



Wissenschaftliche Berichte | Scientific reports

Ökologische Transformation in Technik, Wirtschaft und Gesellschaft?

Nr. 002, 2021

Tagungsband zur 26. Interdisziplinären Wissenschaftlichen Konferenz Mittweida





Ökologische Transformation in Technik, Wirtschaft und Gesellschaft? Tagungsband zur 26. Interdisziplinären Wissenschaftlichen Konferenz Mittweida

Impressum

Herausgeber:

Hochschule Mittweida
University of Applied Sciences
Der Rektor
Prof. Dr. phil. Ludwig Hilmer
Der Prorektor Forschung
Prof. Dr.-Ing. Uwe Mahn

Redaktion dieser Ausgabe:

Hochschule Mittweida | Referat Forschung
University of Applied Sciences

Tagungsgruppenleiter

Prof. Dr.-Ing. Leif Goldhahn
Prof. Dr.-Ing. Jörg Hübler
Prof. Dr.-Ing. René Ufer
Prof. Dr.-Ing. habil. Alexander Winkler
Prof. Dr. rer. nat. habil. Thomas Villmann
Prof. Dr.-Ing. Christian Schulz
Prof. Dr.-Ing. Jan Thomanek
Prof. Dr. rer. nat. habil. Röbbbe Wünschiers
Prof. Dr.-Ing. Anika Dittmar
Prof. Dr.-Ing. Jörg Mehlis
Prof. Dr.-Ing. Jan Schaaf
Prof. Dr. phil. Stephan Beetz
Prof. Dr. med. habil. Dr. rer. soc. Stefan
Brunnhuber

Kontakt:

Hochschule Mittweida
University of Applied Sciences
Referat Forschung
Postfach 1457
D-09644 Mittweida

Tel.: +49 (0) 3727 / 58-1264
Fax: +49 (0) 3727 / 58-21264
forschung@hs-mittweida.de
www.forschung.hs-mittweida.de

ISSN 1437-7624

Erscheinungsweise:

Unregelmäßig

Auflage:

Belegexemplare sowie bestellte Druckexemplare

Druck:

Hochschuldruckerei Hochschule Mittweida

Titelseite: Foto: greenbutterfly-stock.adobe.com |
Romolo Tavani-stock.adobe.com

Bildnachweise werden direkt am Foto bzw. im
jeweiligen Artikel aufgeführt.

Im Scientific Report gelten grammatikalisch
maskuline Personenbezeichnungen gleicher-
maßen für Personen jeglichen Geschlechts.

Die Scientific Reports/Wissenschaftliche Berichte
als Wissenschaftliche Zeitschrift der Hochschule
Mittweida - University of Applied Sciences lösen
die bisherigen Scientific Reports mit allen Volume
I-III ab und erscheinen mit Nr.1, 1998 ab
November 1998 in neuem Layout und in neuer
Zählung.

Für den Inhalt der Beiträge sind die Autoren
verantwortlich.

Im laufenden Kalenderjahr sind bereits
erschienen:

Nr. 001, 2021

Entwicklung innovativer Verfahrens- und
Betriebsmittelmodelle sowie Qualifizierungs-
konzepte für die ressourceneffiziente Fertigung
hochbeanspruchter Bauteile -MoQuART

SCIENTIFIC REPORTS | WISSENSCHAFTLICHE BERICHTE

The main aspect of the Scientific Reports is to promote
the discussion of modern developments in research and
production and to stimulate the interdisciplinary
cooperation by information about conferences,
workshops, promotion of partnerships and statistical
information on annual work of the Hochschule
Mittweida (FH) University of Applied Sciences. This issue
will be published sporadically. Contributors are
requested to present results of current research,
transfer activities in the field of technology and applied
modern techniques to support the discussion among
engineers, mathematicians, experts in material science
and technology, business and economy and social work.

Die Scientific Reports der Hochschule Mittweida sind online verfügbar unter:

www.forschung.hs-mittweida.de/veroeffentlichungen/scientific-reports

Eine Veröffentlichung einzelner Beiträge erfolgt entsprechend der Open Access Strategie der Hochschule
Mittweida auf dem Hochschulschriftenserver: <https://monami.hs-mittweida.de>

Herzlich willkommen zur 26. IWKM

Der wachsenden Weltbevölkerung und dem steigenden Lebensstandard auf der einen Seite stehen knapper werdende Ressourcen und Nutzflächen, Klimawandel und Rückgang der Artenvielfalt auf der anderen Seite entgegen.

Wie begegnen wir diesen großen globalen Herausforderungen? Wie können wir nachhaltiger leben, Ressourcen schonen und gleichzeitig unseren Lebensstandard sichern? Welchen Nutzen können wir aus Wissenschaft und Forschung ziehen? Welche Möglichkeiten bietet der digitale Wandel für die Entwicklung nachhaltiger Lösungen? Und wie profitieren davon Wirtschaft und Gesellschaft?

Auf der 26. Interdisziplinären Wissenschaftlichen Konferenz an der Hochschule Mittweida wollen wir gemeinsam mit Vertreterinnen und Vertretern aus Wissenschaft, Wirtschaft und Gesellschaft Technologien und Verfahrensweisen für eine nachhaltige und biobasierte Wirtschaftsform aufzeigen, Denkansätze diskutieren und erste Schritte gehen, hin zu einem nachhaltigen und kreislauforientierten Wirtschaften. Dabei beleuchten wir das Thema aus ökologischer, technischer, ökonomischer und sozialer Sicht.

Im Jahr 2021 feiern wir ein außergewöhnliches Jubiläum unserer wissenschaftlichen Konferenzen. Die allererste wissenschaftliche Konferenz fand vor 50 Jahren am 09. Juni 1971 an der damaligen Ingenieurhochschule Mittweida statt. Im Programm, welches in unserem Hochschularchiv verfügbar ist; steht:

„Weitere Veranstaltungen dieser Art werden in einem festen Rhythmus folgen. Die erste Tagung soll vor allem der Selbstverständigung des Lehrkörpers und des gesamten wissenschaftlichen Personals unserer Hochschule dienen. Rektor und Wissenschaftlicher Rat sind bestrebt, diese Veranstaltungsreihe zu einem solchen Erfolg zu führen, daß sie internationale Anerkennung findet.“

Die 26. IWKM setzt die 1971 begonnene Konferenzreihe in bewährter und stetiger Weise fort und baut, wie schon vor 50 Jahren, auf den Austausch zwischen Wissenschaft, Wirtschaft und Gesellschaft.

Aufgrund der Corona-Pandemie findet die 26. IWKM als Online-Veranstaltung statt und trägt dadurch selbst ein wenig zur Nachhaltigkeit bei.

Seien Sie herzlich willkommen auf unserer digitalen Tagung, wählen Sie aus dem vielfältigen Programm der Tagungsgruppen und treten Sie in einen fruchtbaren Austausch mit den Wissenschaftlerinnen und Wissenschaftlern. In den Eröffnungsveranstaltungen an beiden Tagen freuen wir uns auf inspirierende Keynotes und Grußworte sowie auf die Vergabe von drei Preisen, mit denen wissenschaftliche Arbeiten gewürdigt werden. Im Rahmenprogramm in den Konferenzpausen können Sie an einem virtuellen Campusrundgang teilnehmen, Labore besichtigen oder sich über die Transferaktivitäten an der Hochschule informieren.

Ich wünsche Ihnen eine erfolgreiche Konferenz!

