werden sollen.

Struktur des technischen Entwurfs

Im Folgenden wird zunächst von rechtlichen Anforderungen (Grundlagen) ausgegangen, ohne auf eine etwaige Hierarchie einzugehen.

Die technischen Anforderungen folgen aus den rechtlichen Grundlagen. Sie werden definiert durch entsprechende Systemfunktionen, die als technische Systeme realisiert werden. Diese Systeme sind aus Hardwarelementen zusammengesetzt. Durch logische Funktionen und Ereignisse wird eine Schnittstelle zur funktionalen Sicherheitsanalyse und auch zu Fragestellungen der Security gebildet. Basisereignisse und Gatter können sowohl in Fehlerbaumanalysen als auch in Angriffsbaumanalysen einfließen.

Indem die rechtlichen Anforderungen somit auf Bestandteile des technischen Entwurfs abgebildet werden, wird eine Information dokumentiert, die man sonst mühevoll extrahieren muss.

Die Bild 2 unten zeigt die Struktur, die in einer entsprechenden Datenbankapplikation mit historisch verwalteten Datensätzen implementiert worden ist.

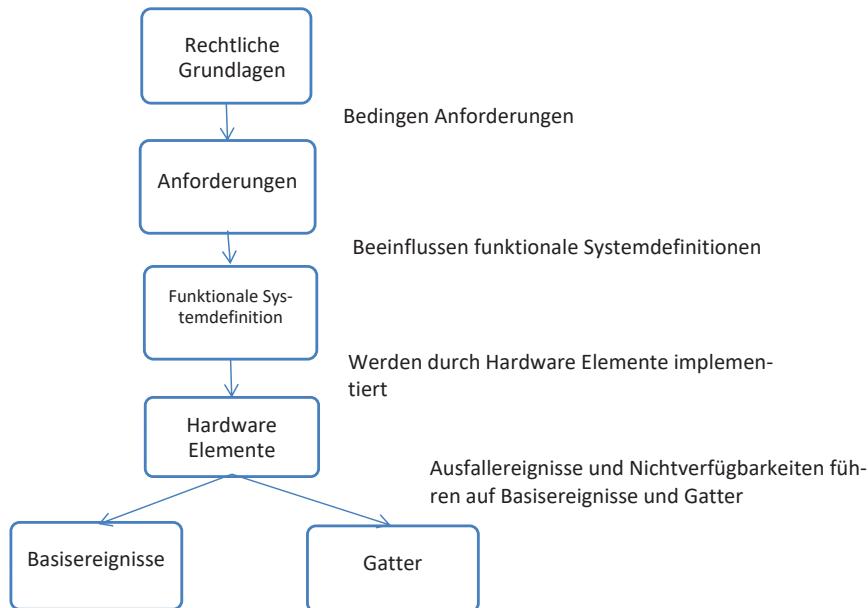


Bild 2: Grobdarstellung der rechtlichen und der technischen Konstrukte

In der Nutzeroberfläche wird die Struktur baumartig dargestellt, wie es in Abb gezeigt wird.

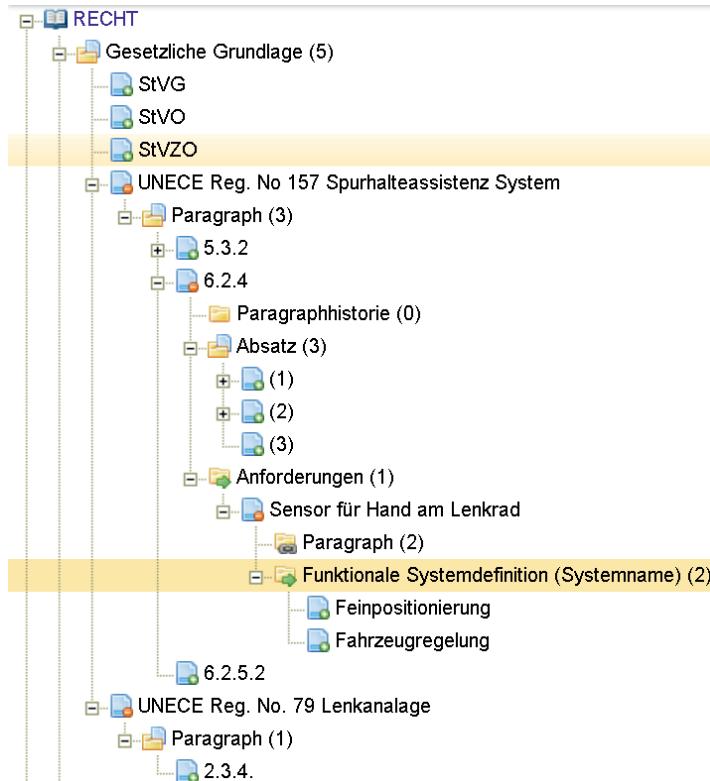


Bild 3: Benutzeroberfläche des Rechtsinformationssystems

Man erkennt, wie man von der UNECE-Regelung ausgehend über die sog. „Paragraphen“ (die Unterabschnitte der Regelung wurden im Rahmen des Projekts als „Paragraphen“ aufgefasst) zu einzelnen Abschnitten kommt. Beispielsweise ist eine Anforderung aus diesem Paragraphen eingefügt worden, nämlich die Anforderung, die Konzentration des Fahrers muss überwacht werden, indem z.B. geprüft wird, ob er eine Hand am Lenkrad hat.

Auch von der Anforderungsliste ausgehend kommt man auf die Anforderung und auch auf die Paragraphen, zu denen eine Relation besteht (siehe Bild 4).

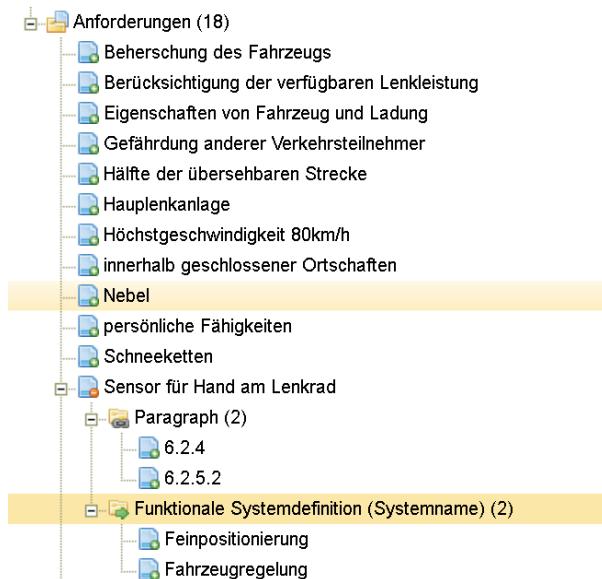


Bild 4: Traversierung von den Anforderungen zu den Paragraphen, welche diese bedingen und zu den Systemdefinitionen

Anwendungsbeispiele

Ein Entwicklungsprozess wird praktisch nie statisch geplant um in genau der Weise umgesetzt zu werden, wie er einmal geplant war. Der Prozess muss dynamisch reagieren, wenn sich zu einem beliebigen Zeitpunkt im Produktlebenszyklus von außen eingeprägten Ereignissen funktionale Anforderungen ändern, so muss der Prozess eine eindeutige Reaktion implementieren. Ein Beispiel für diesen Vorgang zeigt die Bild 5. Externe Anforderungen können die verschiedensten Ursachen haben. Eine dieser Ursachen sind Änderungen im Regelwerk, zu welchem auch die gesetzlichen Grundlagen zählen. In der Tat entstehen auch Änderungen im Regelwerk nicht ex nihilo, sondern sie reflektieren z.B. neuartige technologische Möglichkeiten, die aber aufgrund eines nicht vernachlässigbaren Gefährdungspotenzials entsprechende Anpassungen erfordern. Dies wird hier aber nicht weiter verfolgt; an dieser Stelle dient der Anstoßprozess lediglich zur Generierung von Anforderungen, auf die der Entwicklungsprozess zu reagieren hat.

Ein Beispiel einer solchen Änderung besteht im Zusammenhang mit der UNECE-Regel zu den Lenkanlagen. Mit dem Aufkommen von Spurhalteassistenten werden Überlegungen zur

Gestaltung einer Lenkanlage erforderlich, die Regeln zur sicheren Handhabung solcher Assistanten festlegen.

Sofern ein Verweis auf die UNECE Regelungen Nr. 157 für Spurhalteassistentensysteme und Nr. 79 für die Lenkanlage erfolgt, ergeben sich mit der Einführung der Regelung Nr. 157 die in **Fehler! Verweisquelle konnte nicht gefunden werden.** angegebenen Erweiterungen.

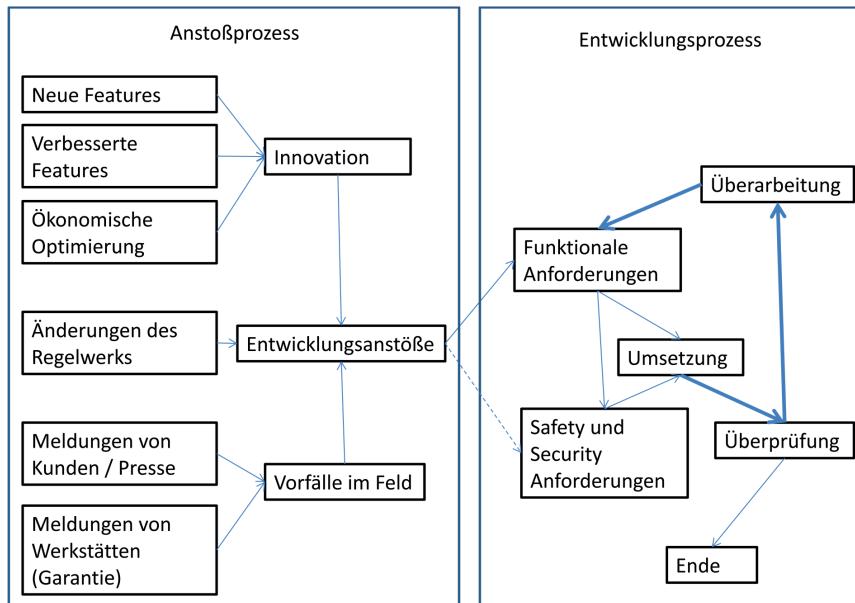


Bild 5: Externe Entwicklungsanstöße wirken auf den Entwicklungsprozess ein

Folgende mögliche automatische Lenkfunktionen sind in der UNECE Regelung für Fahrsysteme geregelt.

- Automatische Lenkfunktionen in 2.3.4.1. geregelt: 6 Kategorie (A, B1,B2, C,D,E)
(Einparken (A), Spurhalten (B1,B2), Spurwechsel (C, D, E)
 - Unterscheidung hinsichtlich „was muss der Fahrzeugführer veranlassen“
- Korrigierende Lenkfunktionen in 2.3.4.2. geregelt
- Notlenkfunktion in 2.3.4.3. geregelt
- Bauvorschriften sind in Nr. 5 geregelt. Für korrigierende Lenkfunktionen und Notlenkfunktionen gibt es weitergehende Anforderungen

Hinzu kommen

- Anforderungen an
- Safety
 - generelle Anforderungen 9 +
 - Dynamische Fahrfunktion (5.2): 5 +5 Unteranforderungen
 - Emergency Manoeuvre (5.3): 7 Anforderungen
 - Transition Demand (5.4): 4, Unteranforderungen: 8
 - Minimum Risk Manoeuvre (5.5)
- HMI (human machine interface)
- Anforderungen an die Sensorik (Kapitel7) für Umgebungserkennung → 7 Anforderungen
- DSSAD (data storage system for automated driving) → 11 Anforderungen
- Cybersecurity

Im Abschnitt 5.3.2 der UNECE-Regelung Nr. 157 steht beispielsweise, dass die verfügbare Lenk- und Bremsleistung bei einem Ausweichmanöver zu Anwendung zu kommen hat. Sollte aufgrund von Störungen die volle Leistung nicht zur Verfügung stehen, so muss dies berücksichtigt werden.

Diese rechtliche Anforderung hat für die funktionalen Anforderungen die Konsequenz, dass die Feinpositionierung und die Fahrzeugregelung genau diese Information benötigen. An Lenk- und Bremsanlage sind entsprechend Sensoren vorzusehen, die diesbezügliche Einschränkungen bekannt machen.

Alle rechtlichen und funktionalen Anforderungen werden mit einem Gültigkeitsdatum versehen. Sie werden ferner mit den Systemfunktionen des Fahrzeugs verlinkt, auf die sie Einfluss haben. Ist diese Arbeit getan, so kann das System mit einer geeigneten Abfrage zu jedem Zeitpunkt diejenigen Funktionen und Teilfunktionen angeben, auf die eine Regelung zeigt, deren Gültigkeitsdatum größer (jünger) ist, als das eigene. Anhand eines solchen Abfrageergebnisses kann jederzeit nachvollzogen werden, ob es noch Teile in der Systemdefinition gibt, die nicht mit neueren Regelungen abgeglichen sind.

Sobald ein solcher Abgleich erfolgt wird, wird der Zeitstempel im Systemmodell aktualisiert, so dass diese Stelle beim nächsten Aufruf der Abfragefunktion nicht mehr angezeigt wird.

Kontext des DEFEnD Projekts und Anschluss

Im Kontext des DEFEnD Projektes stellt die oben vorgestellte Methode einen Baustein dar, um sowohl Safety als auch Security Aspekte mit Hilfe eines Toolings, um Entwicklungsprozesse abzubilden. Im Anschluss an dieses Projekt, hat sich erneut ein Konsortium zusammen gefunden, um im Projekt UNCOVER¹¹ Verfahren und Werkzeuge zu erforschen, die der systematischen und strukturierten Erfassung von Sicherheitsvorfällen bei autonomen Fahrfunktionen dienen. Dabei sollen Sicherheitsvorfälle und Erkenntnisse aus dem Feldbetrieb in die Entwicklungswerzeuge fließen.

¹¹ UNCOVER ist ein vom BMBF unter der Rubrik KMU – Innovativ gefördertes Forschungsprojekt. Es steht für Datenbasierte Entwurfsunterstützung durch kontinuierliches Monitoring von Sicherheitsvorfällen.

Real World Scenarios for the Safety Validation and Development of Highly Automated Driving Functions

Tobias Moers, Christoph Klas, Dr. Adrian Zlocki,

fka GmbH, Aachen;

Lennart Vater,

Institute for Automotive Engineering, RWTH Aachen University

Zusammenfassung

Reale Szenario-Daten sind für die Weiterentwicklung und Validierung des automatisierten Fahrens von wichtiger Bedeutung. Darunter fallen beispielsweise im realen Verkehr gemessene Szenarien oder Referenzmessungen für die Umgebungswahrnehmung eines Prototyps. Die fka hat mit levelXdata Ansätze entwickelt, um mit Hilfe von Drohnen hochgenaue Messdaten zu erzeugen, die nicht den typischen Nachteilen etablierter bodengebundener Messmethoden wie Fahrzeug-zu-Fahrzeug-Verdeckungen und Beeinflussung der Verkehrsteilnehmer unterliegen. Mit Hilfe dieser Daten können Trajektorien und damit Szenarien aus dem Realverkehr erfasst und fahrzeugunabhängige Referenzdaten für die Umweltwahrnehmung erzeugt werden.

Abstract

Real scenario data is important for the development and validation of automated driving. This includes, for example, scenarios measured in real traffic or reference measurements for the environmental perception of a prototype. With levelXdata, fka has developed approaches for using camera-equipped drones to generate highly accurate measurement data that are not subject to the typical disadvantages of established ground-based measurement methods such as vehicle to vehicle occlusions and influencing other road users' behavior. These data can be used to extract trajectories and therefore scenarios from real traffic and produce vehicle-independent reference data for environmental perception.

1. Introduction

Highly automated driving promises multiple benefits for future traffic. This is on the one hand, in form of improved comfort for individual passengers and, on the other hand, in form of increased traffic safety and flow. The development and validation of highly automated driving systems, however, are a highly complex challenge. Conventional driver assistance systems are limited to certain operational design domains and individual scenarios that can be analyzed on the basis of expert knowledge. This is no longer possible with highly automated systems (ADS), due to their complexity and their use of data-driven application for e.g. environment perception.

For ADS, it must be proven systematically that they drive safer than human drivers. For a pure statistical validation, test drives of hundreds of millions of kilometers would be needed [2]. Therefore, instead of relying purely on expert knowledge, data-driven methods are used, that are based on large datasets of sensor data and naturalistic driving behavior in the form of trajectory datasets. These trajectory datasets of road users measured in real traffic are crucial for several application areas in the development of automated driving, as depicted in Fig. 1. Naturalistic trajectory data can be used, for example, for developing models for predicting road user behavior, for evaluations of trajectory planning algorithms or in the scenario-based safety validation process of a complete ADS [4], [5]10.

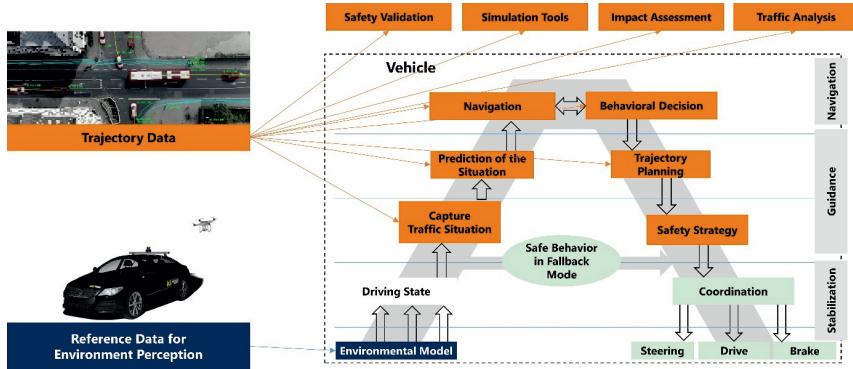


Fig. 1: Applications of trajectory and reference data for automated driving, based on the A-model [3].

Such trajectory datasets containing naturalistic driving data can, however, not be created in simulations or on test tracks. The recorded behavior in those settings is either derived from simple models or is instructed. Thus, neither the natural diversity nor the frequency distribution

of driving behavior and road infrastructure found in public road traffic can be adequately captured and recorded.

An additional unsolved challenge is the generation of highly accurate reference data to prove that the sensor technology of the highly automated vehicle captures the environment with sufficient accuracy and reliability. To meet both of these challenges, fka has developed a new method named levelXdata [1] to generate such data from unmanned aerial vehicles ("drones") equipped with high-resolution cameras.

In the following the dedicated camera drone-based measurement method is introduced and compares to other similar method. Chapter 3 focuses on the generation of trajectory datasets and Chapter 4 covers the generation of reference measurements. The content is summarized and concluded in Chapter 5.

2. Drones as a Measurement System

There exist different approaches to capturing real world trajectory data by means of sensor technology. Fig. 2 shows an extensive comparison of four different measurement methods. As all of these methods use different sensors and setups, they also offer different advantages and disadvantages.

Series-production vehicles equipped with basic environment perception sensors, dedicated measuring vehicles, or vehicles equipped with reference systems, have a high degree of flexibility in terms of the recording location and the sensors used. They can also record surrounding traffic participants with high accuracy. Vehicles equipped with additional sensors, however, might be more noticeable and, thus, influence the behavior of other road users. That hinders recording of naturalistic driving data. In addition, the perspective of the used sensors might lead to incomplete data collection caused by vehicle to vehicle occlusions.

Another approach for trajectory data collection are infrastructure sensors. These are sensors mounted on poles that are positioned close to a traffic area. Each of the poles can host multiple sensors and by using more poles at e.g. an intersection, the traffic participants there can be captured with high accuracy and reliability. Incomplete data collection due to vehicle to vehicle occlusions can also be addressed by fusing data from multiple poles. That means, that these systems are efficient for collecting highly accurate trajectory data from a dedicated traffic area over a long time span. Disadvantages are the high costs of the sensors and the installation. Moreover, they can usually not be used flexibly to collect data at different traffic areas, since relocation of the poles is also connected with a higher amount of work.

The camera drone-based measurement method benefits from the perspective of the drone. It enables the entire traffic scene to be recorded without vehicles being occluded by each other.

In addition, measurements with a drone are time efficient because all road users in the field of view are recorded simultaneously. At heights of 80 m to several hundred meters, the drone is barely visible from a passing vehicle, resulting in completely unaffected, naturalistic driving behavior of all capture traffic participants. Compared to infrastructure sensors or special measurement vehicles, the acquisition costs are quite low. Another benefit of this approach is its flexibility. Different data recording sites can be chosen based on the traffic domain of interest (urban, highway, rural) or short-time construction sites. The measurement method is, however, limited by weather conditions, since most drones cannot fly during heavy rain, snow or wind.



Fig. 2: Comparison of four measurement methods for trajectory dataset generation. Series-production vehicles, measurement vehicles equipped with reference systems, infrastructure sensors and the aerial perspective representing the drone-based method.

3. Drone-based Trajectory Dataset Generation

For the dataset generation, a camera-equipped drone records high-resolution videos of public traffic or a test scenario from a bird's eye view. With the help of state-of-the-art object detection and tracking algorithms, the positions and movements of all visible road users can be extracted from the video recordings with high precision. Incorrectly detected data can then be identified and corrected by a quality management process. This ensures consistent and high data quality.

Based on this approach, multiple datasets from different driving domains have already been created. Fig. 3 shows a section of the highD dataset, which consists of road user trajectories from German highways. The figure illustrates the trajectory extraction process from the aerial

video footage. Due to the deep learning and tracking algorithms, the vehicles can be detected and tracked very accurately. The highD dataset contains about 110 500 tracked vehicles from six different highway sections between Cologne and Aachen. Since its release in 2018 it has been used and cited by many researchers.



Fig. 3: A cutout from the video recordings of the highD dataset. The upper part shows the whole highway section and the lower part provides a closer look.

Since not only highway trajectory data is of interest, the same method has been used to create other publicly and non-publicly available datasets. The urban area is a highly complex area where there are many challenging scenarios for a highly automated vehicle. These scenarios can be detected and extracted from the levelXdata datasets. The inD dataset contains data from four intersections in Aachen. All relevant road users were detected: cars, trucks, buses, pedestrians and bicycles/motorcycles. A first impression of the dataset can be found in Fig. 4. Especially the interaction between vulnerable road users and vehicles is interesting, since these can end in serious accidents.



Fig. 4: A sample image of the inD dataset showing the detections and trajectories of each road user.

The uniD dataset includes even more pedestrian to vehicle interactions. The video footage for this dataset was recorded in front of the university building in Aachen. Fig. 5 shows a section of the uniD dataset in which many pedestrians cross the street at different points without the direct support of a traffic light. This leads to interaction between vehicles and pedestrians and, thus, provokes interesting scenarios.



Fig. 5: A sample image from the uniD dataset showing the detected road users and their trajectories.

Since there has been a great deal of interest in highway data in recent years, another highway dataset was created to supplement the highD dataset. HighD includes many interesting scenarios due to the large amount of included vehicles. The exiD dataset includes data from seven on- and off-ramps, which are shown in Fig. 6. Data were specifically recorded at on- and off-ramps. The reason is that these areas usually require merging actions that lead to interaction between vehicles. Due to the selection of recording sites many of those lateral movement scenarios including interaction of oncoming vehicles with vehicles already on the highway can be found in the dataset.

In order to support research, the datasets mentioned are available free of charge for non-commercial public research. In recent years, such research support has already been provided by levelXdata for the highD dataset [6], the intersection drone dataset inD [7], the roundabout drone dataset rounD [8], the university drone dataset uniD [9] and the exiD dataset (entries and exits drone dataset) [10]. A license to use these datasets free of charge for non-commercial public research can be requested online at www.drone-dataset.com.

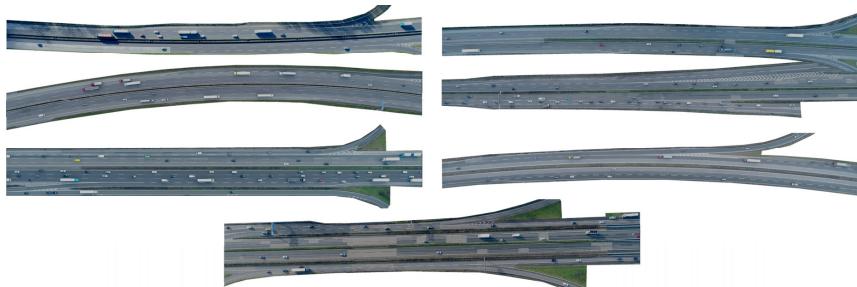


Fig. 6: An overview of the recording sites of the exiD dataset.

As mentioned in Chapter 1, the resulting trajectory datasets, can be used as a data source for a scenario-based validation approach. For such an approach, relevant traffic scenarios need to be available. These can be extracted from the trajectory datasets in a data-driven way. An example for the usage is the Japanese SAKURA project [4] and the German funded project PEGASUS [5].

The UNECE Regulation *R157 Automated Lane-Keeping Systems* [12], which represents the first internationally agreed framework for the introduction of passenger cars with Level 3 systems, requires scenario-based testing of the systems to be introduced and provides a scenario catalog. No specific parameters are specified for this scenario catalog, so fka recognized selected scenarios of the regulation in recorded measurement data and extracted scenario parameters [11]. For this purpose, the dataset of German highways highD [6] was used.

4. Drone-based Reference Measurement System

The environment detection of a highly automated vehicle is a highly complex module due to the use of several different sensors. For the evaluation of these sensors, reference systems are needed provide reference data for the individual sensors, but also for the entire perception system (see Figure 7).

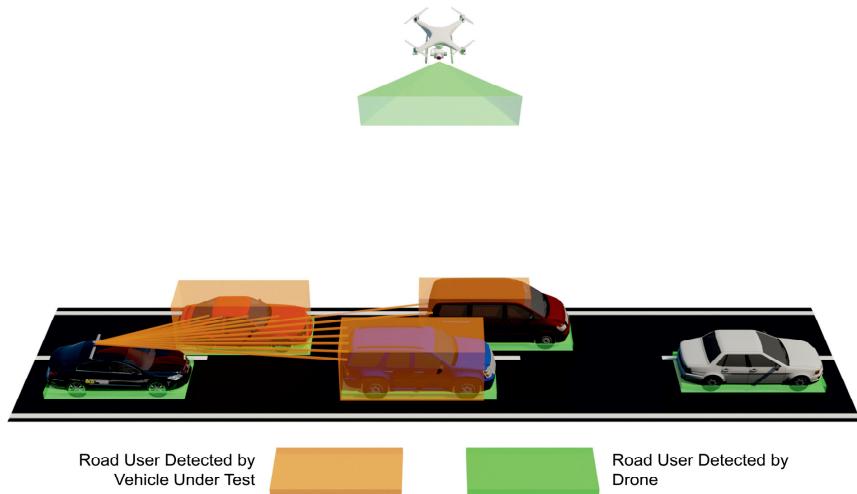


Fig. 7: During the reference measurements of the drone-based reference system, the drone follows the test vehicle, which is equipped with the sensor system under test. The goal is to obtain the reference measurements (green) to compare them to the sensor-system-under-test's measurements (orange).

Fig. 8 compares state of the art methods with the proposed camera drone-based method developed by fka. Known on-board sensor systems include multiple high-resolution 360° LiDAR or camera systems that can provide reference objects. These reference systems must have even higher accuracies than the sensors themselves. The systems used so far are, for example, high-resolution on-board supplementary sensors or IMU-RTK-GNSS systems (hereafter just RTK-systems). However, these must be superior to the sensor-under-test to provide valuable data. The RTK-system provides highly accurate positions of the objects and can thus be used to record data with a few equipped vehicles.

fka has developed a system in which the drone actively follows a test vehicle and can thus take reference measurements on test tracks or in real traffic. An illustration can be found in Fig. 7. The drone can recognize the road users around the test vehicle and provide temporally synchronous reference objects. Thanks to highly accurate temporal synchronization, these can be compared with the objects of vehicle perception.

In general, the drone leverages the same advantages for this measurement method as for the normal trajectory dataset creation. For example, when comparing the RTK-system with the drone in terms of captured road users, every vehicle or object that wants to be observed by

the RTK-system has to be equipped with this system, whereas the drone can detect and observe all visible road users on demand. Another advantage of the drone is cost efficiency, as most on-board reference systems are still very expensive and the cost of RTK-systems scale per installation.

Requirement	More sophisticated sensor technology	IMU-RTK-GNSS	VUT Followed by Drone
Capture All Road Users	Directly surrounding vehicles	Only equipped vehicles	All road users in sensor's field-of-view Both instructed an uninstructed road users
Precision	Must be superior to onboard sensors, Suffers from occlusions	Centimeter accuracy, No occlusions	Centimeter accuracy
Flexibility	Useable on test tracks and public roads	Useable on test tracks and public roads	Useable on test tracks and (most) public roads
Costs	Very high (LiDAR)	Very high, scales with # of captured road users	Reasonable, Does not scale with # of captured road users
Use case	High diversity of different measurement locations	Scenarios with 1-2 object vehicles	Complex scenarios with multiple road users, esp. pedestrians and cyclists

Fig. 8: Comparison of three different measurement systems.

The drone-based reference measurement system of fka is a useful complement to the existing measurement methods. This new methodology makes it easier to evaluate and improve perceptual systems of highly automated vehicles.

5. Conclusion

The development and safety validation of automated driving systems is a complex and challenging topic. Data-driven approaches, which are in the focus for these tasks, require a high amount of highly accurate and diverse real world traffic data. There exist multiple methods to collect this kind of data. However, each of them has limitations for the collection of trajectory as well as reference data. The use of the drone offers a valuable addition for data generation due to its high flexibility, cost efficiency and precision. Therefore, the drone is an effective tool for the generation of datasets on highways and in urban areas, but also on test tracks. Due to the high precision of the data from real traffic, new valuable insights into the interaction between road users can be extracted. In addition, camera drone-based approaches can also be used to evaluate complex sensor systems with respect to its reliability and precision. In contrast to other reference measurement approaches, a single camera sensor can be used also for more complex traffic scenarios and multiple involved vehicles as well as pedestrians.

Bibliography

- [1] www.levelxdata.com: levelXdata - Highly accurate traffic data captured from aerial perspective for all levels of automated driving.
- [2] Wachenfeld, W., Winner, H (2015): Die Freigabe des autonomen Fahrens. - Autonomes Fahren: Technische, rechtliche und gesellschaftliche Aspekte. Vol. 116: p. 439–464., Heidelberg, Germany: Springer.
- [3] Lampe, B., Woopen, T., Eckstein, L. (2019): Collective Driving - Cloud Services for Automated Vehicles in UNICARagil. 28tes Aachener Kolloquiums Fahrzeug- und Motorentechnik
- [4] Jacobo, A., et al. (2019): Development of a safety assurance process for autonomous vehicles in Japan. Proceedings of ESV Conference
- [5] Winner, H., Lemmer, K., Form, T., Mazzega, J. (2019): Pegasus—first steps for the safe introduction of automated driving. Road Vehicle Automation 5, 2019
- [6] Krajewski, R., Bock, J., Kloeker, L., Eckstein, L. (2018): The highD Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems. 21st IEEE ITSC Conference
- [7] Bock, J., Krajewski, R., Moers, T., Runde, S., Vater, L., Eckstein, L. (2019): The inD Dataset: A Drone Dataset of Naturalistic Road User Trajectories at German Intersections. IEEE ICCVE Conference
- [8] Krajewski, R., Moers, T., Bock, J., Vater, L., Eckstein, L. (2020): The rounD Dataset: A Drone Dataset of Road User Trajectories at Roundabouts in Germany. 23rd IEEE ITSC Conference
- [9] uniD: The University Drone Dataset (uniD). Available: <https://www.unid-dataset.com>
- [10] Moers, T., Vater, L., Krajewski, R., Bock, J., Eckstein, L. (2022): The exiD Dataset: A Real-World Trajectory Dataset of Highly Interactive Highway Scenarios in Germany. Submitted to 33rd IEEE Intelligent Vehicles Symposium
- [11] Tenbrock, A., et al. (2021): The ConScenD Dataset: Concrete Scenarios from the highD Dataset According to ALKS Regulation UNECE R157 in OpenX. CoRR. Available: <https://arxiv.org/abs/2103.09772>
- [12] U. N. Economic and S. Council. (2020) Ece/trans/wp.29/2020/81: Proposal for a new UN regulation on uniform provisions concerning the approval of vehicles with regards to automated lane keeping system. [Online]. Available: <https://un-docs.org/ECE/TRANS/WP.29/2020/81>

Entwicklung, Erprobung und Regelbetrieb automatisierter Fahrzeuge aus datenschutzrechtlicher Perspektive

Dipl.-Jur. **Hans Steege**, CARIAD SE, Hannover

Zusammenfassung

Der Beitrag befasst sich mit datenschutzrechtlichen Aspekten bei der Entwicklung, Erprobung sowie dem Regelbetrieb automatisierter und autonomer Fahrzeuge. Betrachtet werden Herausforderungen bei der Fahrzeugkonstruktion, der verwendeten Hard- und Software sowie datenschutzrechtliche Prinzipien, die Hersteller bei der Entwicklung berücksichtigen müssen. Des Weiteren wird thematisiert, wie Erprobungsfahrten datenschutzkonform möglich sind und welche Aspekte es zu beachten gilt. Schließlich befasst sich der Beitrag mit datenschutzrechtlichen Aspekten beim Betrieb von automatisierten und autonomen Fahrzeugen. Hierbei werden anwendbare Rechtssätze aus dem Straßenverkehrsgesetz (§§ 63a, 1g StVG) betrachtet. Letztlich wird aufgezeigt, dass die Datenschutz-Grundverordnung nicht auf den Regelbetrieb automatisierter und autonomer Fahrzeuge zugeschnitten ist.