Prof. Dr.-Ing. Uwe Mahn
Prorektor Forschung an der Hochschule Mittweida
Tagungsleiter

Inhaltsverzeichnis

Nachhaltige Produkt- und Prozessentwicklung im disruptiven Umfeld

Nachhaltige Produktentwicklung / Digitaler Zwilling

Entwicklung von Material-Mastermodellen für die Nutzung in CAE- und PLM-Systemen011

René Ufer¹, Thies Marwitz¹, Uwe Diekmann²

¹ Hochschule Mittweida, ² MATPLUS GmbH,

Nachhaltige Produktentwicklung durch Crowdsourcing.....015

Sarah Uhlig, Roman Nuss, André Schneider

Hochschule Mittweida

Ressourcen- und umwelteffiziente Gestaltung von Fördersystemen mit Kunststoffketten019

Jens Sumpf, Lynn Lüdemann, Marcus Bona, Markus Golder

TU Chemnitz, Professur Förder- und Materialflusstechnik

Prototyping / Predictive Maintenance

Impedimetrische Sensoren zur Überwachung von Deichbauwerken 025

Wolfgang Fichtner¹, Michael Mertig^{1,2}, Dirk Fleischer³, Ulf Helbig³

¹ Kurt-Schwabe-Institut für Mess- und Sensortechnik Meinsberg e.V.

² Institut für Physikalische Chemie, Technische Universität Dresden

³ Institut für Wasserbau und Technische Hydromechanik, Technische Universität Dresden

Kunststoffgleitlager mit intelligenter Verschleißüberwachung.....029

Norman Katzer, Hagen Bankwitz, Jörg Hübler

Hochschule Mittweida

Wissenschaft und Forschung zur additiven Fertigung an der Hochschule Mittweida033

Hagen Bankwitz, Jörg Hübler, Claudius Petzold

Hochschule Mittweida

Mikromobilität

EMOTION – Electric Mobility in L-Category Vehicles for all Generations 038

Nicole Perterer, Martin Perterer

KTM E-Technologies

Recovery of electronic waste (lithium battery) for the industrial competitiveness..... 043 of electronic recycling

Dra. María Luisa Valenzuela¹, Dr. Rene Rojas², Dra. Marjorie Segovia, Sebastián Quintana,

¹ Universidad Autónoma, Instituto de Ciencias Químicas Aplicadas, Facultad de Ingeniería.

² Universidad Católica, Facultad de Química y de Farmacia.

Robotik und Automation 1

Untersuchung von Kraftverlusten in der Antriebskinematik von Umformmaschinen.....048

Armin Schleinitz, Henry Kirchner, André Sewohl, Holger Schlegel, Martin Dix

Technische Universität Chemnitz, Institut für Werkzeugmaschinen und Produktionsprozesse (IWP)

Adaptive Produktion von Holzbauteilen mittels Part-to-Tool052

Marten Stepputat, Florian Beuß, Uwe Pfletscher, Jan Sender, Wilko Flügge

Fraunhofer-Institut für Großstrukturen in der Produktionstechnik IGP

| | |
|--|------------|
| Möglichkeiten der Kraftregelung an Leichtbaurobotern..... | 057 |
| Christian Thormann, Alexander Winkler Hochschule Mittweida | |
| Robotik und Automation 2 | |
| Nutzung und Potenziale von Instandhaltungsmanagementsystemen für die Optimierung..... | 061 |
| automatisierter Produktionsprozesse Christian Brenner, Martin Hiersemann, Rolf Hiersemann Hiersemann Prozessautomation GmbH | |
| Einsatz von AutomationML zur automatischen Generierung von Digitalen Zwillingen | 065 |
| von Werkzeugmaschinen zur Entwicklung energieeffizienter Produktionssysteme Nicolai Beisheim, Tobias Ott, Sebastian Amann Hochschule Albstadt-Sigmaringen | |
| Sprachsteuerung für einen industriellen Leichtbauroboter | 072 |
| Thomas Wachsmuth, Christian Thormann, Alexander Winkler Hochschule Mittweida | |
| Digitalisierung in der Montage 1 - Abschlussworkshop Projekt FrEDA - Framework Enriched Data Assembly | |
| Einsatzszenarien für Electronic Shelf Label in der Produktion | 076 |
| Leif Goldhahn, Sebastian Roch Hochschule Mittweida | |
| Einsatz des webbasierten Frameworks der Montage (FrEDA) in Verbindung mit..... | 080 |
| digitalen Beschriftungssystemen Leif Goldhahn, Dorit Bock Hochschule Mittweida | |
| Digitalisierung in der Montage 2 | |
| Entwicklungsbegleitende Planung neuer Montageprozesse in der Zulieferindustrie | 084 |
| Michael Willnich Pierburg Pump Technology | |
| Standard Assembly Time Setting in an Early Stage of Product Development..... | 088 |
| Izabela Kutschenreiter-Praszkiewicz University of Bielsko-Biala, Poland | |
| Gestaltung von Anleitungen für die Montage variantenreicher Erzeugnisse..... | 092 |
| Sven Hinrichsen Technische Hochschule Ostwestfalen-Lippe | |
| Gestaltung industrieller Arbeits- und Lernprozesse 1 | |
| Ergonomie 4.0 – Herausforderungen und Chancen der Industrie 4.0 für die | 097 |
| Gestaltung menschengerechter Arbeitsplätze Lars Fritzsche imk automotive GmbH | |

| | |
|---|-----|
| Leitidee Ökologie als Nachhaltigkeitsdimension in Lernprozessen für zukünftige | 100 |
| Fachkräfte in der Kunststofftechnik – Beispielhafte Lern- und Arbeitsaufgabe zur Wahl von Fertigungsparametern mit einem Augmented Reality Lernsystem | |
| Marvin Goppold ¹ , Simon Kreisler ¹ , Maren Klatt ¹ , Lukas Dunsche ² , Tobias Döhring ² , Sven Tackenberg ³ , Martin Frenz ¹ , Verena Nitsch ¹ | |
| ¹ Institut für Arbeitswissenschaft der RWTH Aachen | |
| ² Institut für Kunststoffwirtschaft | |
| ³ Labor für Industrial Engineering der Technischen Hochschule Ostwestfalen-Lippe | |
| Standardisiertes Übungsszenario für HMD-basierte virtuelle Lernumgebungen | 105 |
| Leif Goldhahn, Robert Eckardt, Katharina Müller-Eppendorfer Hochschule Mittweida | |
| Gestaltung industrieller Arbeits- und Lernprozesse 2 | |
| Auswahl und Einsatz von Smart Devices in der Arbeitsanalytik | 109 |
| Torsten Merkel Westfälische Hochschule Zwickau | |
| Konzept zur Simulation des Entscheidungsverhaltens von Arbeitspersonen | 113 |
| bei der Produktion kundenindividueller Produkte | |
| Sven Tackenberg ¹ , Jan-Phillip Herrmann ¹ , Marvin Goppold ² | |
| ¹ Labor für Industrial Engineering der Technischen Hochschule Ostwestfalen-Lippe | |
| ² Institut für Arbeitswissenschaft der RWTH Aachen University | |
| Dynamische Losbildung für eine dezentral geplante und gesteuerte Produktion | 118 |
| Martin Krockert, Marvin Matthes und Torsten Munkelt HTW Dresden | |
| Entwicklung eines Systems zur Planung und Fertigung patientenspezifischer | 122 |
| Schädelimplantate unter Verwendung additiver Fertigungsverfahren und Techniken des maschinellen Lernens | |
| Michael Kaiser, Patrick Knorr Westfälische Hochschule Zwickau | |
| KI und Maschinelles Lernen | |
| Sorting of Single-Molecule Trajectories by means of Machine Learning - a status update on the annotation procedure | 126 |
| Lisa Krenkel ¹ , Tobias Schlosser ² , Danny Kowanko ² , Richard Börner ¹ | |
| ¹ Laserinstitut Hochschule Mittweida | |
| ² Department of Informatics, Technical University Chemnitz | |
| Analyse von Thorax-Aufnahmen zur Erkennung von Virusinfektionen | 130 |
| Lena Kopp, Roman Radtke, Alexander Jesser Hochschule Heilbronn | |
| Intelligent Gait Analysis using Marker Based Motion Capturing System | 133 |
| Danny Möbius ^a , Marika Kaden ^b , Daniel Staps ^b , and Thomas Villmann ^b | |
| ^a Institut für Mechatronik, Chemnitz, ^b SICIM, Hochschule Mittweida | |
| Quantum Computing for Efficient Learning in Prototype-based Vector Quantization | 137 |
| Thomas Villmann, Alexander Engelsberger Univ. of App. Sciences Mittweida, Saxon Institute for Computational Intelligence and Machine Learning, | |
| An attempt to explain double descent in modern machine learning | 141 |
| Jochen Merker, Gregor Schuldt HTWK Leipzig | |

| | |
|--|-----|
| Sensors Data Fusion for Smart Decisions Making Using Interpretative Machine Learning Models | 145 |
| Feryel Zoghلامي ^a , Marika Kaden ^b , Thomas Villmann ^b , Germar Schneider ^a , Harald Heinrich ^a | |
| ^a Automation, Maintenance and Factory Integration Infineon Technologies Dresden GmbH & Co. KG | |
| ^b SICIM, Hochschule Mittweida | |
| Recurrent Learning Vector Quantization | 147 |
| Jensun Ravihandran and Thomas Villmann | |
| SICIM, Hochschule Mittweida | |
| Machine Learning Algorithms for Classifying Component Defects for Predictive Maintenance | 149 |
| Christoph Kammerer, Micha Küstner, Michael Gaust, Pascal Starke, Roman Radtke, Alexander Jesser | |
| University of Applied Sciences Heilbronn, CeraCon GmbH | |
| Edge detection in Images with multiplicative Noise by using Ant Colony System | 154 |
| Sergio Baltierra ¹ , Marco Mora ² | |
| ¹ Facultad de Ingeniería, Universidad Autónoma de Chile | |
| ² Department of Computer Science, Universidad Católica del Maule | |
| The challenge of using artificial intelligence in autonomous driving | 157 |
| Michael Sachse, André Müller, Klaus Krumbiegel | |
| Hochschule Mittweida | |
| Transformation der Mobilität - ökologisch, autonom und vernetzt | |
| Autonomes Fahren – Traum der Ingenieure, Alptraum der Juristen? | 162 |
| Daniel Graewe | |
| NORDAKADEMIE Hochschule der Wirtschaft | |
| Betrieb Sächsischer Testfelder für das automatisierte und vernetzte Fahren | 166 |
| Christian Andrä | |
| Sächsische Energieagentur – SAENA GmbH | |
| Smarte Mobilitätsketten als Zukunftsbild für den ländlichen Raum | 170 |
| Marco Rehme, Udo Wehner, Julia Döring, Steve Rother | |
| Technische Universität Chemnitz | |
| Mobility 4ALL – Mobilität weiterdenken in Mittweida | 175 |
| Ralf Hartig, Alexander Knauer, Vivien Lorenz, René Härtel | |
| Hochschule Mittweida | |
| Umstieg des ÖPNV auf alternative Antriebstechnologien im Regionalverkehr | 179 |
| Andreas Seltmann | |
| REGIOBUS Mittelsachsen GmbH | |
| Application of Green Hydrogen in Mobility Sector | 183 |
| Lorenzo Reyes-Bozo ¹ , Carlos Fúnez Guerra ² , Claudia Sandoval Yañez ¹ , Guillermo Schaffeld ¹ | |
| ¹ Universidad Autónoma de Chile | |
| ² National Hydrogen Center, Puertollano, Spain | |
| Autonomes und Vernetztes Fahren auf der Schiene - ATLAS-Modellprojekt der HS Mittweida | 187 |
| Christian Schulz, Jan Thomanek | |
| Hochschule Mittweida | |

| | |
|---|-----|
| Implementing and Evaluating a Tracking-By-Detection Algorithm for a Camera | 193 |
| Monitoring System | |
| Khaled Jbaili, Jan Thomanek Hochschule Mittweida | |
| Suffizienzstrategien für Immobilien - darf es etwas weniger sein? | |
| Wie muss man bauen, um suffizientes Wohnen zu ermöglichen? | 204 |
| Margarete Over (ifeu), Patrick Zimmermann (BTU), Lars-Arvid Brischke (ifeu) ifeu – Institut für Energie- und Umweltforschung Heidelberg gGmbH, BTU Cottbus-Senftenberg | |
| Wohnraum: Wird die Qualität in Quadratmetern gemessen? | 210 |
| Lena-Katharina Peter, Anja Bierwirth Wuppertal Institut für Klima, Umwelt, Energie gGmbH | |
| Wohnflächeneffizienz möchte erhöht werden | 215 |
| Arne Steffen werk.um architekten | |
| Ökologischer Wandel und soziale Teilhabe | |
| Ökologischer Wandel und Transformation: | 221 |
| Neue Herausforderungen für die Hochschulkommunikation | |
| Markus Drowatzky Hochschule Mittweida | |
| StakeholderInnen- und Kommunikationsmanagement sächsischer Energie- und | 225 |
| Rohstoffunternehmen | |
| Stefanie Walter Hochschule Mittweida | |
| Biologische Systeme verstehen | |
| FRET-assisted structural model of the GAAA RNA tetraloop receptor | 230 |
| Felix Erichson ¹ , Fabio D. Steffen ² , Richard Börner ¹ ¹ Laserinstitut Hochschule Mittweida ² Department of Chemistry, University of Zurich | |
| Optimization of an optical setup for combined TIRF/confocal fluorescence microscopy | 234 |
| Hannes Michaelsen, Richard Börner Laserinstitut Hochschule Mittweida | |
| AnthroWorks3D: Digitalisierung von Skelettfundstücken und die virtuell-osteologische | 237 |
| Untersuchung | |
| Hanjo Tim Fritzschn ¹ , Fabian Schmiedel ¹ , Marie Heuschkel ¹ , Marleen Mohaupt ¹ , Dirk Labudde ¹ , Toralf Kirsten ¹ , Birgit Grosskopf ² , Anna Lena Flux ² , Stefanie Meier ³ , Patricia van der Burgt ³ ¹ Hochschule Mittweida ² Georg-August-Universität Göttingen ³ Landesamt für Archäologie Sachsen | |
| Engineering cyanobacteria for direct solar chemical and fuel production from CO2 | 241 |
| Peter Lindblad Uppsala University, Department of Chemistry-Ångström, Sweden | |