Abstract

The article deals with data protection aspects in the development, testing and regular operation of automated and autonomous vehicles. The article looks at challenges in vehicle design, the hardware and software used, as well as data protection principles that manufacturers must take into account during the development. Furthermore, the article discusses how test drives can be carried out in compliance with data protection law and which aspects need to be taken into account. Finally, the article deals with data protection aspects in the operation of automated and autonomous vehicles. Applicable legal principles from the Road Traffic Act (§§ 63a, 1g StVG) are considered. In the end, it is shown that the General Data Protection Regulation is not tailored to the regular operation of automated and autonomous vehicles.

1. Einleitung

Autonomes Fahren ist derzeit in aller Munde und tangiert wie kaum eine andere Technologie zahlreiche Rechtsgebiete.^[1] Dabei kommt dem Schutz von Daten eine besondere Bedeutung zu. Sowohl für die Entwicklung, Erprobung als auch den Regelbetrieb automatisierter Fahrzeuge ist eine Vielzahl von Daten nötig – von Testdaten bis hin zu Umgebungsdaten, die das Fahrzeug zur Bewältigung der Fahraufgabe benötigt. Das Recht schützt verschiedene Arten

von Daten und hat dabei unterschiedliche Schutzzwecke, etwa der Schutz von Geschäftsgeheimnissen nach dem Geschäftsgeheimnisgesetz. Für das automatisierte Fahren von erheblicher Relevanz ist indes der Schutz personenbezogener Daten, so dass der Datenschutz-Grundverordnung (DS-GVO) eine Schlüsselposition zukommt. Denn datenschutzrechtliche Vorgaben gilt es bereits bei der Konstruktion des Fahrzeugs zu beachten. Dieser Beitrag soll einen kurzen Überblick über datenschutzrechtliche Aspekte bei der Entwicklung, Erprobung und dem Regelbetrieb automatisierter Fahrzeuge bieten.

2. Datenarten

Insbesondere durch die Vernetzung der Fahrzeuge entstehen bei vernetzten und automatisierten Fahrzeugen stetig große Datenmengen. Diese sind unterschiedlicher Natur und reichen von auf den ersten Blick neutralen Daten – wie etwa Daten über den momentanen Motorzustand oder den Ladezustand einer Batterie – hin zu Daten über die durch die Sensorik aufgenommene Fahrzeugumgebung, wie beispielsweise Personen, Fahrzeuge und Kennzeichen. Aber auch vom Automobilhersteller oder Dritten bereitgestellte Dienste sowie die Konnektivität mit dem Smartphone führen zu personenbezogenen Daten. Die erzeugten Daten lassen sich in verschiedene Kategorien von Fahrzeugdaten aufteilen. Diese sind unter anderem Fahrzeugdaten zum Zweck von Wartung und Reparatur, Daten im Rahmen der Produktbeobachtung, resultierend aus deliktischen Produktbeobachtungspflichten, oder die Weiterentwicklung und Verbesserung künftiger Dienste und Fahrzeugkomponenten sowie Fahrzeuggenerationen.[2]

Datenschutz hat dabei eine grundrechtliche Dimension,[3] die insbesondere den Freiheitsaspekt betont, so dass der Einzelne einerseits jederzeit Kenntnis darüber hat, was und wo etwas mit seinen Daten passiert und andererseits nicht dadurch beeinflusst wird, also der Habitus nicht aus Angst geändert wird. Letzterer Aspekt betont primär den freiheitsrechtlichen Aspekt des grundrechtlichen Datenschutzes. Das Recht auf informationelle Selbstbestimmung ist Ausfluss des allgemeinen Persönlichkeitsrechts sowie der Menschenwürde nach Art. 2 Abs. 1 i.V.m. Art. 1 Abs. 1 GG und geht auf das bekannte Volkszählungsurteil des Bundesverfassungsgerichts zurück.[4] Neben dieser grundrechtlichen Dimension ist seit dem 25. Mai 2018 die europaweit harmonisierte Datenschutz-Grundverordnung maßgeblich. Das nationale Bundesdatenschutzgesetz hat dadurch an Bedeutung verloren und ist nur noch für einzelne Fälle relevant.

3. Fahrzeugdaten, Innenraumdaten sowie Umgebungsdaten als personenbezogene Daten?

Die DS-GVO findet lediglich bei personenbezogenen Daten Anwendung, Art. 2 Abs. 1 DS-GVO. Daher kommt es bei der Frage, ob Daten erhoben, gespeichert und verarbeitet werden dürfen, maßgeblich darauf an, ob es sich um personenbezogene Daten i.S. des Art. 4 Nr. 1 DS-GVO handelt, in dem der Terminus legaldefiniert wird. Relevant ist bei dieser Definition, ob eine Person mittels dieser Informationen identifizierbar ist.[5] Die in der DS-GVO zugrunde gelegte Definition wird dabei weit ausgelegt, was sich spürbar auswirkt.

Als Beispiele für personenbezogene Daten lassen sich Name, Anschrift, Telefonnummer, E-Mail-Adresse, Geburtsdatum, Standortdaten, Bild-/Filmmaterial, Kennzeichen, Fahrzeugidentifikationsnummer (FIN), (dynamische) IP-Adressen,[6] Cookies sowie Log-Files nennen, so lange diese einer bestimmten Person zuordbar sind.[7] Mit Blick auf Fahrzeugdaten ist es also entscheidend, ob die Identifizierung einer Person möglich ist.[8] Dabei werden auch Daten berücksichtigt, die mit den vorliegenden Daten noch kombiniert werden können. Diese Daten müssen nicht im Besitz desjenigen sein, der bestimmte Fahrzeugdaten hat. Relevant sind in diesem Kontext die Umstände des jeweiligen Einzelfalls sowie eine Zeit- und Kostenabschätzung. Melden sich Insassen mit einem Benutzerkonto an, so lässt sich ein Personenbezug deutlich schneller annehmen, sofern anfallende Daten mit der entsprechenden User-ID verknüpft werden.

Werden Daten umgehend gelöscht und auch nicht transferiert, so liegen keine personenbezogenen Daten vor.[9] Nicht anders verhält es sich diesbezüglich hinsichtlich des grundrechtlichen Datenschutzes.[10] Als plakatives Beispiel kann die eingestellte Innenraumtemperatur genannt werden. Hier wird die Innenraumtemperatur eingestellt und gemessen, eventuell auch ausgegeben, aber muss nicht gespeichert werden. In diesem Fall liegt kein personenbezogenes Datum vor.

4. Entwicklung automatisierter und autonomer Fahrzeuge

4.1 Einbeziehung für Datenschutz zuständiger Mitarbeiter

Bei der Entwicklung automatisierter und autonomer sowie vernetzter Fahrzeuge sind datenschutzrechtliche Vorgaben zu berücksichtigen. Es empfiehlt sich, bereits frühzeitig die für Datenschutz zuständigen Mitarbeiter in einem Unternehmen einzubeziehen, da diese dann von Beginn an eine Einschätzung liefern können, so dass sich auch die entwickelnden Ingenieure hierauf einstellen können, damit nicht später aufgrund von datenschutzrechtlichen Vorgaben

gravierende Änderungen notwendig sind. Dieser Aspekt der frühzeitigen Einbeziehung betrifft insbesondere die Unternehmensorganisation.

4.2 Datenschutzrechtliche Vorgaben bei der Entwicklung

Die Entwicklung und juristisch besehen die Konstruktion eines Produkts ist der Dreh- und Anelpunkt. Spätere haftungsrechtliche Ansprüche gründen nicht selten auf einem Konstruktionsfehler. Auch datenschutzrechtliche Anforderungen müssen in dieser frühen Phase der Technik- und Softwaregestaltung berücksichtigt werden, damit das spätere Produkt mit Anforderungen aus der DS-GVO im Einklang steht und folglich „compliant“ ist.

In der DS-GVO finden sich in Art. 25 Abs. 1, 2 die beiden Prinzipien „Privacy by Design“ und „Privacy by Default“.

Das Konzept „Privacy by Design“ findet sich in Art. 25 Abs. 1 DS-GVO. Prämisse ist, dass fundamentale Bedingungen für die Datenverarbeitung bereits durch die eingesetzte Soft- und Hardware determiniert werden.[11] Prägnant lässt sich dieser Ansatz als „Datenschutz durch Technologie als Regulierungsansatz“ beschreiben.[12] Systeme müssen daher derart ausgelegt sein, dass später lediglich eine geringe Datenverarbeitung stattfindet und der Eingriff in die Rechte des Betroffenen ebenfalls gering bleibt.[13] Es bietet sich etwa an, dass durch das System eine Pseudonymisierung stattfindet.[14] So lässt sich bereits frühzeitig eine Datenminimierung erreichen.[15]

In Art. 25 Abs. 2 DS-GVO findet sich das Prinzip „Privacy by Default“, welches an einer anderen Ebene ansetzt und zum Ziel hat, dass Unternehmen nicht so viele Daten wie möglich verarbeiten. Dies kontrahiert mit dem Aspekt, dass Daten eines der zentralen Wirtschaftsgüter unserer Zeit sind,[16] so dass konträre, insbesondere monetäre Anreize bestehen. Bei dem Ansatz „Privacy by Design“ geht es um die Datenmenge, Verarbeitungsumfang sowie Speicherdauer und Datenzugang.[17] Ein wichtiger Aspekt, um dies zu erreichen, sind die Technischen und Organisatorischen Maßnahmen (TOM).

Mit Blick auf den Umfang der personenbezogenen Daten gilt es zu beachten, dass nur diejenigen Daten verarbeitet werden, die für den jeweiligen Zweck und dessen Erreichung absolut erforderlich sind.[18]

Auch auf Benutzereinstellungen wirkt sich das Konzept „Privacy by Design“ aus. So müssen Benutzereinstellungen schon ab Werk datenschutzfreundlich voreingestellt sein. Dies hat vor allem positive Effekte für Personen, denen technische Kenntnisse fehlen, um entweder Einstellungen zu ändern oder das Ausmaß der Datenverarbeitung zu überblicken.[19]

5. Erprobung automatisierter und autonomer Fahrzeuge

5.1 Grundlage für die Erprobung nach der StVG-Novelle 2021

Bei der Erprobung automatisierter und autonomer Fahrzeuge ist zunächst zu beachten, dass seit der StVG-Novelle 2021 § 1i StVG lex specialis gegenüber § 19 Abs. 6 i.V.m. § 70 Abs. 1 Nr. 1, 2 StVZO ist. Ausgehend vom Wortlaut des § 1i Abs. 1 StVG erfasst die neu eingefügte Rechtsnorm zur Erprobung sowohl „automatisierte“ als auch „autonome“ Kraftfahrzeuge. Autonome Kraftfahrzeuge meint dabei solche i.S. des § 1d StVG, wobei der Gesetzgeber auf Level 4 des SAE J3016 Standards rekurriert. Was mit automatisiert gemeint ist, ist umstritten, da in der StVG-Novelle 2017 in den §§ 1a ff. StVG ausschließlich die Rede vom hoch- und vollautomatisierten Fahren ist. Ob der in § 1i Abs. 1 StVG verwendete Term „automatisiert“ das hoch- und vollautomatisierte Fahren erfassen soll oder sämtliche Automatisierungsstufen neben dem autonomen Fahren, ist offen, wirkt sich aber spürbar aus.

5.2 Datenschutzrechtliche Aspekte bei der Erprobung automatisierter Fahrzeuge

Für die Erprobung automatisierter und autonomer Fahrzeuge unerlässlich sind Testfahrten. Hierbei können Fahrzeugalgorithmen und das Zusammenspiel mit der Sensorik zur Umfelderfassung und anschließenden Objekterkennung sowie Prädiktion getestet werden. Die Erprobung ist aus haftungsrechtlicher Sicht besonders relevant, um spätere Konstruktionsfehler auszuschließen, aus denen eine Produkt- und Produzentenhaftung des Automobilherstellers nach dem harmonisierten Produkthaftungsgesetz (ProdHaftG) oder § 823 Abs. 1 BGB resultieren kann.[20]

Durch die Anforderungen aus der Entwicklung und dem Produkthaftungsrecht werden zugleich die anfallenden Daten(mengen) determiniert. So wird bei Testfahrten die Fahrzeugumgebung aufgezeichnet. Erfasst werden daher beispielsweise Gesichter, Passanten, Kennzeichen. Bei allen dreien handelt es sich um personenbezogene Daten, so dass es für die Rechtmäßigkeit der Verarbeitung der Daten (Erhebung, Verarbeitung, Speicherung) auf Art. 6 DS-GVO ankommt.[21] Dieser nennt die Erlaubnisträger zur Datenverarbeitung.

5.2.1 Datenverarbeitung auf Grundlage einer Einwilligung

In Art. 6 Abs. 1 lit. a DS-GVO findet sich zuerst die Einwilligung. Diese muss jedoch „für einen oder mehrere bestimmte Zwecke gegeben“ werden. Dies folgt aus dem Zweckfestlegungsgebot. Eine Legaldefinition der Einwilligung findet sich in Art. 4 Nr. 11 DS-GVO. Danach ist eine „Einwilligung“ der betroffenen Person jede freiwillig für den bestimmten Fall, in informierter Weise und unmissverständlich abgegebene Willensbekundung in Form einer Erklärung oder einer sonstigen eindeutigen bestätigenden Handlung, mit der die betroffene Person zu verstehen gibt, dass sie mit der Verarbeitung der sie betreffenden personenbezogenen Daten einverstanden ist.“ Maßgeblich hierbei ist, dass die Einwilligung freiwillig, in informierter Weise und eindeutig durch eine Erklärung erfolgt. Schriftform ist nicht erforderlich, wie auch aus Erwägungsgrund 32 DS-GVO hervorgeht.

Einwilligen können daher Testfahrer, falls dies zusätzlich zu dem bestehenden Arbeitsvertrag erforderlich ist.[22] Personen, die sich außerhalb des Fahrzeugs befinden, können indes nicht einwilligen.[23]

5.2.2 Datenverarbeitung auf Grundlage einer Vertragserfüllung

In Art. 6 Abs. 1 lit. b DS-GVO findet sich sodann die Möglichkeit der Datenverarbeitung personenbezogener Daten zur Erfüllung eines Vertrags, sofern dies erforderlich ist. Jedoch scheidet dies sowohl bei Testfahrten als auch beim regulären Betrieb aus, denn Personen außerhalb des Fahrzeugs sind nicht von etwaigen Vertragsverhältnissen des Fahrers oder der Insassen betroffen.[24]

5.2.3 Datenverarbeitung auf Grundlage eines berechtigten Interesses

Da sowohl eine Einwilligung als auch die Datenverarbeitung zur Erfüllung eines Vertrags ausscheiden, bleibt Art. 6 Abs. 1 lit. f DS-GVO. Hierbei handelt es sich um die „zentrale Interessenabwägungsklausel“ der DS-GVO, oft auch als „zentrale Stellschraube“ bezeichnet.[25] Denn die Verarbeitung basiert auf dem berechtigten Interesse des Verantwortlichen. Durch den Wortlaut des Rechtssatzes wird dieser Abwägungscharakter besonders deutlich: „Die Verarbeitung ist zur Wahrung der berechtigten Interessen des Verantwortlichen oder eines Dritten erforderlich, sofern nicht die Interessen oder Grundrechte und Grundfreiheiten der betroffenen Person, die den Schutz personenbezogener Daten erfordern, überwiegen, insbesondere dann, wenn es sich bei der betroffenen Person um ein Kind handelt.“ Soll die Datenverarbeitung auf ein berechtigtes Interesse gestützt werden, kommt es maßgeblich auf den Zweck der

Verarbeitung an. Die Frage, ob ein berechtigtes Interesse anzunehmen ist, ist normativ zu beurteilen.