| | |
|--|-----|
| Development of a genetic biomonitoring test for the investigation of pollinator-plant-interactions | 244 |
| Lisa Prudnikow ¹ , Robert Leidenfrost ¹ , Catrin Westphal ² , Röbbbe Wünschiers ¹ | |
| ¹ Hochschule Mittweida | |
| ² Georg-August-Universität Göttingen | |
| Which pollen do bees like best? - Analysis of mixed pollen samples with metabarcoding | 248 |
| Svenja Bänsch ^{1,2} , Teja Tschardt ³ , Röbbbe Wünschiers ⁴ , Catrin Westphal ¹ | |
| ¹ Functional Agrobiodiversity, University of Göttingen, Germany | |
| ² Centre for Biodiversity Monitoring, Zoological Research Museum Alexander Koenig Bonn, Germany | |
| ³ Agroecology, University of Göttingen, Germany | |
| ⁴ Department of Biotechnology and Chemistry, Mittweida University of Applied Sciences, Germany | |
| Posterbeiträge | |
| Long reads for microbial bioaerosol surveillance | 253 |
| Robert Maximilian Leidenfrost ¹ , Dierk-Christoph Pöther ² , Henriette Mietke ³ , Udo Jäckel ² , Röbbbe Wünschiers ¹ | |
| ¹ Mittweida University of Applied Sciences | |
| ² Federal Institute of Occupational Safety and Health | |
| ³ State Operating Company for Environment and Agriculture | |
| Learning by doing: Integrating bioprocess engineering and undergraduate education | 255 |
| Robert Maximilian Leidenfrost, René Kretschmer Mittweida University of Applied Sciences | |
| Into the genome of the German Sheep Poodle | 256 |
| Robert Maximilian Leidenfrost, Nils Schön, Rebecca Prause, Röbbbe Wünschiers Mittweida University of Applied Sciences | |

Nachhaltige Produkt- und Prozessentwicklung im disruptiven Umfeld

Entwicklung von Material-Mastermodellen für die Nutzung in CAE- und PLM-Systemen

René Ufer¹, Thies Marwitz¹, Uwe Diekmann²
¹ Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida
² MATPLUS GmbH, Hofaue 55, 42103 Wuppertal

Im Produktlebenszyklusmanagement (PLM) spielt die Möglichkeit der breiten Abbildung von Daten über diverse Systeme, welche ähnliche softwaretechnische Funktionen und Inhalte abdecken, eine große Rolle. Zwischen den Systemen für die Materialmodellierung und -visualisierung, die einen Beitrag zur Erstellung digitaler Zwillinge leisten, ergeben sich große Differenzen in der Ausführung der Modelle bei ähnlichen Kategorien (Plastizität, Schäden, Temperatur etc.). Angefangen mit der unterschiedlichen Interpretation von allgemeinen Materialmodell-Ansätzen bis hin zur Bezeichnung und Definition von notwendigen Parametern, gilt es, eine Übertragbarkeit und Kompatibilität der spezifischen Modelle systemübergreifend zu ermöglichen. Dieser Beitrag verdeutlicht eine Methodik zur Erstellung und Entwicklung von Material-Mastermodellen, welche die benötigten physikalischen Eigenschaften, Korrekturparameter und Konstanten zur Abbildung diverser Materialmodelle unterschiedlicher Systeme abbilden. Auf Basis herkömmlicher (bspw. Zugversuch) und fortschrittlicher (bspw. DIC) Werkstoffprüfungen, sowie Werkstoffdatenbanken wird dafür ein Datengerüst erstellt. Der konzipierte digitale Material-Zwilling ist in Hinblick auf unterschiedliche Anwendungen im PLM nutzbar.

1. Einleitung

Um Materialien und ihre Eigenschaften digital und als Digitalen Zwilling nutzen zu können, benötigt man Materialmodelle, welche physikalische Änderungen durch unterschiedliche Einwirkungen abbilden können. Im Rahmen des Projektes „AMMICAL – Advanced Materials Model Integration for CAE-Applications“ werden Werkstoffdaten in die Industrie 4.0 Wertschöpfungsketten integriert und ein generisches Materialmodell für die Prozess- und Systemsimulation abgeleitet. Dafür wird ein Prototyp geschaffen, welcher die Datenerfassung, -verdichtung, -auswertung, Modellbildung und -distribution entlang der gesamten Prozesskette von fortschrittlichen Werkstoffprüfungen unterstützt. [1][2]

Die Materialmodellierung bedingt aufgrund ihrer Vielzahl an möglichen Modellarten und Anwendungsgebieten sowie deren Ausprägungen einen großen Datensatz

an Ausgangswerten und -parametern. In Anbetracht diverser Software zur Erstellung von Materialmodellen und deren unterschiedlichen Auslegungen von Format, Anordnung der Materialkarten und Anwendung der Basisformel zur Berechnung des Modells werden, um Parameter zu mappen und einen Zusammenhang zwischen den Modellen zu erstellen, sogenannte Mastermodelle benötigt.

Folgend wird anhand der Abbildung 1 das Umfeld der Mastermodelle für die Kategorien Plastizität, Schaden und Temperatur beschrieben.

2. Umfeld des Mastermodells

Das Mastermodell bildet die Verknüpfung von der Datenerfassung und -verarbeitung auf Basis von bspw. Werkstoffprüfungen oder Werkstoffsimulationswerk-

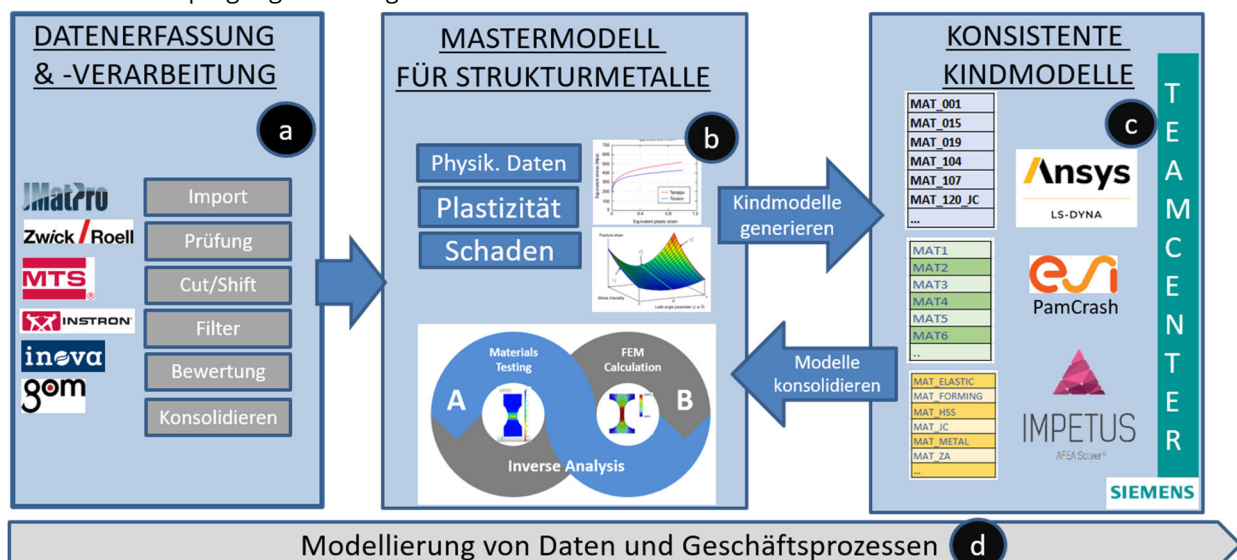


Abbildung 1: Prozessmodellierung Mastermodell/Kindmodell (Erstellt von der Matplus GmbH)

zeugen **(a)** und der möglichen Generierung verschiedener Materialmodelle von FEM-Simulationsprogrammen (Anwendungssoftware), wie bspw. *Ansys LSDyna*, *PamCrash* und *Impetus Afea*, auf Basis des gleichen Datensatzes **(c)**. Das Mastermodell basiert auf verschiedenen Modellkategorien (Physikalische Daten, Plastizität und Schaden). Aufgrund einer Inversen Analyse können die jeweiligen Variationen der zerstörenden und zerstörungsfreien Prüfverfahren genauer nachvollzogen und aufeinander abgestimmt werden, was ökonomische und ökologische Folgewirkungen hat **(b)**.

2.1 Datenerfassung und -verarbeitung

Die Datenerfassung und -verarbeitung bildet den Anfang des Prozesses der „Modellierung von Daten und Geschäftsprozessen“ (Abbildung 1). Der Datensatz kann neu erstellt, importiert oder verändert werden. Darüber hinaus werden temperaturabhängige Werkstoffeigenschaften, wie thermische Ausdehnung und Wärmekapazitäten, mithilfe des Werkzeuges „MatPro“ [3] ermittelt und dem Mastermodell hinzugefügt. Chemische Zusammensetzungen der Materialien werden über Datenbanken dem Datensatz zugeordnet. Des Weiteren werden die Output-Dateien der Werkstoffprüfmaschinen per Kurvenanpassung komprimiert, in dem die Messdaten geglättet, vermindert und angeglichen werden. Damit wird für die spätere Verwendung eine schnellere Datenverarbeitung bei komplexen und umfangreichen Materialmodellen ermöglicht.

2.2 Erstellung bzw. Ableitung der Kindmodelle mit dem AMMICAL-Prototypen

Die Kindmodelle und deren Verwendung in CAE-Applikationen und PLM-Systemen bilden im abgebildeten Prozess, abgesehen von einer möglichen Weiterentwicklung und Revisionierung der Modelle, das Prozesskettenende.

Grundsätzlich hat ein Materialmodell (Kindmodell) mehrere Materialkarten, welche verschiedene Parameter und Informationen enthalten. Zwischen den Softwares zur Erstellung von Materialmodellen gibt es Unterschiede in der Ausführung der Materialkarten, so sind es beispielsweise bei *LSDyna* acht Parameter pro Materialkarte, wohingegen es bei *PamCrash* mehr als acht sein können.

Der Ablauf zum Erstellen eines Kindmodells beginnt mit den Werkstoffprüfdaten, die zusammengefasst und validiert werden. Die reale Werkstoffprüfung ist notwendig, um Eigenschaftskennnisse der Werkstoffe zu gewinnen. Die Auswahl des Materialmodelles bzw. Kindmodelles richtet sich nach dem Zeitpunkt der zu untersuchenden Belastung. Ist die Versuchsdurchführung stoß- oder ruckartig, so kommt ein Modell in Betracht, das plastische bzw. elastisch-plastische Eigenschaften widerspiegelt. Bei monoton und langsam ansteigenden Belastungen der Werkstoffprüfung hingegen werden hauptsächlich elastische Eigenschaften in den Vordergrund der

Modellierung gestellt [4]. Die Anwendungsbereiche können sich jedoch überschneiden, bspw. wenn der elastische Bereich bei einer stoß- bzw. ruckartigen Belastung untersucht werden soll.

Aufgrund der Vielzahl von Software zur Erstellung und Bearbeitung von Materialmodellen, haben die notwendigen Parameter diverse Bezeichnungen und Positionen. Daher ist es notwendig, die Parameter auf Basis der Variablen und Definition zwischen den Anwendungen zu mappen. In Tabelle 1 ist dieser Sachverhalt exemplarisch dargestellt. Am Beispiel des „Young´s Modulus“ wurden die Parameterbezeichnungen zusammengefasst und erkennbar gemacht, dass diese sich in den verschiedenen FEM-Simulationsprogrammen überschneiden.

Tabelle 1: Parameterbezeichnungen der FEM-Simulationsprogrammen

| Parameter | FEM-Simulationsprogramm | | |
|-----------------|---|--|--|
| | LSDyna | PamCrash | ImpetusAfea |
| Young´s Modulus | <ul style="list-style-type: none"> • EL • YMRT • YM • EA – EC • E • E1-E3 | <ul style="list-style-type: none"> • Ec • Eswi • Et • E • E1-E3 | <ul style="list-style-type: none"> • E • E1-E3 |

Im AMMICAL-Prototypen werden die Parameter mit ihren möglichen Bezeichnungen gemappt und gespeichert, sodass diese per Programmierung angesprochen werden können.

2.3. Mastermodell und dessen Parameter sowie Einheitensystem

Das Mastermodell ermöglicht die Ableitung aller integrierten Kindmodelle verschiedener Anwendungssoftwares [5] auf Basis eines gleichen Datensatzes. Somit setzen sich die Mastermodell-Parameter aus den Kindmodell-Parametern zusammen und werden mit ihren unterschiedlichen Bedeutungen für eventuell gleiche Bezeichnungen gemappt. Es gibt zwei Herangehensweisen für die Erstellung des Mastermodells: Zum einen auf Basis der Werkstoffprüfdaten und zum anderen durch Auflösung mehrerer bzw. verschiedener Materialmodelle für gleiche Materialien und Umstände. Die abgeleiteten Kindmodelle können weiterbearbeitet bzw. optimiert oder für andere Umstände angepasst werden. In dem Fall werden diese dann versioniert oder Varianten je nach Anwendungsfall gebildet. Das Mastermodell kann aktualisiert und ebenfalls versioniert werden. Der Ablauf kann in Abbildung 2 nachvollzogen werden.

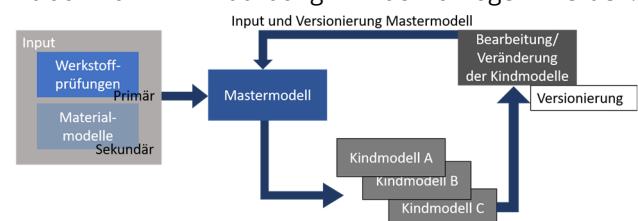


Abbildung 2: Verarbeitungsprozess zur Erstellung der Mastermodelle

Die Materialmodelle bzw. die verschiedenen Anwendungssoftwares geben nicht die Einheiten der Modellparameter an. Anhand von Werten/Angaben für bspw. Dichte und der Art des geprüften Materials lassen sich die Einheiten von vorhandenen Materialmodellen ermitteln. Stellt man ein Mastermodell anhand geprüfter Werkstoffinformationen auf, so ist die Angabe eines konsistenten Einheitensystems notwendig und ermöglicht in weiterer Bearbeitung leichte Umrechnungen. Wird jedoch das Mastermodell durch die Auflösung der Materialmodelle aufgestellt, so ist eine Einheitszuordnung meist nicht gegeben. Durch Betrachtung der Größenordnung der Werte für die Dichte und den „Young's Modulus“ lassen sich für eine Vielzahl an metallischer und nichtmetallischer Werkstoffe die genutzten Einheiten konsistent zuordnen. Die Dichte (Masseneinheit*Längeneinheit³) umfasst die Masse- und die Längeneinheit, wobei der „Youngs Modulus“ (Masseneinheit/(Längeneinheit*Zeiteinheit²)) noch die Zeiteinheit enthält. In Kombination kann ein konsistentes Einheitensystem (Masse – Länge – Zeit) den Modellen zugeordnet werden [6]. Analog dazu können die Einheiten für die Kindmodelle zugeordnet werden.