Interesse meint dabei nicht nur rechtliche Interessen des Verarbeiters, sondern auch wirtschaftliche sowie ideelle Interessen. Den Verarbeiter trifft zudem die Pflicht, den Betroffenen über seine Interessen zu informieren (Art. 13 Abs. 1 lit. d, 14 Abs. 2 lit. b DS-GVO).[26] Der Betroffene hat die Möglichkeit, gemäß Art. 6 Abs. 1 lit. f DS-GVO der Datenverarbeitung zu widersprechen. Bei Testfahrten kann der im datenschutzrechtlichen Sinne Verantwortliche seiner Pflicht gegenüber dem Betroffenen am einfachsten dadurch nachkommen, dass die Fahrzeuge als „Kamerafahrzeug“ mittels Aufkleber / Magnetschilder gekennzeichnet sind. Auf diesen Labeln muss zudem eine Webseite angegeben sein, auf der Betroffene weitergehende Informationen finden. Hier müssen der Zweck der Datenverarbeitung, die Betroffenenrechte sowie weitere zwingende Informationen vorhanden sein.[27]

Ein legitimer Zweck ist beispielsweise die Erprobung von Softwarealgorithmen, da hierdurch die Erhöhung der Sicherheit des Straßenverkehrs bezweckt wird. Aber auch die Weiterentwicklung, die Entwicklung neuer Technologien oder die Überprüfung der Konstruktion im Rahmen des fortlaufenden Konstruktionsprozesses bis zum Zeitpunkt des Inverkehrbringens stellen einen möglichen Zweck dar.[28]

Bereits vor Beginn entsprechender Testfahrten muss darüber nachgedacht werden, ob Daten außerhalb der EU gespeichert werden sollen. Sollen Daten außerhalb gespeichert werden, wirkt sich dies entsprechend auf die Abwägung der jeweiligen Interessen aus. Auch die beabsichtigte Weitergabe an Konzerngesellschaften muss im Vorhinein erwähnt werden und ist ansonsten nicht möglich. Unproblematisch ist indes die Weitergabe der späteren Ergebnisse einer Analyse, wenn es sich bei diesen nicht mehr um personenbezogene Daten handelt.[29]

6. Betrieb automatisierter und autonomer Fahrzeuge

6.1 „Blackbox“ bei automatisierten Fahrzeugen

Mit der StVG-Novelle 2017 wurde für das hoch- und vollautomatisierte Fahren § 63a StVG in das Straßenverkehrsgesetz eingefügt. Dieser regelt die Datenverarbeitung bei Kraftfahrzeugen mit hoch- oder vollautomatisierter Fahrfunktion. Auf autonome Fahrzeuge ist die Rechtsnorm nicht anwendbar. Nach § 63a Abs. 1 StVG müssen Fahrzeuge i.S. des § 1a StVG Positions- und Zeitangaben speichern, wenn ein Wechsel der Fahrzeugsteuerung zwischen Fahrzeugführer und dem hoch- oder vollautomatisierten System erfolgt.[30] Wenngleich

Normadressat des § 63a StVG etwa mit Blick auf die Datenübermittlung der Fahrzeughalter ist, muss der Automobilhersteller nicht nur den Datenspeicher und das entsprechende System implementieren, sondern dem Fahrzeughalter auch ermöglichen, die Daten auszulesen und sodann zu übermitteln. Damit wird § 63a StVG für die Entwicklung von hoch- und vollautomatisierten Fahrzeugen i.S. der §§ 1a ff. StVG relevant.

6.2 „Blackbox“ bei autonomen Fahrzeugen

Mit der jüngsten StVG-Novelle 2021 zum autonomen Fahren wurde unter anderem § 1g StVG in das Straßenverkehrsgesetz eingefügt. Darin findet sich eine – im Vergleich zu § 63a StVG – detailliertere Regelung zur Datenverarbeitung bei autonomen Fahrzeugen des SAE J3016 Level 4. Gespeichert werden müssen etwa die Fahrzeugidentifikationsnummer (FIN), Positionsdaten, Umwelt- und Wetterbedingungen oder auch die Geschwindigkeit, um nur 4 von 13 nach § 1g Abs. 1 StVG zu speichernde Daten zu nennen. Wenngleich auch hier oftmals der Fahrzeughalter Normadressat ist, adressiert § 1g StVG auch Pflichten an den Hersteller autonomer Fahrzeuge. Nach § 1g Abs. 3 StVG hat der Hersteller eines autonomen Kraftfahrzeugs dieses dergestalt auszustatten, dass die Datensicherung nach § 1g Abs. 1 StVG dem Fahrzeughalter auch tatsächlich möglich ist. Er muss diesen zudem präzise und in klar und leichter Sprache über die Einstellungsmöglichkeiten zur Privatsphäre und zur Datenverarbeitung informieren. Zudem muss die Software Wahlmöglichkeiten zur Art und Weise der Speicherung und der Übermittlung der in der autonomen Fahrfunktion verarbeiteten Daten vorsehen und dem Halter entsprechende Einstellungen ermöglichen. Dies muss bereits bei der Fahrzeugkonstruktion berücksichtigt und umgesetzt werden.[31]

6.3 Rechtsgrundlage für die Datenverarbeitung beim Regelbetrieb automatisierter und autonomer Fahrzeuge

6.3.1 Datenverarbeitung von Insassen

Eine Datenverarbeitung von Insassen kann sowohl bei automatisierten als auch autonomen Fahrzeugen auf einer Einwilligung basieren. Betroffene können diese jedoch nach Art. 7 Abs. 3 DS-GVO jederzeit widerrufen. Ist die Datenverarbeitung einzig auf die Einwilligung gestützt, müssen in diesem Fall sämtliche personenbezogene Daten gelöscht werden. Hinsichtlich der Freiwilligkeit einer Einwilligung ist darauf zu achten, dass Art. 7 Abs. 4 DS-GVO ein sogenanntes Kopplungsverbot bezüglich Einwilligung und Vertragserfüllung aufstellt.[32] Sollen Insassen im Fahrzeug einwilligen, stellt sich zudem die praktische Frage, wie die notwendigen Datenschutzhinweise angezeigt werden. Berücksichtigung finden müssen unter anderem Schriftgröße und Lesbarkeit. Auch ehemals für Verarbeiter beliebte Opt-out Verfahren sind nicht

mehr zulässig.[33] Müssen personenbezogene Daten zwingend zur Erfüllung eines Vertrags verarbeitet werden, so ist Art. 6 Abs. 1 lit. b DS-GVO einschlägig, so dass die Verarbeitung hierauf gestützt werden kann. Letztlich ist auch hier die Verarbeitung auf Grundlage eines berechtigten Interesses möglich. Hier gilt indes nichts anderes, als oben ausgeführt.

6.3.2 Datenverarbeitung von Personen außerhalb des Fahrzeugs

Für den Betrieb von automatisierten und autonomen Fahrzeugen ist es erforderlich, dass Daten der Fahrzeugumgebung verarbeitet werden. Dadurch rückt die Frage in den Mittelpunkt, auf welcher Grundlage dies geschieht. Weder § 63a StVG noch § 1g StVG lassen sich hierfür heranziehen.

Zur Vereinfachung der Problematik soll angenommen werden, dass die Insassen Verantwortliche im datenschutzrechtlichen Sinne sind. Wie aufgezeigt, kommen Einwilligung und Verarbeitung zur Vertragserfüllung nicht in Betracht. Damit bliebe lediglich ein berechtigtes Interesse. Für private Fahrten könnte Art. 2 Abs. 2 Nr. c DS-GVO zu Gute kommen, wonach bei ausschließlich persönlichen und familiären Tätigkeiten der Anwendungsbereich der DS-GVO nicht eröffnet ist. Gegen die Anwendbarkeit dieser Ausnahmeregelung spricht jedoch, dass die Datenverarbeitung im öffentlichen Raum und nicht in einem abgegrenzten Privatbereich stattfindet.[34]

Wird das Kraftfahrzeug gewerblich genutzt, scheidet die Ausnahme des Art. 2 Abs. 2 lit. c DS-GVO ohnehin aus.

Auch hier mündet alles wieder bei der Frage, ob ein berechtigtes Interesse des Verantwortlichen nach Art. 6 Abs. 1 lit. f DS-GVO vorliegt. Ist in der Praxis nicht nur der Insasse Verantwortlicher, ergeben sich weitere Schwierigkeiten und das Ganze wird weitaus komplexer. Ansonsten kommt es bei einer Abwägung der berechtigten Interessen insbesondere darauf an, wie viele Daten gespeichert werden, auf Speicherdauer, Speicherort, Datenzugriff sowie mögliche Datentransfers. Jeder dieser Punkte erhöht die Intensität für den Betroffenen und erschwert – je nach Beantwortung und Ausgestaltung – ein berechtigtes Interesse des Verantwortlichen, wofür ein überwiegendes Interesse vorliegen muss.[35]

7. Fazit und Ausblick

Datenschutz nimmt bei der Entwicklung, Erprobung und dem Regelbetrieb von automatisierten und autonomen Fahrzeugen einen entscheidenden Stellenwert ein. Daher sollte frühzeitig an

die Rechtsmaterie „Datenschutz“ gedacht werden und damit einhergehend die für diese Thematik zuständigen Personen eingebunden werden. Bei der Fahrzeugkonstruktion gilt es verschiedene datenschutzrechtliche Prinzipien zu beachten, welche sich auch auf Hard- und Software auswirken.

Während die Erprobung von automatisierten und autonomen Fahrzeugen eher unproblematisch ist, bringt der Regelbetrieb erhebliche Rechtsunsicherheit mit sich. Diese wird auch von den §§ 63a, 1g StVG nicht beseitigt. Mit Blick auf die Datenverarbeitung von hoch- und vollautomatisierten sowie autonomen Fahrzeugen finden sich Anforderungen in § 63a StVG sowie § 1g StVG. Letzterer normiert zudem ausdrücklich Anforderungen und Pflichten an den Hersteller eines autonomen Fahrzeugs.

Alles in allem lässt sich festhalten, dass Datenschutz bei der Entwicklung, Erprobung und dem Regelbetrieb von großer Relevanz ist und die Entwicklung maßgeblich beeinflusst.

- [1] Ein Überblick über die zahlreichen Rechtsgebiete findet sich bei Chibanguza, K.J./Kuß, C./Steege, H., Künstliche Intelligenz, Recht und Praxis automatisierter und autonomer Systeme, Baden-Baden 2021.
- [2] Ausführlich zu den Kategorien Steege, H./Stender-Vorwachs, J., in: Chibanguza, K.J./Kuß, C./Steege, H., Künstliche Intelligenz, § 3, K. Automatisiertes Fahren – Datenschutzrecht, S. 369 Rn. 7.
- [3] Eingehend dazu Stender-Vorwachs, J./Steege, H., in: Oppermann, B./Stender-Vorwachs, J., Autonomes Fahren, 2. Aufl., München 2020, Kap. 3.6.1. Rn. 157-258.
- [4] BVerfG, Urt. v. 15.12.1983 – 1 BvR 209/83 = NJW 1984, 419.
- [5] Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 510.
- [6] EuGH, MMR 2012, 471 Rn. 51; LG Berlin, MMR 2007; 799.
- [7] Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 510.
- [8] Weichert, T., Der Personenbezug von Kfz-Daten, NZV 2017, 507, 511.
- [9] Metzger, A., Digitale Mobilität - Verträge über Nutzerdaten, GRUR 2019, 129, 131; Buchner, B., Datenschutz im vernetzten Automobil, DuD 2015, 372, 374.
- [10] Zu dieser Thematik Stender-Vorwachs, J./Steege, H., in: Oppermann, B./Stender-Vorwachs, J., Autonomes Fahren, 2. Aufl., München 2020, Kap. 3.6.1. Rn. 192 ff.

- [11] Voigt, P./von dem Bussche, A., The EU General Data Protection Regulation, Cham 2017, S. 62.
- [12] Voigt, P./von dem Bussche, A., The EU General Data Protection Regulation, Cham 2017, S. 25.
- [13] Wybitul, T./Draf, O., Projektplanung und Umsetzung der EU-Datenschutz-Grundverordnung im Unternehmen, BB 2016, 2101, 2104; Schulz, S., Privacy by Design, CR 2012, 204, 204.
- [14] Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 512.
- [15] Gierschmann, S., Was "bringt" deutschen Unternehmen die GS-DVO, ZD 2016, 51, 53.
- [16] Stender-Vorwachs, J./Steege, H., Wem gehören unsere Daten?, NJOZ 2018, 1361, 1361; Bisges, M., Personaldaten, Wertzuordnung und Ökonomie, MMR 2017, 301, 301; Fezer, K.-H., Dateneigentum, MMR 2017, 3, 3.
- [17] Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 512.
- [18] von dem Bussche, A./Zeiter, A./Brombach, T., Die Umsetzung der Vorgaben der EU-Datenschutz-Grundverordnung durch Unternehmen, Der Betrieb 2016, 1359, 1359.
- [19] Dienst, S./Falke, M., in: Bräutigam/Rücker, E-Commerce, München 2017, S. 60.
- [20] Steege, H., Auswirkungen von künstlicher Intelligenz auf die Produzentenhaftung in Verkehr und Mobilität, NZV 2021, 6, 8.
- [21] Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 511.
- [22] Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 511.
- [23] Klink-Straub, J./Straub, T., Nächste Ausfahrt DS-GVO - Datenschutzrechtliche Herausforderungen beim automatisierten Fahren, NJW 2018, 3201, 3205.
- [24] Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 511.
- [25] Albrecht, J.P., Das neue EU-Datenschutzrecht - von der Richtlinie zur Verordnung, CR 2016, 88, 91.
- [26] Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 512.
- [27] Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 512.

- [28] Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 512.
- [29] Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 512.
- [30] hierzu: Stender-Vorwachs, J./Steege, H., Kleine SIM-Karte - große Konsequenz: Automobilhersteller als TK-Anbieter?, MMR 2018, 212, 216; Stender-Vorwachs, J./Steege, H., Facetten Automatisierten Fahrens - von Datenschutz über Fahrerflucht bis zur Verkehrskontrolle, Bonner Rechtsjournal 2017, Heft 2, 114, 118.
- [31] Eingehend zu dem Inhalt und den Auswirkungen des § 1g StVG, Steege, H., Gesetzesentwurf zum autonomen Fahren (Level 4), SVR 2021, 128, 134.
- [32] Albrecht, J.P., Das neue EU-Datenschutzrecht - von der Richtlinie zur Verordnung, CR 2016, 88, 91; Kühling, J./Martini, M., Die Datenschutz-Grundverordnung: Revolution oder Evolution im europäischen und deutschen Datenschutzrecht?, EuZW 2016, 448, 451; Schantz, P., Die Datenschutz-Grundverordnung - Beginn einer neuen Zeitrechnung im Datenschutzrecht, NJW 2016, 1831, 1845; Metzger, A., Digitale Mobilität - Verträge über Nutzerdaten, GRUR 2019, 129, 131 f.
- [33] Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 511.
- [34] Eingehend Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 513.
- [35] Steege, H., Ist die DS-GVO zeitgemäß für das autonome Fahren?, MMR 2019, 509, 513.