2.4 Import/Export Kindmodelle mit dem AMMICAL-Framework

Mit dem AMMICAL-Prototyp lassen sich Material- bzw. Kindmodelle von den Anwendungssoftwares importieren, bearbeiten, versionieren und wieder exportieren. Der Import wird durch das Einlesen der Ausgangsdatei und die Umwandlung in eine JSON-Datei ermöglicht. Bei dem Export wird die JSON-Datei ausgewertet, die Werte den jeweiligen Materialmodellparametern zugeordnet und per *Django* Template in ein HTML-Output mit Download-Link im Zielformat gewandelt.

3. Materialmodelle als wichtiger Bestandteil für den Digitalen Zwilling

Materialmodelle leisten einen erheblichen Beitrag zum digitalen Zwilling, welcher alle relevanten Informationen zu Prozessen und Objekten enthält [7]. Im Grunde sind Materialmodelle ein Teil des digitalen Zwillings, da sie Dinge, in dem Fall Eigenschaften von Materialien, aus der realen Welt repräsentieren. Bei der Materialmodellierung werden die Ergebnisse bspw. einer durchgeführten oder noch durchzuführenden Werkstoffprüfung simuliert.

Der Zusammenhang zwischen Materialmodell und einem umfassenden digitalen Zwilling liegt im Lebenszyklusverhalten von Produkten, welches aufgrund verschiedener Materialeigenschaften beeinflusst wird. Über die Materialmodelle können bspw. Druck- und Zugbeanspruchungen simuliert und der mögliche Einfluss auf Form oder Funktionsfähigkeit untersucht werden.

4. Materialmodelle im PLM-System

Mithilfe von PLM-Systemen ist eine strukturierte Übersicht von Materialmodellen möglich. Es lassen sich Variantenräume, für bspw. ähnliche Materialmodelle, wie die Johnson-Cook-Modelle der Anwendungssoftwares, erstellen bzw. zuordnen. Durch die fortlaufende Optimierung und Entwicklung der Materialmodelle, auf Basis gleicher Werkstoffprüfungsvoraussetzungen, werden Versionen bzw. Revisionen erzeugt. Somit wird der zeitliche Entwicklungsprozess, die Reproduzierbarkeit sowie die Nachvollziehbarkeit der Modelle ermöglicht.

5. Zusammenfassung

Zusammenfassend ist es mit dem AMMICAL-Prototyp erstmalig möglich, die Prozesse und Schritte zur Werkstoffdatenbearbeitung und -verarbeitung in einem Werkzeug zu vereinen. Dadurch wird eine Reduzierung der Fehlerquellen und Inkonsistenzen im Vergleich zu derzeitigen Bearbeitungsprozessen erzielt. Der Umgang mit Materialmodellen für diverse Softwares wird vereinfacht und durch die Implementierung des Mastermodells vereinheitlicht. So ist es für den Anwender möglich mit wenig Aufwand verschiedene Modelle auf Basis eines Datensatzes zu generieren.

Materialmodelle sind im PLM-Umfeld und im Rahmen des Digitalen Zwillings zur Abbildung physikalischer Eigenschaften oder auch für die Entstehungsprozesse von Produkten ein wichtiger Bestandteil. Im Produktlebenszyklus können sich die Materialmodelle ändern. Mittels Konfigurationsmanagement in PLM-Systemen sind diese Modelle zu jedem Zeitpunkt reproduzierbar.

Danksagung

Das Forschungs- und Entwicklungsprojekt wird durch das Bundesministerium für Bildung und Forschung (BMBF) im Programm „Innovationen für die Produktion, Dienstleistung und Arbeit von morgen“ (FKZ: 02P17W503) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt dieser Veröffentlichung liegt bei den Autoren. Dank gilt außerdem den Kooperationspartnern im Rahmen des AMMICAL-Projektes.

GEFÖRDERT VOM



**Bundesministerium
für Bildung
und Forschung**

Literaturverzeichnis

- [1] [Online]. Verfügbar unter: <https://mat-plus.de/de/ammical/>
- [2] [Online]. Verfügbar unter: <https://www.inw.hs-mittweida.de/webs/kt/forschung/digitale-produkt-entwicklung/>
- [3] Matplus: „JMatPro – Werkstoffsimulation“, abgerufen am 05.01.2021 unter: <https://mat-plus.de/de/jmatpro/>
- [4] M. Stommel, M. Stojek, W. Korte: „FEM zur Berechnung von Kunststoff- und Elastomerbauteilen“, Hanser Verlag, 2018
- [5] U. Diekmann, N. Herzig, J. Boll, R. Ufer, P. Rostami, I. Alperovich, T. Alder, S. Rzepa: „Towards Integration of Advanced Material Models into PLM“, MEFORM 2020, 2020
- [6] LS-DYNA Support: „Consistent units“, abgerufen am 18.01.2021 unter: <https://www.dynasupport.com/howtos/general/consistent-units>
- [7] T. Kuhn: „Digitaler Zwilling“, Springer-Verlag, 2017

Nachhaltige Produktentwicklung durch Crowdsourcing

Sarah Uhlig, Roman Nuss, André Schneider
Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

Mit einer qualitativen Studie in Form von 36 leitfadengestützten Experteninterviews soll ein entwickeltes Erklärungsmodell ergänzt werden, welches die relevanten Einflussfaktoren auf das Nutzungsverhalten von Crowdsourcing in kleinen und mittelständischen Unternehmen identifizieren soll. Besonderes Augenmerk liegt dabei auf den Antezedenzen, welche die Einstellung zur Nutzung von Crowdsourcing beeinflussen. Die Interviews wurden transkribiert und eine Inhaltsanalyse mithilfe der Analyse-Software MAXQDA durchgeführt. Auf der Grundlage der Ergebnisse dieser Studie kann das Erklärungsmodell um die ermittelten nutzungsfördernden und -hemmenden Gründe konkretisiert werden. Aus praktischer Sicht liefern die Erkenntnisse über die Determinanten der Einstellung und Intention zum Crowdsourcing wichtige Anknüpfungspunkte für die Ableitung von Implikationen für den Einsatz von Crowdsourcing in Unternehmen, welches auch zur nachhaltigen Produktentwicklung genutzt werden kann.

1. Einleitung

Um Produkte nachhaltig zu gestalten, müssen auch die Entwicklungsprozesse in den Unternehmen angepasst werden. Oftmals ist die Entwicklung von Produkten in den Unternehmen ein geschlossener Prozess und die Kunden bleiben dabei außen vor. Das kann dazu führen, dass die Produkte später am Markt scheitern, da vielleicht der Bedarf nicht gegeben ist oder die Kunden ganz andere Anforderungen an das Produkt stellen. Durch die Öffnung des Innovationsprozesses können Kundenwünsche und -bedürfnisse von vornherein mit in die Entwicklung eines Produktes einbezogen und so Kosten für Anpassungen und Ressourcen eingespart werden¹. Der Einbezug von Crowdsourcing im Unternehmen stellt eine Möglichkeit dar, wie man den Innovationsprozess öffnen und potentielle Kunden in die Entwicklungsprozesse einbeziehen kann. Crowdsourcing bezeichnet die Auslagerung von Aufgaben oder Tätigkeiten durch eine Organisation an eine undefinierte Masse an Individuen mittels eines offenen Aufrufs^{2,3}. Mit diesem Open-Innovation-Ansatz soll also die Weisheit der Vielen genutzt und somit versucht werden, bessere Ergebnisse zu erreichen¹.