Do you want me to shift lanes? – Investigating automated driving strategies in urban areas

**Patrick Roßner, Konstantin Felbel, Dr.-Ing. André Dettmann,
Prof. Dr. Angelika C. Bullinger,**
Chemnitz University of Technology, Chemnitz

Zusammenfassung

Der Fahrstil eines hochautomatisierten Fahrzeugs hat großen Einfluss auf die Akzeptanz und das Vertrauen der Passagiere. Studien zeigen, dass ein defensiver Fahrstil auf Autobahnen und Landstraßen einem dynamischen Fahrstil bevorzugt wird. Es gibt jedoch nur wenige Erkenntnisse über die Wahrnehmung des Fahrstils in städtischen Gebieten. Um diese Forschungslücke zu schließen, wurde eine explorative Fahrsimulatorstudie mit 42 Probanden durchgeführt. Die Probanden erlebten vier hochautomatisierte Fahrten mit 9 verschiedenen Fahrsituationen. In jeder Versuchsbedingung führte das Ego-Fahrzeug entweder eine dynamische oder eine defensive Fahrweise aus. Während jeder Testfahrt gaben die Versuchspersonen über eine Handregler kontinuierlich Rückmeldung zu ihrem Diskomfort und füllten Fragebögen zur Akzeptanz, zum Vertrauen und zur subjektiv erlebten Fahrleistung der Automatisierung aus. Für eine detaillierte Bewertung der beiden Fahrstile wurde am Ende der Studie ein Abschlussinterview durchgeführt. Insgesamt wurde eine signifikant bessere Bewertung des defensiven Fahrstils festgestellt. Die Probanden gaben signifikant höhere Werte für die Akzeptanz, das Vertrauen und die subjektiv empfundene Fahrleistung für den defensiven Fahrstil an. Durch die kontinuierliche Auswertung des Diskomforts konnte jede Situation im Nachhinein separat untersucht werden. Die Situationsanalyse zeigt hohe Diskomfortwerte für Fahrsituationen mit Quer- und Gegenverkehr, insbesondere beim dynamischen Fahrstil. Im Gegensatz dazu waren die Diskomfortwerte in Situationen mit Längsverkehr gering. In diesen Situationen wurde die dynamische Fahrweise ebenso positiv bewertet wie die defensive Fahrweise, in zwei Situationen war sie sogar erwünscht. Die Befragungen bestätigten die Ergebnisse und zeigten auch die Wünsche der Probanden nach einer situationsspezifischen Anpassung der automatisierten Fahrweise. Das Fahrzeug sollte sich je nach Verkehrssituation und persönlichen Präferenzen unterschiedlich verhalten. Die Ergebnisse der Studie tragen dazu bei, akzeptierte und vertrauenswürdige Fahrstile zu entwickeln, indem sie Empfehlungen für das Fahrverhalten in bestimmten städtischen Situationen geben.

Abstract

The driving style of a highly automated vehicle has major influence on acceptance and trust ratings of passengers. Studies show that a defensive driving style is preferred over a dynamic one in highway and rural road scenarios. However, there are only few findings on driving style perception in urban areas. To fill that research gap, an exploratory driving simulator study with 42 subjects was conducted. Subjects experienced four highly automated drives with 9 different driving situations. For each experimental condition, the Ego-vehicle performed either a dynamic or a defensive driving style. During each test drive, subjects gave continuous feedback on their discomfort level via handset controller and filled in questionnaires regarding acceptance, trust and subjectively experienced driving performance of the automation. For a detailed assessment of the two driving styles, a final interview was conducted at the end of the study. Overall, a significantly better evaluation of the defensive driving style was found. Subjects submitted significantly higher acceptance, trust and subjectively experienced driving performance ratings for the defensive driving style. Due to the continuous evaluation of the discomfort level, every situation could be examined separately afterwards. The situation-based analysis showed high discomfort ratings for driving situations with cross- and oncoming traffic, especially in the dynamic driving style. In contrast, situations with longitudinal traffic resulted in low discomfort ratings. In these situations, the dynamic driving style was rated as positive as the defensive driving style and in two situations was even desired. The interviews confirmed the results and showed also subjects wishes for a situation-specific adjustment of the automated driving style. The vehicle should behave differently depending on the traffic situation and personal preferences. The findings of the study help to develop accepted and trustworthy driving styles by providing recommendations for driving behaviour in specific urban situations.

1. State of literature and knowledge

Highly automated driving in urban environments has difficult prerequisites from both a technical as well as a human factors viewpoint. First, Sensor data from a demanding urban environment must be collected and fused [1]. Second, based on the sensory data real time motion planning is performed. This motion planning is not only dependent on safety [2] but also on aspects of users' acceptance. Ergonomic studies on highly-automated driving constitute essential aspects for later acceptance and use of highly automated vehicles [3, 4]. Therefore, the automated motion needs to be considered in the designing of such vehicles. The main starting point for a high level acceptance in highly automated driving is the design of automated motion respectively the driving style itself. Studies on highways and rural roads show, that an automated driving style should be parameterized for specific situations and manoeuvres [5], the

situation in urban environments is yet unclear. Hence, the present study focused on investigating automated driving strategies in urban areas. From the aforementioned studies it can be shown, that subjects prefer their own or a very similar driving style and reject other driving styles that include e.g. very high acceleration and deceleration rates or small longitudinal and lateral distances to other road users [6, 7]. Also, a swift, anticipatory, safe and naturally-looking driving style is prioritized, but no overall accepted driving style is found yet [8, 9, 10]. In the present study, we try to fill the identified research gap by investigating two different automated driving styles in typical urban traffic situations. Based on previous experiments [6, 11], the driving situations were divided into its basic elements and implemented in a driving simulator.

2. Method and variables

A fixed-based driving simulator with a programmable automated driving function (Fig. 1) was used to conduct the experiment. Three projectors with a resolution of 1400 x 1050 pixels each were used to simulate the driving scene with the simulation software SILAB version 5.1. The study was conducted with 42 subjects, 22 younger and 20 older participants. They experienced four automated drives, to times the offensive and two times the defensive driving style, in an urban scenario on a two-lane road. The drives were presented in a randomized order and the driving speed was constantly set to 70 km/h. At all times, participants could oversteer the automation.

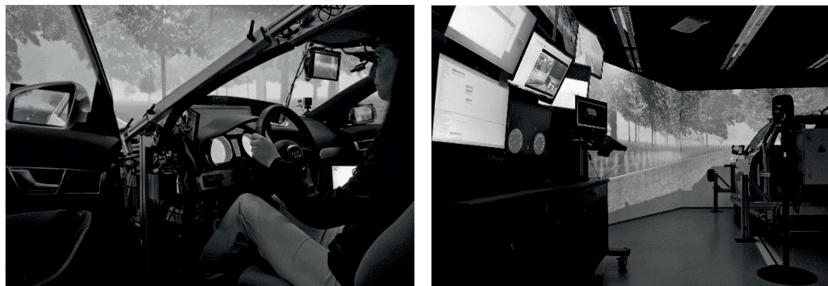


Fig. 1: Driving simulator with instructor centre (left) and an exemplary subject (right)

Subjects were at least 21 years old and had a valid driver's licence (see Table 1 for details). Due to one technical problem and one case of simulator sickness two participants were excluded of further analysis. The test track consisted of 9 different urban traffic situations that required a reaction of the automated vehicle (see Table 2 for details).

Table 1: Subject characteristics

	#	Age		Driver's license holding [years]		Mileage last five years [km]	
		N	M	SD	M	SD	M
female younger	10	29.1	5.4		11.7	5.3	32.760
male younger	10	29.5	5.9		11.8	6.4	78.300
female older	10	69.5	4.3		46.8	6.4	25.100
male older	10	69.3	4.5		51.2	4.0	92.700
total	40	49.4	5.0		30.4	5.3	57.215
							44.967

Table 2: Description of the driving situations

#	Situation description
1	merging bus out of bus stop without following traffic to the rear left of the Ego-vehicle
2	traffic light approach with following traffic to the left of the Ego-vehicle; red phase = 1s
3	traffic light approach with following traffic to the left of the Ego-vehicle; red phase = 5s
4	merging bus out of bus stop with following traffic to the rear left of the Ego-vehicle
5	traffic light approach; Traffic changes state from green to yellow momentarily
6	traffic light approach without following traffic to the left of the Ego-vehicle; red phase = 1s
7	Cooperative intersection / oncoming traffic; Ego-vehicle needs to change the direction
8	traffic light approach without following traffic to the left of the Ego-vehicle; red phase = 5s
9	Cooperative intersection / traffic from left side – Ego-vehicle needs to give way

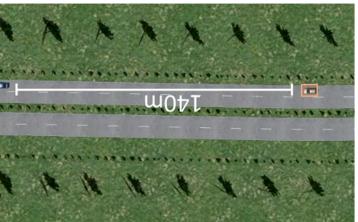
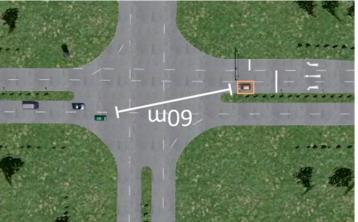
The driving styles were designed differently regarding the longitudinal and lateral trajectory. The offensive driving style accelerated quickly, changed lanes in situation 1 – 4, 6 and 8 and kept a short distance to road users in front within legal regulations. The defensive driving style accelerated slowly, did not change lanes and generally kept a greater distance to other road users. Table 3 shows the experienced traffic situations through a bird's eye view. The highlighted vehicle displays the Ego-Vehicle. At a specific distance to other road users or infra-

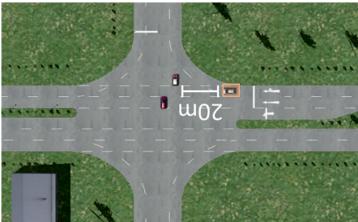
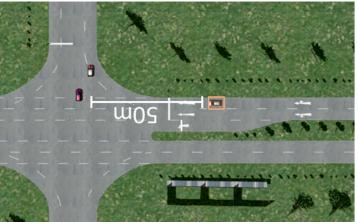
structure, the Ego-Vehicle either changed the lane (offensive driving style) or decelerated (defensive driving style). In the offensive driving style condition, the automated Ego-Vehicle decelerated after the lane change was completed.

Table 3: Birds Eye View of the driving situations with trigger distances

	Situation 1	Situation 2	Situation 3	Situation 4
Trigger distance for the offensive driving manoeuvre (Lane change)				

Trigger distance
for the defensive
driving manoeuvre
(breaking)

	Situation 5	Situation 6	Situation 7	Situation 8
Trigger distance for the offensive driving manoeuvre (Lane change)	 A 3D-style diagram showing a road with a dashed line. A car is on the left lane, and a truck is on the right lane. A horizontal line with arrows at both ends spans the distance between the two vehicles, labeled '30m'. A small orange square marks the point where the manoeuvre begins.  A 3D-style diagram showing a road with a dashed line. A car is on the left lane, and a truck is on the right lane. A horizontal line with arrows at both ends spans the distance between the two vehicles, labeled '50m'. A small orange square marks the point where the manoeuvre begins.  A 3D-style diagram showing a road with a dashed line. A car is on the left lane, and a truck is on the right lane. A horizontal line with arrows at both ends spans the distance between the two vehicles, labeled '60m'. A small orange square marks the point where the manoeuvre begins.  A 3D-style diagram showing a road with a dashed line. A car is on the left lane, and a truck is on the right lane. A horizontal line with arrows at both ends spans the distance between the two vehicles, labeled '50m'. A small orange square marks the point where the manoeuvre begins.	 A 3D-style diagram showing a road with a dashed line. A car is on the left lane, and a truck is on the right lane. A horizontal line with arrows at both ends spans the distance between the two vehicles, labeled '140m'. A small orange square marks the point where the manoeuvre begins.  A 3D-style diagram showing a road with a dashed line. A car is on the left lane, and a truck is on the right lane. A horizontal line with arrows at both ends spans the distance between the two vehicles, labeled '60m'. A small orange square marks the point where the manoeuvre begins.  A 3D-style diagram showing a road with a dashed line. A car is on the left lane, and a truck is on the right lane. A horizontal line with arrows at both ends spans the distance between the two vehicles, labeled '160m'. A small orange square marks the point where the manoeuvre begins.  A 3D-style diagram showing a road with a dashed line. A car is on the left lane, and a truck is on the right lane. A horizontal line with arrows at both ends spans the distance between the two vehicles, labeled '90m'. A small orange square marks the point where the manoeuvre begins.		
Trigger distance for the defensive driving manoeuvre (breaking)				

Situation 9	
Trigger distance for the offensive driving manoeuvre (Lane change)	
Trigger distance for the defensive driving manoeuvre (breaking)	

The participants were required to observe the driving as a passenger of an automated car. Subjects were informed via display of the oncoming manoeuvres. During the drive subjects' main feedback tool was an online handset control to measure perceived discomfort as shown in Fig. 2. This tool provides information about the occurrence of safety concerns in each location of the track and could be recorded in sync with video, eye-tracking, physiological or driving data [12]. After each experimental condition subjects filled in questionnaires regarding acceptance [13], trust in automation [14] and subjectively experienced driving performance [15] and were interviewed at the end of the study.

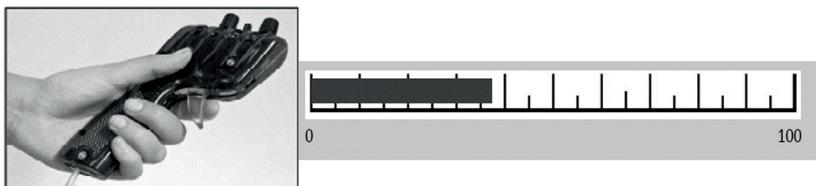


Fig. 2: Handset control (left) and visual feedback (right) to measure perceived discomfort while driving highly automated. Higher values indicate higher perceived discomfort.

3. Results

Both driving styles are evaluated positively regarding the main questionnaire constructs acceptance, trust and subjectively experienced driving performance (see Fig. 3). Although acceptance ratings show participant's high usefulness and satisfaction ratings in one summarised item for both age groups and driving styles (defensive driving style: $M = 1,04$; offensive driving style: $M = 0,81$), the defensive driving style is evaluated significantly better, $F(1, 36) = 9.766$, $p = .004$, $\eta_p^2 = .213$. Trust values illustrate a similar evaluation behaviour. Both driving styles are trusted generally with high trust values (defensive driving style: $M = 5.58$; offensive driving style: $M = 4.75$). Again, the defensive driving style is evaluated significantly better, $F(1,36) = 26.839$, $p = .000$, $\eta_p^2 = .427$.

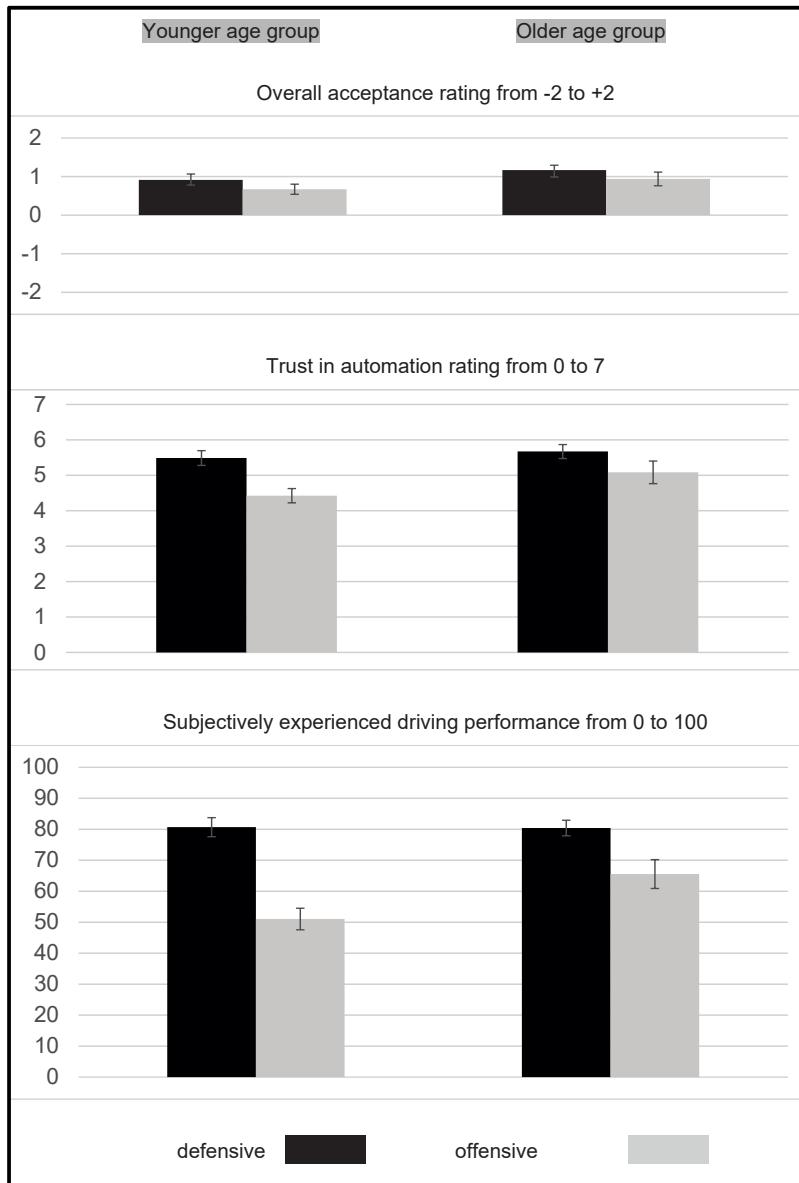


Fig. 3: Mean values of acceptance, trust and subjectively experienced driving performance.
Error bars display ± 1 SE.

Ratings for subjectively experienced driving performance show similar results (defensive driving style: $M = 80.55$; offensive driving style: $M = 58.30$). The defensive driving style is perceived as significantly more pleasant, $F(1,36) = 43.703$, $p = .000$, $\eta_p^2 = .548$. All three significant differences indicate a high effect size, which shows a strong preference for the defensive driving style. There is a tendency to age depended ratings in the evaluation of acceptance, trust and subjectively experienced driving performance, but only the evaluation of subjectively experienced driving performance shows significant differences of the factor age group, $F(1,36) = 4.816$, $p = .035$, $\eta_p^2 = .118$.

Furthermore, the online handset control delivers more detailed information, because participants were able to give continuous feedback of their perceived discomfort during the drive.

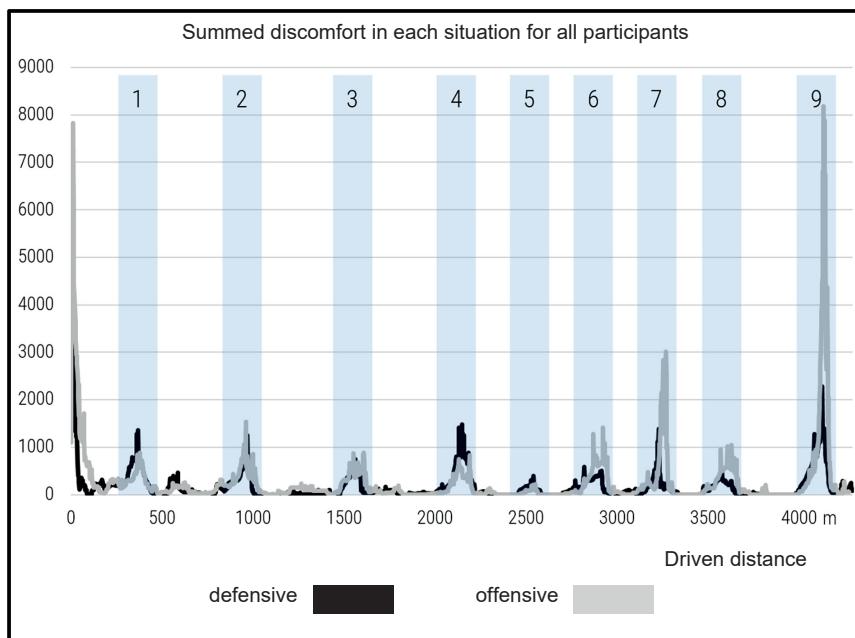


Fig 4: Summed perceived discomfort for each driving style over the entire test track

These data help to identify comfort-critical traffic situations or driving manoeuvres of each subject. Fig. 4 shows the summed perceived discomfort for each driving style and all subjects over

the entire test track. All 9 driving situations, that were performed with situation-specific manoeuvres, are marked. When comparing the two driving styles, the evaluation of perceived discomfort only differs visibly in situations 7 and 9, whereby the greatest difference is seen in situation 9. In these two situations, the automated vehicle had to cross an intersection cooperatively, whereby the driving manoeuvre of the offensive driving style produced a noticeably higher level of discomfort (Fig. 5). A possible conclusion is that an automated vehicle should behave more defensively in situations with lateral or oncoming traffic. In situations with exclusively longitudinal traffic, offensive driving manoeuvres are also a possible implementation scenario (situations 1 and 4: a merging bus; situations 2, 3, 5, 6 and 8: approaching a red traffic light with traffic in front of the Ego-vehicle). In situations 1 and 4 (merging bus), the defensive driving style with its reserved manoeuvre even caused more perceived discomfort.



Fig. 5: Cooperative crossing: left Situation 7 - offensive driving style: turns into the intersection before oncoming traffic, defensive driving style turns into the intersection after oncoming traffic; right Situation 9 - offensive driving style brakes late and only slightly before crossing traffic, defensive driving style brakes hard and early before crossing traffic.

4. Conclusion and outlook

The aim of the study was to investigate different driving styles in an urban scenario to better understand people's preferences regarding driving styles that do have an essential influence on the use of an automated vehicle [4]. The study focused on the investigation of lane change manoeuvres. Data from the studies' questionnaire revealed significantly higher acceptance, trust and subjectively experienced driving performance ratings for the defensive driving style. There is a tendency to age depended ratings in the evaluation, but only the evaluation of subjectively experienced driving performance shows significant differences of the factor age group. Additionally, the handset control data delivers a more situation-specific analysis. An automated vehicle should behave more defensively in situations with lateral or oncoming traffic. In situations with exclusively longitudinal traffic, offensive driving manoeuvres are also a possible implementation scenario, in two situations even a better solution. All these findings go along with previous study results [5 - 9]. In order to better understand the impact of the different aspects, further inference statistical analysis of the handset control data is going to be conducted. Furthermore, we found out that metrics distinguishing between driving styles gathered through real-world driving also function in a static driving simulator for implementing highly-automated driving styles [6, 8]. We assume that the defensive driving style can be used as a basic, but adjustable driving style in further studies. It remains unclear, whether subjects want to adapt the driving style in specific driving situations [as seen in 7 and 9]. A basic driving style with customization options or selection of different modes has the potential to positively influence driving experience and therefore needs to be investigated in more detail.

Acknowledgement

This research was partially supported by Sächsische Aufbaubank (SAB) in cooperation with the European Union (EU) and its funds for regional development (research project: SynCar, funding code: 100293982) as well as by the Deutsche Forschungs-gemeinschaft (DFG, German Research Foundation) – Project-ID 416228727 – SFB 1410. The sponsor had no role in the study design, the collection, analysis and interpretation of data, the writing of the report, or the submission of the paper for publication. We are very grateful to Marty Friedrich, Maximilian Hentschel, Madeleine Bankwitz and Paul Marcion for their assistance with data collection and analysis.

5. References

1. Liu, S., Tang, J., Zhang, Z. and Gaudiot, J.-L. "Computer architectures for autonomous driving," Computer, vol. 50, no. 8, pp. 18–25, 2017.
2. Artuñedo, A., Villagra A. and Godoy, J. "Real-Time Motion Planning Approach for Automated Driving in Urban Environments," in IEEE Access, vol. 7, pp. 180039-180053,2019, doi: 10.1109/ACCESS.2019.2959432.
3. Banks, V. A., & Stanton, N. A.: Keep the driver in control: Automating automobiles of the future. *Applied Ergonomics* (2015)
4. Elbanhawi, M., Simic, M., & Jazar, R.: In the Passenger Seat: Investigating Ride Comfort Measures in Autonomous Cars. *IEEE Intelligent Transportation Systems Magazine*, 7(3), 4–17 (2015) doi:10.1109/MITS.2015.2405571.
5. Dettmann A, Hartwich F, Roßner P, Beggiato M, Felbel K, Krems J, Bullinger AC. Comfort or Not? Automated Driving Style and User Characteristics Causing Human Discomfort in Automated Driving. *International Journal of Human–Computer Interaction*:331–339. 2021. doi:10.1080/10447318.2020.1860518.
6. Festner, M., Baumann, H., & Schramm, D.: Der Einfluss fahrfremder Tätigkeiten und Ma-növerlängsdynamik auf die Komfort- und Sicherheitswahrnehmung beim hochautomatisier-ten Fahren. 32. VDI/VW- Gemeinschaftstagung Fahrerassistenz und automati-siertes Fahren, Wolfsburg (2016)
7. Griesche, S.; Nicolay, E.; Assmann, D.; Dotzauer, M.; Kähnner, D.: "Should my car drive as I do? What kind of driving style do drivers prefer for the design of automated driving func-tions?" Contribution to 17th Braunschweiger Symposium Automatisie-rungssysteme, Assis-tenzsysteme und eingebettete Systeme für Transportmittel (AAET), ITS automotive nord e.V., ISBN 978-3-937655-37-6, pp. 185–204 (2016)
8. Bellem, H., Schönenberg, T., Krems, J. F., & Schrauf, M.: Objective metrics of comfort: Developing a driving style for highly automated vehicles. *Transportation Research Part F: Traffic Psychology and Behaviour*, 41, pp. 45–54 (2016)
9. Hartwich, F., Beggiato, M., Dettmann, A., & Krems, J. F.: Drive me comfortable: Cus-tom-ized automated driving styles for younger and older drivers. 8. VDI-Tagung „Der Fahrer im 21. Jahrhundert“ (2015)

10. Felbel K., Dettmann A., Lindner M., Bullinger A.C. (2021) Communication of Intentions in Automated Driving – the Importance of Implicit Cues and Contextual Information on Freeway Situations. In: Krömker H. (eds) HCI in Mobility, Transport, and Automotive Systems. HCII 2021. Lecture Notes in Computer Science, vol 12791. Springer, Cham. https://doi.org/10.1007/978-3-030-78358-7_17
11. Wu, B.-F., Yeh, C.-H., & Chen, Y.-H.: Driving behaviour-based event data recorder. IET Intelligent Transport Systems, 8(4), 361–367 (2014)
12. Hartwich, F., Beggiato, M., Dettmann, A., & Krems, J. F.: Drive me comfortable: Custom-ized automated driving styles for younger and older drivers. 8. VDI-Tagung „Der Fahrer im 21. Jahrhundert“ (2015)
13. Van der Laan, J.-D., Heino, A. & Waard, D. de. (1997). A simple procedure for the assessment of acceptance of advanced transport telematics. Transportation Research Part C: Emerging Technologies, 5 (1), 1-10.
14. Jian, J.-Y., Bisantz, A.-M. & Drury, C.-G. (2000). Foundations for an Empirical Determined Scale of Trust in Automated Systems, 53-71.
15. Voß, G.-M.-I. & Schwalm, M. (2017). Bedeutung kompensativer Fahrerstrategien in Kontext automatisierter Fahrfunktionen (Berichte der Bundesanstalt für Straßenwesen F, Fahrzeugtechnik, Bd. 118). Bremen: Fachverlag NW.

Nachweis der Betriebsbewährung automatisierter und autonomer Fahrzeuge

Proof of operational reliability of automated and autonomous vehicles

U. Steininger, TÜV SÜD, München;
J. Mazzega, S. Witkowski, Prof. T. Form,
Volkswagen, Wolfsburg;
Prof. K. Lemmer,
Deutsches Zentrum für Luft- und Raumfahrt, Braunschweig

Zusammenfassung

Das vom BMWi geförderte Verbundprojekt PEGASUS lieferte herstellerübergreifende Methoden für die Absicherung automatisierter Fahrfunktionen, erarbeitete ein einheitliches Vorgehen und setzte eine Werkzeugkette exemplarisch um. Ein wichtiger Arbeitsinhalt war die Klärung der notwendigen und hinreichenden Voraussetzungen für das Inverkehrbringen. Ein validier statistischer Nachweis, dass die Fahrzeuge die entsprechenden Voraussetzungen vollständig erfüllen und sich im realen Betrieb bewähren, kann allerdings vor Markteinführung nicht erbracht werden. Der Nachweis der Betriebsbewährung ist daher für die gesellschaftliche Akzeptanz und damit für die erfolgreiche Einführung und Verbreitung der Technologie von esenzieller Bedeutung. Nationale und internationale Gesetze, Verordnungen und Normen tragen dem bereits Rechnung. In dem Beitrag werden konkrete Anforderungen an und Lösungsansätze für die zukünftige Produktbeobachtung und Marktüberwachung automatisierter und autonomer Fahrzeuge abgeleitet. Die Autoren haben dem BMVI die Einsetzung einer Kommission empfohlen, die die organisatorischen und technischen Voraussetzungen für den Nachweis der Betriebsbewährung schafft und die Markteinführung unternehmens- und technologieunabhängig begleitet. Ziel des Beitrags ist es u.a., der Kommission Arbeitsgrundlagen zum Nachweis der Betriebsbewährung an die Hand zu geben.

Abstract

The joint project PEGASUS, funded by the BMWi, provided cross-manufacturer methods for safeguarding automated driving functions, developed a uniform procedure and implemented a tool chain as an example. An important part of the work was the clarification of the necessary

and sufficient requirements for market introduction. However, valid statistical proof that the vehicles completely fulfil the corresponding prerequisites and prove themselves in real operation cannot be provided before market launch. Proof of probation in the field is therefore essential for social acceptance and thus for the successful introduction and dissemination of the technology. National and international laws, regulations and standards already take this into account. In the article, certain requirements and solutions for the future product monitoring and market surveillance of automated and autonomous vehicles are derived. The authors have recommended that the BMVI set up a commission to create the organisational and technical prerequisites for proving operational reliability and to accompany the market launch independently of companies and technologies. The aim of the contribution is, among other things, to provide the Commission with a working basis for proving the operational probation.

1 Ausgangssituation

Das vom Bundesministerium für Wirtschaft und Energie (BMWi) geförderte Verbundprojekt PEGASUS lieferte herstellerübergreifende Methoden für die Absicherung automatisierter Fahrfunktionen, erarbeitete ein einheitliches Vorgehen und setzte eine Werkzeugkette exemplarisch um [1]. Die Ergebnisse wurden zunächst im nationalen Rahmen erarbeitet, auf internationaler Ebene diskutiert, in den Nachfolgeprojekten vertieft und finden zum Teil bereits Berücksichtigung in nationalen und internationalen Gesetzen, Verordnungen und Normen. Ein wichtiger Arbeitsinhalt war die Klärung der notwendigen und hinreichenden Voraussetzungen für das Inverkehrbringen von automatisierten Fahrzeugen.

Eine notwendige Voraussetzung für das Inverkehrbringen von Produkten besteht darin, dass sie den Anforderungen an die Produktsicherheit genügen. Den gesellschaftlichen Konsens, was als hinnehmbares, erlaubtes Risiko gilt, regelt das Produktsicherheitsgesetz (ProdSG)¹. Gemäß §5(2) ProdSG darf davon ausgegangen werden, dass ein Produkt, das Normen oder anderen technischen Spezifikationen entspricht, den Anforderungen an die Produktsicherheit genügt, soweit diese von den betreffenden Normen oder anderen technischen Spezifikationen abgedeckt sind. Die für Fahrzeuge mit elektrischen und elektronischen (E/E) Systemen zum automatisierten und autonomen Fahren anzuwendenden Normen sind

¹ Internationale Abkommen, Richtlinien, Verordnungen und Regelungen, nationale Gesetze und Verordnungen sowie Normen, die durch die Verweise im Text eindeutig gekennzeichnet sind, werden in diesem Beitrag nicht gesondert referenziert. Wenn keine weiteren Angaben zur Ausgabe gemacht werden, beziehen sich die Verweise immer auf die aktuelle Ausgabe.

- ISO 26262:2018 für die funktionale Sicherheit,
- ISO/PAS 21448:2019 für die Sicherheit der Sollfunktion (SOTIF) und
- ISO/SAE 21434:2021 für die Cybersicherheit.

Die Einhaltung dieser Normen gewährleistet, dass von der bestimmungsgemäßen oder vorhersehbaren Verwendung des Produkts kein unvertretbares Risiko ausgeht.

Als hinreichende Voraussetzung formuliert die vom Bundesministerium für Verkehr und Digitale Infrastruktur (BMVI) eingesetzte Ethik-Kommission Automatisiertes und Vernetztes Fahren in ihrem Bericht aus dem Jahr 2017 [2],

- dass automatisiertes Fahren vertretbar ist, wenn es im Vergleich zu menschlichen Fahrleistungen eine Verminderung von Schäden im Sinne einer positiven Risikobilanz verspricht und
- dass bei Vorliegen einer grundsätzlich positiven Risikobilanz technisch unvermeidbare Restrisiken einer Einführung nicht entgegenstehen.

2 Maßnahmen vor Inverkehrbringen

Im vorangegangenen Kapitel wurde die normenkonforme Entwicklung der Systeme als notwendige Voraussetzung für deren Inverkehrbringen identifiziert. Dies schließt bereits Maßnahmen zur Absicherung und Erprobung der Systeme ein. Diese Maßnahmen müssen gewährleisten, dass die Systeme mindestens menschliche Fahrleistung erreichen und die Automationsrisiken lediglich technisch unvermeidbare Restrisiken darstellen.