Ziel der Arbeit ist es, ein theoretisches Modell zu entwickeln, mit dem die Einstellung und Absicht der Nutzung von Crowdsourcing im Innovationsprozess in kleinen und mittelständischen Unternehmen erklärt werden kann, um entsprechende Handlungsempfehlungen abzuleiten. Von ganz besonderem Interesse sind dabei die Antezedenzen, welche die Einstellung zur Nutzung von Crowdsourcing beeinflussen. Da bisher nur wenige Erkenntnisse zum Nutzungsverhalten von Crowdsourcing in kleinen und mittelständischen Unternehmen existieren, sollen zunächst kontextrelevante Einflussfaktoren dieses Verhaltens identifiziert werden. Von der Entwicklung des Totalmodells wird jedoch in der Arbeit abgesehen, da die Nutzungsentscheidung von Crowdsourcing im Innovationsmanagement zu komplex ist. Dementsprechend liegt das Augenmerk im weiteren

Verlauf der Untersuchung auf der Erarbeitung eines Partialmodells, welches ausgewählte Konstrukte des Entscheidungsprozesses zum Crowdsourcing enthält und die kausalen Zusammenhänge zwischen diesen berücksichtigt. Einen geeigneten Erklärungsansatz für die Einstellung zum Crowdsourcing von Unternehmen liefert die Einstellungs-Verhaltens-Forschung⁴. Die Behavioral Reasoning Theory, welche eine Weiterentwicklung der Theorie des geplanten Verhaltens darstellt^{5,6,7,8,9}, wird in der Arbeit als geeigneter integrativer Ansatz zur Erklärung und Prognose des Verhaltens genutzt. Diese Theorie enthält Konstrukte und motivationale Determinanten, die auch für die Untersuchung der Nutzungsabsicht von Crowdsourcing von Bedeutung erscheinen. Um die Konstrukte des theoretischen Partialmodells zu spezifizieren, wurde im ersten Schritt eine Literaturanalyse sowie eine qualitative Untersuchung durchgeführt. Das Hauptaugenmerk lag dabei auf der Identifizierung der Gründe, die für und gegen eine Nutzung von Crowdsourcing im Entwicklungsprozess sprechen. Auf der Grundlage der Ergebnisse dieser ersten Studie kann das Erklärungsmodell um die ermittelten nutzungsfördernden und -hemmenden Gründe konkretisiert werden. Aus praktischer Sicht liefern die Erkenntnisse über die Determinanten der Einstellung und Intention zum Crowdsourcing wichtige Anknüpfungspunkte für die Ableitung von Implikationen für den Einsatz von Crowdsourcing in Unternehmen, welches zur nachhaltigen Produktentwicklung genutzt werden kann.

2. Methode

Für die Ermittlung der Antezedenzen der globalen Motive und der Nutzungsintention wurde eine qualitative Vorstudie in Form von leitfadengestützten problemzentrierten Interviews mit Experten durchgeführt^{10,11,12,13,14}. Auf Grundlage der aus der qualitativen Untersuchung gewonnenen Erkenntnisse resultiert die Ergänzung des Erklärungsmodells,

insbesondere um die fördernden und hemmenden Faktoren des Einsatzes des Crowdsourcing sowie die Konkretisierung der Hypothesen.

Insgesamt konnten bisher 36 Interviews mit Experten aus kleinen und mittelständischen Unternehmen verschiedener Branchen (z. B. Textil/Bekleidung, Maschinen- und Anlagenbau, Software) aus Sachsen geführt werden. Der überwiegende Teil der Gesprächspartner waren die Geschäftsführer selbst (n =28) und das durchschnittliche Geburtsjahr der Teilnehmer war 1980. Für die Akquise der Teilnehmer wurde ein Anschreiben entwickelt, welches über den groben Kern des Interviews, den Ablauf, Informationen zum Datenschutz und Kontaktdaten informiert. Um eine Beeinflussung der Teilnehmer durch das Einholen von Informationen zum Thema zu vermeiden, wurde im Anschreiben keinerlei Bezug zum Thema Crowdsourcing hergestellt. Das Anschreiben wurde per Post an 90 kleine und mittelständische Unternehmen in Sachsen gesandt. Wenn die Unternehmen sich nicht von selbst zurückgemeldet haben, wurde nach einem Zeitraum von mindestens zwei Wochen nochmals telefonisch nachgefragt, ob eine Bereitschaft für die Teilnahme am Interview besteht. Die Interviews haben im Zeitraum Juni 2020 bis Februar 2021 stattgefunden und wurden am Telefon durchgeführt. Ursprünglich war vorgesehen, die Interviews in den Unternehmen von Angesicht zu Angesicht zu führen. Dies konnte jedoch aufgrund der COVID-19-Pandemie in dem genannten Zeitraum nicht umgesetzt werden. Für das Interview wurde ein

Leitfaden entwickelt, welcher der Strukturierung und der zielgerichteten Fragestellung und Informationserhebung diente¹². Mit Erlaubnis der Interviewpartner wurden die Interviews aufgezeichnet und im Anschluss vollständig transkribiert. Es wurden Interviews von 9 bis 45 Minuten gehalten. Die durchschnittliche Interviewzeit lag bei 22 Minuten. Nach der Transkription wurde eine Inhaltsanalyse unter Zuhilfenahme der Analyse-Software MAXQDA durchgeführt.

3. Ergebnisse

Die qualitative Studie hat ergeben, dass die meistgenannten Gründe für den Einsatz von Crowdsourcing im Entwicklungsprozess das Wissen und die Kompetenz der Crowd, das Erfassen von Kundenbedürfnissen, eine Finanzierung durch die Crowd und ein Kostenvorteil für das Unternehmen sind. Die meistgenannten Gründe, welche gegen den Einbezug von Crowdsourcing im Entwicklungsprozess sprechen, sind das Infragestellen der Kompetenz und der Zusammensetzung der Crowd, die Transparenz gegenüber der Konkurrenz und die damit eventuell einhergehende gleichzeitige Schaffung neuer Konkurrenz, der firmeneigene Wissensverlust sowie die Vernachlässigung des eigenen Potentials und der Individualität. In Tabelle 1 und 2 werden die meistgenannten Gründe für und gegen den Einsatz von Crowdsourcing im Entwicklungsprozess nochmals zusammengefasst.