Für die Absicherung konventioneller Fahrzeugsysteme hat es sich bewährt, eine überschaubare Anzahl von Testfahrzeugen dem Strecken- und Beanspruchungsprofil auszusetzen, dem sie im Laufe ihres Fahrzeuglebens vorhersehbar standhalten müssen. Dieser streckenbasierte Ansatz funktioniert bei Systemen zum automatisierten und autonomen Fahren grundsätzlich nicht, weil deren bestimmungsgemäße Funktion sich nicht durch gefahrene Kilometer nachweisen lässt, sondern dadurch, dass sie konkreten Verkehrssituationen ausgesetzt werden, die sie bewältigen sollen. Wir sprechen hier von einem szenarienbasierten Ansatz. Die Herausforderung besteht in der Gewährleistung der Vollständigkeit der Szenarien bzw. ihrer Repräsentativität hinsichtlich der abzusichernden Funktionen. Mit der Komplexität der Funktionen steigt die Zahl von Szenarien, die zu testen sind. Das Problem lässt sich durch den Einsatz virtueller Tests lösen. Neben Hardware-, Software- oder Vehicle-in-the-Loop-Tests geht es dabei vor allem um Simulationsanalysen. Wenn die erforderlichen Simulationsmodelle vorliegen und die Parameterräume definiert sind, können die Testfälle automatisch erzeugt und so nahezu beliebig große Testräume abgedeckt werden. Reale Tests auf Prüfgeländen und im

Feld werden dadurch keineswegs überflüssig. Sie dienen vielmehr der Verifikation der Ergebnisse der virtuellen Tests und der Validierung der zu Grunde gelegten Szenarien. Insbesondere muss dabei entlang der Systemgrenzen getestet werden. Die Anzahl der realen Tests kann durch das beschriebene Testkonzept in einem durchführbaren Rahmen gehalten werden. Eine entsprechende Vorgehensweise zur Absicherung des automatisierten Fahrens wurde in [3] vorgestellt und anschließend im Rahmen von PEGASUS [1] herstellerübergreifend erarbeitet und exemplarisch umgesetzt.

Komplizierter gestaltet sich die Situation beim Versuch nachzuweisen, dass die Systeme eine positive Risikobilanz versprechen. Der von der Ethik-Kommission verwendete Begriff der Risikobilanz macht deutlich, dass es hier nicht um ein einzelnes Fahrzeug oder Fahrzeugsystem geht, sondern um den Einfluss der Einführung des automatisierten und autonomen Fahrens auf die Sicherheit des Straßenverkehrs in seiner Gesamtheit. Wir sind also zunächst darauf angewiesen, dass Experten und Institutionen im Ergebnis möglichst umfassender a priori Be trachtungen positive Erwartungen ableiten. Exemplarisch und ohne jeden Anspruch auf Vollständigkeit soll hier auf die Literaturstellen [4] bis [8] verwiesen werden, in denen sich Vertreter von Regierungen, Behörden und Industrieverbänden entsprechend artikulieren.

3 Produktbeobachtung und Marktüberwachung

Ein valider statistischer Nachweis, dass automatisierte und autonome Fahrzeuge die vorgenannten Voraussetzungen vollständig erfüllen und sich im realen Betrieb bewähren, lässt sich vor Markteinführung nicht erbringen. Der Nachweis der Betriebsbewährung ist aber für die individuelle und gesellschaftliche Akzeptanz und damit für die erfolgreiche Einführung und Verbreitung der Technologie von essenzieller Bedeutung. Diesem Nachweis dienen im Wesentlichen die heute schon im Produktsicherheitsgesetz geforderten Maßnahmen zur Produktbeobachtung durch die Hersteller und die Marktüberwachung durch die dafür zuständige Behörde. In Deutschland ist das das Kraftfahrt-Bundesamt (KBA).

Darüber hinaus hält das Thema Einzug ins Zulassungsrecht und gewinnt in einschlägigen Normen an Bedeutung. Beispielhaft anzuführen sind hier folgende Aktivitäten.

- Auf europäischer Ebene trat zum 01.09.2020 die Verordnung (EU) 2018/858 über die Genehmigung und die Marktüberwachung von Kraftfahrzeugen und Kraftfahrzeuganhängern sowie von Systemen, Bauteilen und selbstständigen technischen Einheiten für diese Fahrzeuge in Kraft. Sie unterscheidet sich von der EG-Typgenehmigungsrichtlinie 2007/46/EG, die sie ersetzt, durch entsprechende Regelungen zur Marktüberwachung.

- Die Wirtschaftskommission für Europa der Vereinten Nationen (UNECE) erarbeitet derzeit unter der Bezeichnung „New Assessment/Test Method for Automated Driving (NATM)“ einen Ansatz zur Absicherung des automatisierten Fahrens. Die Marktüberwachung als so genanntes „In-Service Monitoring and Reporting“ ist ein wesentlicher Bestandteil dieses Ansatzes [9].
- Im nationalen Gesetz zum autonomen Fahren in festgelegten Betriebsbereichen vom 12.06.2021, das am 27.07.2021 im Bundesgesetzblatt Teil 1 Nr. 48 veröffentlicht wurde, wird eine Marktüberwachung explizit gefordert. In der zugehörigen Rechtsverordnung, die sich noch in Bearbeitung befindet, soll diese detailliert geregelt werden.
- Im aktuellen Normenentwurf ISO/PAS 21448:2019 wird ein Prozess zur Überwachung der Sicherheit der Sollfunktion im Feld in Abhängigkeit vom Automatisierungsgrad und von der Komplexität der vorgesehenen Funktionalität gefordert.

4 Grenzen bestehender Maßnahmen

Neben der Produktbeobachtung durch die Hersteller und die Marktüberwachung durch das KBA verfügen wir in Deutschland mit GIDAS (German In-Depth Accident Study) über eines der weltweit größten Projekte der herstellerunabhängigen Verkehrsunfallforschung [10]. Hier werden reale Verkehrsunfälle dokumentiert, rekonstruiert und der Unfalleinlauf separat simuliert². Die Analysen bieten dem Gesetzgeber die Möglichkeit, das Unfallgeschehen genau zu beobachten und positive sowie negative Entwicklungen festzustellen. Automobilhersteller und Zulieferer können anhand der Ergebnisse ihre eigene Technik verbessern. Darüber hinaus führt auch die Versicherungswirtschaft eigene systematische Unfallanalysen durch. Flankiert und ergänzt werden Unfallanalysen z.B. durch Fahrstudien im Feld und in Fahrsimulatoren. Grundsätzlich ist also alles vorhanden, was für den Nachweis der Betriebsbewährung benötigt wird.

² Seit Juli 1999 werden in GIDAS in den Großräumen Hannover und Dresden pro Jahr etwa 2.000 Verkehrsunfälle mit Personenschäden erhoben. Durch ein definiertes Stichprobenverfahren ist die Repräsentativität der Daten im Vergleich zur Bundesstatistik gewährleistet. Träger des Projektes sind die Bundesanstalt für Straßenwesen (BAST) und die Forschungsvereinigung Automobiltechnik e.V. (FAT). Die Unfallerhebungen und -rekonstruktionen werden von der Verkehrsunfallforschung (VUFO) der TU Dresden und der Unfallforschung der Medizinischen Hochschule Hannover durchgeführt.

Im Hinblick auf die bevorstehende Einführung von Systemen zum automatisierten und autonomen Fahren weisen die vorgenannten Maßnahmen allerdings systematische Schwachstellen auf. Das hat u.a. folgende Ursachen.

- (1) Die Modelle, die für Prognosen zur Wirksamkeit von technischen Systemen und zur Rekonstruktion der Unfälle herangezogen werden, sind darauf ausgerichtet, das Handeln „menschlicher“ Verkehrsteilnehmer abzubilden. Sie sind nicht ohne weiteres geeignet, um automatisierte und autonome Fahrfunktionen zu bewerten. Das wird besonders deutlich, wenn man selbst von Experten immer wieder hört, dass der Mensch über 90% der heutigen Verkehrsunfälle verursache, von denen die Mehrzahl wegfallen würden, wenn technische Systeme die Fahraufgabe übernehmen [11]. In der Tat werden technische Systeme niemals unaufmerksam und können Routineaufgaben mit größerer Präzision ausführen als Menschen. Dabei darf aber nicht übersehen werden, dass Menschen durch ihre Fähigkeit zu antizipieren und sich auf neue und komplexe Situationen einzustellen auch bewirken, dass kritische Situationen im Straßenverkehr überwiegend gar nicht zu Unfällen führen. Ob bzw. wann technische Systeme das in vergleichbarer Qualität leisten können, muss sich zeigen. Dazu kommt, dass in den nächsten Jahrzehnten Fahrzeuge mit verschiedensten automatisierten und autonomen Fahrfunktionen und mit jeweils ganz unterschiedlichen Marktanteilen miteinander, mit anderen Verkehrsteilnehmern und mit der sich ebenfalls weiterentwickelnden Verkehrsinfrastruktur in Wechselwirkung treten werden. Eine große Bedeutung wird daher der Frage zukommen, ob automatisierte und autonome Fahrzeuge sich so verhalten, wie es andere Verkehrsteilnehmer erwarten. Und neben der durchaus berechtigten Erwartung, dass diese Fahrzeuge eine Reihe von Unfällen vermeiden oder zumindest ihre Folgen abmildern, wird es auch Ereignisse geben, die ohne die Einführung der Technologie gar nicht aufgetreten wären (Automationsrisiken). Insbesondere diese Ereignisse gilt es, zuverlässig zu erfassen, um erforderlichenfalls entgegenzuwirken.
- (2) Die Entwicklung von E/E Systemen in Kraftfahrzeugen ist bisher darauf ausgerichtet, systematische Fehler auszuschließen sowie zufällige Fehler auf ein hinnehmbares Maß zu reduzieren, ihr Auftreten zuverlässig zu erkennen und die Systeme in diesem Fall möglichst schnell in einen sicheren Zustand zu überführen. Bei hinreichend komplexen Systemen können Gefährdungen auch entstehen, ohne dass zuvor ein Fehler im herkömmlichen Sinne aufgetreten ist. Das ist z.B. der Fall, wenn Systeme an die Grenzen ihrer sensorischen Wahrnehmung oder ihrer algorithmischen Fähigkeiten zur Situationsanalyse und

Verhaltensvorhersage stoßen und die Sicherheit der Sollfunktion dann nicht mehr gewährleistet ist. Die für Systeme zum automatisierten und autonomen Fahren typischen technischen Limitierungen und die daraus resultierenden Begrenzungen ihrer Anwendungsbeziehe werden bisher bei der Produktbeobachtung und Marktüberwachung nicht oder zumindest nicht systematisch berücksichtigt. Indikatoren für die Erkennung möglicher Unzulänglichkeiten automatisierter und autonomer Fahrfunktionen im Feld fehlen vollständig.

- (3) Ereignisse mit schwerwiegenden Folgen treten im Straßenverkehr nur selten auf. Zwischen zwei Unfällen mit Todesfolge auf deutschen Autobahnen liegen z.B. rund 660 Millionen Fahrkilometer [12]. Das damit verbundene Risiko liegt damit weit unterhalb dessen, was ein einzelnes Individuum praktisch wahrzunehmen in der Lage wäre. Wenn sich die Erwartungen an das automatisierte und autonome Fahren bestätigen, sollte die Unfallhäufigkeit insgesamt abnehmen und die mit der Einführung der Technologie einhergehenden Automationsrisiken sollten sehr selten zu schwerwiegenden Ereignissen führen (d.h. noch seltener, als wir es heute beobachten). Ungeachtet dessen haben sehr seltene, aber schwerwiegende Ereignisse u.U. große Auswirkungen auf die individuelle und gesellschaftliche Akzeptanz insbesondere neuer Technologien und können deren Verbreitung massiv entgegenwirken oder sie sogar vollständig vom Markt verdrängen. Mit den bisherigen Analyseumfängen in Unfallforschung, Produktbeobachtung und Marktüberwachung werden sehr seltene Ereignisse nur unzuverlässig erfasst bzw. erst relativ spät erkannt. Eine schnelle Reaktion auf entsprechende Ereignisse ist dann nicht mehr möglich. Paradoxerweise würde das hohe Maß an Sicherheit im heutigen Straßenverkehr so zur Hürde für den Nachweis, dass es sich bei den verbleibenden Risiken um technisch nicht vermeidbare Restrisiken handelt, die der Einführung der Technologie nicht im Wege stehen.

5 Methodische Lösungsansätze

Wie wir gesehen haben, ergeben sich systematische Schwachstellen für den Nachweis der Betriebsbewährung von Systemen zum automatisierten oder autonomen Fahren u.a. daraus, dass bestehende Maßnahmen der Produktbeobachtung und Marktüberwachung auf das Verhalten menschlicher Fahrer ausgerichtet sind und die Komplexität der eingesetzten E/E Systeme und die daraus resultierenden Limitierungen nicht ausreichend berücksichtigen. Nachfolgend werden Modelle vorgestellt, die grundsätzlich geeignet erscheinen, diese Schwachstellen zu beseitigen. Ansätze zur Entstehung von unerwünschten Ereignissen in komplexen sozio-technischen Systemen, z.B. das Barrieren-Modell von Reason [13] oder die Functional Resonance Analysis Method (FRAM) von Hollnagel [14] erlauben es, Verfahren zur Prognose

der Wirksamkeit und der Auswirkungen von automatisierten und autonomen Fahrfunktionen zu entwickeln. Darüber hinaus sollen auch neuere Ansätze aus der Systemtheorie berücksichtigt werden, die sich explizit mit Software- und Design-Fehlern beschäftigen, wie z.B. der System-Theoretic Accident Model and Process Approach (STAMP) von Leveson [15]. Dessen fahrzeugspezifische Umsetzung als System Theoretic Process Analysis (STPA) beschreiben Placke und Mitarbeiter [16].

Reason bildet technische und organisatorische Maßnahmen und menschliches Handeln zur *Vermeidung* von Unfällen in Form von Barrieren ab. Diese Barrieren weisen aktive und latente Fehler auf, die Reason plakativ als Löcher darstellt. Daher führt die Bezeichnung „Schweizer-Käse-Modell“ für seinen Ansatz (Bild 1). Die Löcher können ihre Größe und ihre Lage verändern. Potenziell vorhandene Gefährdungen führen nur dann zu einem Unfall, wenn die Lage der Löcher zufällig bewirkt, dass sämtliche Barrieren nicht greifen.

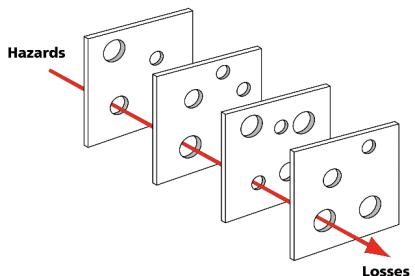


Bild 1: Barrieren-Modell von Reason³

FRAM basiert auf vier Prinzipien:

- Gleichwertigkeit von Misserfolgen und Erfolgen,
- Approximative Anpassungen,
- Emergenz und
- Funktionale Resonanz.

Das erste Prinzip verweist darauf, dass zur Beurteilung komplexer sozio-technischer Systeme nicht nur die seltenen Misserfolge, sondern gleichermaßen die viel häufigeren Erfolge betrachtet werden müssen. Dafür werden die Betrachtungsebenen Sicherheit I und II eingeführt. Si-

³ Davidmack - Eigenes Werk, CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=31679759>

cherheit I beschäftigt sich mit der klassischen Frage, was zum Misserfolg führt. In der Betrachtungsebene Sicherheit II wird nach Ursachen dafür gesucht, dass ein System überwiegend funktioniert, also danach, was zum Erfolg führt. Unter approximativer Anpassung ist zu verstehen, dass sich komplexe sozio-technische Systeme an sich verändernde Umgebungsbedingungen anpassen und dadurch ihre Leistungsfähigkeit verändern. Das Prinzip der Emergenz berücksichtigt, dass die vorgenannten Veränderungen zufällig auftreten und die sich verändernden Systemfunktionen sich auf unvorhersehbare Weise gegenseitig beeinflussen. Das Prinzip der funktionalen Resonanz verweist darauf, dass sich diese Veränderungen unter bestimmten, nicht ohne weiteres vorhersehbaren Umständen - d.h. jenseits einfacher Kausalzusammenhänge - verstärken können. Der grundlegende Schritt bei der Anwendung dieser Methode ist die Identifizierung und Beschreibung von Funktionen und deren Abhängigkeiten. In Bild 2 ist eine FRAM-Funktion dargestellt, die so genannte FRAM-„Schneeflocke“. Jede Funktion wird durch sechs Deskriptoren definiert (Zeit, Steuerung, Ausgang, Ressource, Vorbedingung und Eingabe). Eine Funktion bezieht sich auf die Aktivität - oder auf eine Reihe von Aktivitäten - des Systems, die erforderlich sind, um ein gewünschtes Ergebnis zu erzielen. Über die Deskriptoren werden die verschiedenen Funktionen miteinander verbunden.

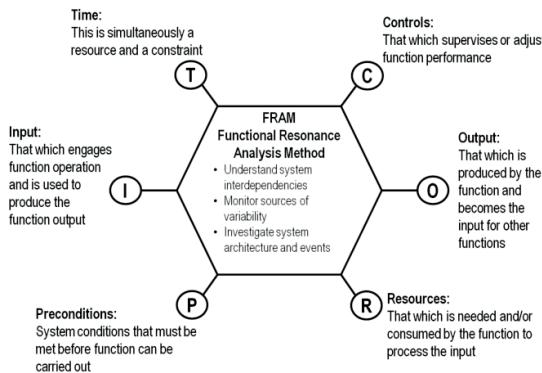


Bild 2: FRAM-Funktion [14]

STAMP ist ein Unfallmodell, das speziell dafür entwickelt wurde, kausale Faktoren wie Software, menschliche Entscheidungsfindung und menschliche Faktoren, neue Technologien, soziale und organisatorische Bedingungen und Sicherheitskultur in die technische Analyse zu integrieren. Zu den Anwendungsbereichen gehören Luft- und Raumfahrt, Verteidigung, Automobilindustrie, Eisenbahn, chemische Industrie, Öl- und Gasindustrie, Medizintechnik, Gesundheitswesen und Arbeitssicherheit. STPA ist eine Technik zur Gefahrenanalyse auf der

Grundlage von STAMP, die speziell im Automobilbereich angewendet wird. STAMP und STPA sind Top-Down-Ansätze, die die zu betrachtenden Systeme als Regelkreise abbilden (Bild 3).

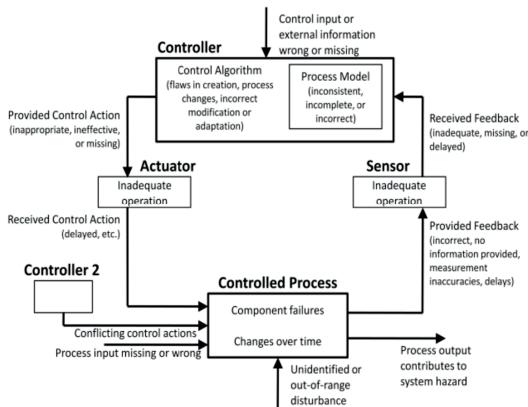


Bild 3: STAMP-Regelkreis [15]

Im ersten Schritt werden - ausgehend von den zu untersuchenden Unfällen - Gefährdungen ermittelt, die zu diesen Unfällen führen können. Anschließend werden unsichere Kontrollaktionen (Unsafe Control Actions, UCA) definiert, die Ursachen für diese Gefährdungen sein können und schlussendlich Szenarien abgeleitet, die zu UCA führen können. Bild 3 enthält beispielhaft Ursachen für unsichere Kontrollaktionen.

6 Betrieblich-organisatorische Lösungsansätze

Der Nachweis der Betriebsbewährung von Systemen zum automatisierten oder autonomen Fahren hängt außer von der zuvor diskutierten Beseitigung der methodischen Schwachstellen insbesondere davon ab, dass es gelingt, auch sehr seltene Ereignisse zuverlässig zu erfassen. Das Problem mit sehr seltenen Ereignissen wird noch dadurch verstärkt, dass am Anfang nur wenig Systeme im Markt sind. Daher muss die Durchführung der Produktbeobachtung und Marktüberwachung optimiert und um gezielte Maßnahmen erweitert werden.

Durch eine bloße Ausweitung, d.h. durch die Analyse von mehr Ereignissen, kann das Problem der sehr seltenen Ereignisse abgemildert, aber voraussichtlich nicht beseitigt werden. Beispielsweise erscheint eine Erfassung *aller* Unfälle in dem für den Nachweis der Betriebsbewährung von Systemen zum automatisierten und autonomen Fahren erforderlichen Detaillierungsgrad selbst in einem begrenzten Erfassungsraum praktisch kaum umsetzbar. Hier könnte

eine gezielte Auswahl Abhilfe schaffen. Dazu müssten Unfälle in automatisierten oder autonomen Fahrmodi als solche erkannt, den betroffenen Einrichtungen gemeldet und dort mit entsprechender Priorität bearbeitet werden. Das erfordert zum einen geeignete technische Maßnahmen zur Erkennung. Dabei können u.U. die in automatisiert oder autonom fahrenden Fahrzeugen vorgesehenen Unfalldatenrekorder helfen. Bisher sind allerdings weder die technischen Randbedingungen zur Erfassung der Unfalldaten noch die Fragen der Cybersicherheit und des Datenschutzes zufriedenstellend geklärt.

Zu Beginn der Einführung, wenn erst wenige Systeme zum automatisierten oder autonomen Fahren im Markt sind, können herstellerübergreifende Feldstudien zu deren Verhalten ebenso helfen, wie die gezielte Befragung von Nutzern der betreffenden Fahrzeuge. Im Rahmen der unmittelbar bevorstehenden Markteinführung von Systemen zum automatisierten Fahren in PKW, läuft das auf Kundenbefragungen hinaus. Beim autonomen Fahren in festgelegten Betriebsbereichen im Rahmen von Mobilitäts- und Transportdienstleistungen sieht das bereits erwähnte nationale Gesetz zum autonomen Fahren eine technische Aufsicht vor. Diese kann verpflichtet werden, relevante Ereignisse zu dokumentieren und zu melden. In jedem Fall bedarf es eindeutiger Indikatoren dafür, dass ein System zum automatisierten oder autonomen Fahren ursächlich war oder zumindest zum Eintritt eines Ereignisses beigetragen bzw. den Verlauf beeinflusst hat.

Darüber hinaus ist zu überlegen, inwieweit z.B. Informationen von Händlerorganisationen der Hersteller oder von Versicherungen als Indikatoren für sich abzeichnende Probleme mit automatisiert oder autonom fahrenden Fahrzeugen genutzt werden können.

7 Hersteller- und technologieunabhängige Begleitung der Markteinführung

In den vorangegangenen Kapiteln wurden Anforderungen an das Inverkehrbringen von Systemen zum automatisierten und autonomen Fahren sowie an die Produktbeobachtung und Marktüberwachung diskutiert. Abschließend sollen das Zusammenwirken der Maßnahmen zum Nachweis der Betriebsbewährung dargestellt sowie die besondere Bedeutung der individuellen und gesellschaftlichen Akzeptanz für die Einführung und Verbreitung von Systemen zum automatisierten und autonomen Fahren beleuchtet werden.

Der Nachweis der hinreichenden Sicherheit der Systeme liegt in der Verantwortung der Hersteller. Besondere Bedeutung wird der Sicherheit der Sollfunktion zukommen, weil wir hier - im Vergleich zur funktionalen und zur Cybersicherheit - normatives Neuland betreten und eine *Best Praxis* noch nicht verfügbar ist. Dessen ungeachtet existieren Lösungsansätze [17]. Die Voraussetzungen für die Erteilung der Straßenzulassung durch das KBA bzw. im Falle des autonomen Fahrens in festgelegten Betriebsbereichen zusätzlich durch die nach Landesrecht

zuständige Stelle für eben jene festgelegten Betriebsbereiche sind damit gegeben. Mit dem Inverkehrbringen setzen Produktbeobachtung und Marktüberwachung ein. Diese liegen grundsätzlich in der Verantwortung der Hersteller bzw. des KBA. Die Ergebnisse von Produktbeobachtung und Marktüberwachung tragen wiederum zur Weiterentwicklung der Systeme durch die Hersteller bei. Dieser Prozess, an dessen Ende der Nachweis der Betriebsbewährung steht, ist in Bild 4 dargestellt.

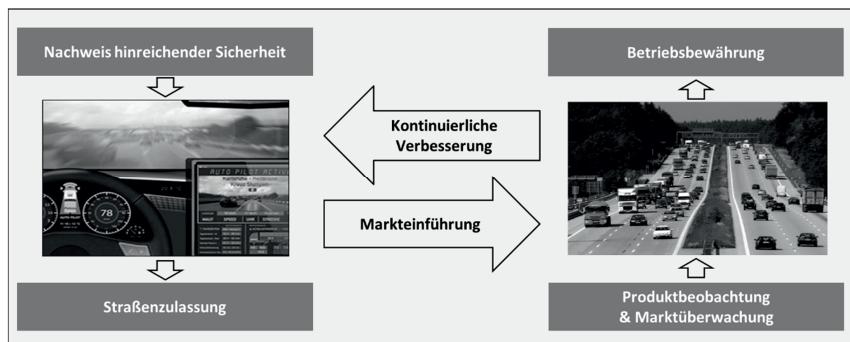


Bild 4: Prozess zum Nachweis der Betriebsbewährung

In diesem Beitrag wurde aufgezeigt, dass die herkömmlichen Maßnahmen zur Produktbeobachtung und Marktüberwachung im Hinblick auf die bevorstehende Einführung von Systemen zum automatisierten und autonomen Fahren systematische Schwachstellen aufweisen. Diese Schwachstellen müssen vor der Markteinführung entsprechender Systeme beseitigt werden.

Wegen der besonderen Bedeutung der individuellen und gesellschaftlichen Akzeptanz für die erfolgreiche Einführung und Verbreitung haben die Autoren dem BMVI die Einsetzung einer Kommission empfohlen, die die Markteinführung vorbereitet und hersteller- und technologieunabhängig begleitet. Die Kommission hat im Wesentlichen zwei Aufgaben.

- Vor der Markteinführung von Systemen zum automatisierten und autonomen Fahren initiiert sie die erforderliche Weiterentwicklung der Maßnahmen zum Nachweis der Betriebsbewährung.
- Nach der Markteinführung unterstützt sie Hersteller und KBA bei der Produktbeobachtung und Marktüberwachung und schlägt Maßnahmen zur Weiterentwicklung der Systeme vor.

Die Kommission besitzt einen beratenden Charakter und nimmt keinen unmittelbaren Einfluss auf Gesetzgebung oder Normung. Mitwirkende sollten neben Vertretern von Bundesministrien und nachgelagerten Behörden vor allem Hersteller und Systementwickler, Sachverständigenorganisationen und Versicherungsunternehmen bzw. deren Verbände sein (VDA, VdTÜV, GDV) sowie Institute, die Verkehrsunfallforschung betreiben (VUFO), Verbraucherschutzorganisationen (ADAC) und wissenschaftliche Einrichtungen. Darüber hinaus erfolgt eine Abstimmung mit internationalen Gremien (z.B. UNECE, WP.29/GRVA).

Ziel der Arbeit der Kommission ist es, hersteller- und technologieunabhängig nachzuweisen, dass die Systeme zum automatisierten und autonomen Fahren die hohen Sicherheitserwartungen, die an sie gestellt werden, tatsächlich erfüllen. Die Autoren sehen darin eine grundlegende Voraussetzung für die Akzeptanz des automatisierten und autonomen Fahrens, die wiederum für eine erfolgreiche Einführung und Verbreitung der Technologie im Markt zwingend erforderlich ist. Das BMVI hat diese Empfehlung im Gesetz zum autonomen Fahren aufgegriffen.

Referenzen

- [1] PEGASUS (2020): Projekt zur Etablierung von generell akzeptierten Gütekriterien, Werkzeuge und Methoden sowie Szenarien und Situationen zur Freigabe hochautomatisierter Fahrfunktion, Abschlussbericht, Version 1.0 https://www.pegasusprojekt.de/files/tmp/pdf/PEGASUS_Abschlussbericht_Gesamtprojekt.PDF
- [2] BMVI (2017): Bericht der Ethik-Kommission Automatisiertes und Vernetztes Fahren https://www.bmvi.de/SharedDocs/DE/Publikationen/DG/bericht-der-ethik-kommision.pdf?__blob=publicationFile
- [3] Steininger, U.; H.-P. Schöner; M. Schiementz (2015): Requirements on tools for assessment and validation of assisted and automated driving systems, In 7. Tagung Fahrerassistenz der TU München, Lehrstuhl für Fahrzeugtechnik, in Zusammenarbeit mit TÜV SÜD Akademie, 25. – 26. Nov. 2015, München <https://www.pegasusprojekt.de/files/tmp/pdf/7.Tagung%20Fahrerassistenz%20Praesentation.pdf>
- [4] U.S. Department of Transportation NHTSA (2017): A vision for safety 2.0 <https://www.nhtsa.gov/document/automated-driving-systems-20-voluntary-guidance>
- [5] Tingvall, C.; N. Haworth (1999): Vision zero. An ethical approach to safety and mobility, In 6th ITE International Conference Road Safety & Traffic Enforcement: Beyond 2000, pp. 6–7 <https://www.monash.edu/muarc/archive/our-publications/papers/visionzero>

- [6] Gasser, T. u.a. (2012): Rechtsfolgen zunehmender Fahrzeugautomatisierung, Berichte der Bundesanstalt für Straßenwesen, Heft F 83, WIRTSCHAFTSVERLAG NW, ISSN: 0943-9307
- [7] Verband der Automobilindustrie e.V. (2015): Von Fahrerassistenzsystemen zum automatisierten Fahren, VDA-Magazin, 2015 <https://www.vda.de/dam/vda/publications/2015/automatisierung.pdf>
- [8] Verein Deutscher Ingenieure e.V. (2018): Berliner Erklärung zur Fahrzeugsicherheit, VDI-Gesellschaft Fahrzeug- und Verkehrstechnik <https://www.vdi.de/ueberuns/presse/publikationen/details/berliner-erklaerung-zur-fahrzeugsicherheit>
- [9] World Forum for Harmonization of Vehicle Regulations - WP.29 (2021): Informal document GRVA-09-07, 9th GRVA, 1 - 5 February 2021, Provisional agenda item 4(b) <https://unece.org/sites/default/files/2021-01/GRVA-09-07e.pdf>
- [10] Liers, H. (2019): Kurzvorstellung des Verkehrsunfallforschungsprojektes GIDAS, In DGUV Fachgespräch Verkehrssicherheit - Detailauswertung und Rekonstruktion von Verkehrsunfällen, 27. - 28. März 2019, Dresden https://www.dguv.de/medien/inhalt/praevention/fachbereiche_dguv/fb-verkehr/veranstaltung2019/vortrag-liers.pdf
- [11] Deutscher Verkehrssicherheitsrat (2012): Vision Zero – Grundlagen & Strategien, Schriftenreihe Verkehrssicherheit Nr. 16, Bonn <https://www.dvr.de/fileadmin/downloads/dvr-schriftenreihe/Schriftenreihe-Verkehrssicherheit-16.pdf>
- [12] Winner, H.; G. Wolf, A. Weitzel (2010): Freigabefälle des Autonomen Fahrens, In 15. Kongress SIMVEC Berechnung und Simulation im Fahrzeugbau, 16. - 17. November 2010, Baden-Baden, VDI-Berichte 2107, VDI-Verlag Düsseldorf, ISBN 9783180921075
- [13] Reason, J. (1991): Human Error, University of Manchester, ISBN: 9780521314190
- [14] Hollnagel, E. (2012): FRAM: The Functional Resonance Analysis Method - Modelling Complex Socio-technical Systems, CRC Press, London, doi:10.1201/9781315255071, eBook ISBN: 9781315255071
- [15] Leveson, N. (2012): Engineering a Safer World: Applying Systems Thinking to Safety, MIT Press
- [16] Placke, S.; J. Thomas, D. Suo (2015): Integration of Multiple Active Safety Systems using STPA. SAE Technical Paper 2015-01-0277, doi: 10.4271/2015-01-0277
- [17] Mazzega, J.; U. Steininger, S. Witkowski (2021): Placing automated driving functions on the market, Auto[nom]Mobil 01. – 02. September 2021, Würzburg



ISBN 978-3-18-092394-9, E-ISBN 978-3-18-102394-5

Bildquelle: © RioPatuca Images – fotolia.com