Tabelle 1

Meistgenannte Gründe für die Nutzung von Crowdsourcing im Entwicklungsprozess (N = 36)

| Grund für Nutzung | Beispiel | Häufigkeit, n (%) |
|--------------------------------|---|-------------------|
| Wissen und Kompetenz der Crowd | „Der erste Grund, und das denke ich auch ist der wichtigste, wäre, dass du an ein viel größeres Spektrum an Wissen gelangst, [...]“ | 24 (66.7) |
| Kundenbedürfnisse erfassen | „Zum Beispiel wenn ich mir nicht sicher bin, ob eine Funktion von einem Produkt in einer breiten Masse wirklich benötigt wird, also sowas wie eben Umfragefunktion oder wenn man bei einem gewissen Design nicht weiterkommt, dass man dort eben potentiell noch externe Expertise braucht oder einen Blickwinkel.“ | 11 (30.6) |
| Finanzierung durch Crowd | „Ich könnte mir das auch für mich selbst vorstellen, wenn ich in der Lage wäre, dass ich Geld benötige, würde ich das vielleicht bevorzugen bevor ich zu einer Bank gehe, würde ich das vielleicht auf diese Weise versuchen.“ | 10 (27.8) |
| Produkt-/Leistungsverbesserung | „Ich kann dann in Verbindung mit den Ergebnissen meine Produkte der Lage entsprechend oder der Anforderung entsprechend anpassen. Insofern hab ich da schon was davon.“ | 7 (19.4) |
| Kostenvorteil | „[...] dass man Sachen auslagern kann an eine breite Masse, die dann Bock hat das zu tun und dafür irgendwie weniger oder kein Geld haben wollen, dann wäre es natürlich auch effizienter von der Kostenstruktur her.“ | 7 (19.4) |

Tabelle 2

Meistgenannte Gründe gegen die Nutzung von Crowdsourcing im Entwicklungsprozess (N = 36)

| Grund gegen Nutzung | Beispiel | Häufigkeit, n (%) |
|--|--|-------------------|
| Kompetenz und Zusammensetzung der Crowd | „[...] und dass da heutzutage viele Leute mitmachen oder da vielleicht auch Fakes dann eher rauskommen, sich Leute bloß profilieren, obwohl sie kein Know-How haben, das weiß ich nicht.“ | 20 (55.6) |
| Transparenz für und Schaffung von Konkurrenz | „[...] weil es unter Umständen vielleicht auch Konkurrenzgeschichten gibt, dass also Wettbewerber dort genau oder relativ schnell wissen, was das Ziel ist oder so, oder wie weit man schon ist mit seinen Entwicklungen bevor man es hat schützen lassen.“ | 17 (47.2) |
| Verlust firmeninternes Know-how | „Und da kann es natürlich auch zu Verlusten, Ideenverlusten oder Know-how-Verlusten kommen. Das wäre der Nachteil, den ich sehe, wenn ich das Ganze öffentlich mache.“ | 7 (19.4) |
| Vernachlässigung eigener Potentiale und Individualität | „Dass meine Produkte sehr individuell sind. Also dass sie halt wirklich auf den Einzelnen, auf die einzelne Person, auf den einzelnen Kunden eben genau zugeschnitten werden und es oft eben genau nicht repräsentativ ist, was für die Masse interessant ist, sondern eben für die Einzelperson.“ | 5 (13.9) |

4. Fazit und Ausblick

Die vorliegenden Ergebnisse der qualitativen Studie ermöglichen das Modell zur Erklärung der Intention zur Nutzung des Crowdsourcings um die fördernden und -hemmenden Gründe zu konkretisieren. Im nächsten Schritt der laufenden Forschung erfolgen die Spezifizierung und quantitative Überprüfung der einzelnen Messmodelle sowie des gesamten Modells zur Erklärung der Intention von Unternehmen zur Nutzung des Crowdsourcings.

Im Hinblick auf die Ergebnisse der qualitativen Studie kann zu den Gründen, die gegen Crowdsourcing im Entwicklungsprozess sprechen, gesagt werden, dass beispielsweise die Wahl der Crowdsourcing-Plattform eine wesentliche Rolle spielt. Man muss herausfinden, wo sich die geeignete Crowd aufhält oder ob es überhaupt eine geeignete Plattform gibt. Dafür muss natürlich eine gewisse Recherchearbeit geleistet werden, die vom Unternehmen selbst oder auch von externen Spezialisten übernommen werden kann. Die Wahl der Plattform spielt auch bei der Transparenz gegenüber der Konkurrenz eine Rolle, da es zum Beispiel Plattformen gibt, bei denen Wettbewerber ausgeschlossen werden können. Eine weitere Möglichkeit zum Schutz vor der Konkurrenz wäre, Geheimhaltungsvereinbarungen mit Crowd-Mitgliedern abzuschließen oder auch nur Teile einer Aufgabe zu veröffentlichen, damit das Problem nicht vollständig nachvollzogen werden kann. Um einem Wissensabfluss entgegenzuwirken muss geklärt werden, was preisgegeben und was veröffentlicht werden darf. Die Formulierung der Aufgabe oder Frage und auch die Moderation des Prozesses müssen deshalb gut durchdacht sein. Auch dafür kann man sich an externe Spezialisten wenden oder ähnliche Beispiele bereits

vorhandener Projekte suchen und auf das eigene Projekt ableiten. Die Vernachlässigung eigener Potentiale und der Verlust der Individualität ist ein Grund gegen Crowdsourcing im Entwicklungsprozess, welcher im Zusammenhang damit steht, ob sich das Unternehmen nach außen öffnen und eine Crowd in Prozesse einbeziehen möchte oder ob es überhaupt das richtige Instrument für die jeweilige Aufgaben- oder Fragestellung ist. Die Aufgaben können zwar in Teilaufgaben zerlegt und es muss nicht der komplette Entwicklungsprozess an die Crowd abgegeben werden, aber natürlich muss jedes Unternehmen im ersten Schritt schauen, ob Crowdsourcing im Innovationsprozess für das Unternehmen geeignet und gewollt ist. Die Unternehmenskultur oder auch Unternehmensstrategie spielen dabei eine wichtige Rolle¹⁵.

Die aus der Studie hervorgegangenen meist-genannten Gründe für den Einsatz von Crowdsourcing im Entwicklungsprozess sprechen auch für eine nachhaltige Produktentwicklung. Der Zugriff auf einen größeren Wissenspool kann zum Beispiel bessere Lösungen oder auch bessere Lösungswege hervorbringen. Die Vergütung der Crowd könnte kostengünstiger als die Vergütung eines Fachexperten sein und durch das Erfassen von Kundenbedürfnissen sowie von Produkt- und Leistungsverbesserungen im Entwicklungsprozess können, wie eingangs bereits erwähnt, Änderungen im Nachhinein verhindert und somit Ressourcen eingespart werden¹.

Danksagung

Die Autoren bedanken sich vor allem bei den Unternehmen, die sich trotz Pandemie und ihrem Alltagsgeschäft Zeit genommen und an der qualitativen Studie teilgenommen haben.

Literaturverzeichnis

- [1] Scholz, U., Pastoors, S. (2018): Grundlagen der betrieblichen Nachhaltigkeit. In U. Scholz, S. Pastoors, J.H. Becker, D. Hofmann & R. van Dun (Hg.): Praxishandbuch Nachhaltige Produktentwicklung –Ein Leitfaden mit Tipps zur Entwicklung und Vermarktung nachhaltiger Produkte (S. 9-48). Berlin: Springer Gabler.
- [2] Howe, J. (2006): The Rise of Crowdsourcing. *Wired Magazine*, 14 (6), 1-4.
- [3] Leimeister, J. M. (2012): Crowdsourcing: Crowdfunding, Crowdvoting, Crowdcreation. *Zeitschrift für Controlling und Management*, 56 (6), 388-392.
- [4] Kroeber-Riel, W.; Weinberg, P. (2003): Konsumentenverhalten, 8. Aufl., München.
- [5] Luo, Q.; Oh, L.-B.; Zhang, L.; Chen, J. (2014): Examining the showrooming intention of mobile-assisted shoppers in a multichannel retailing environment, in: Proceedings of the 18th Pacific Asia Conference on Information Systems (PACIS), Paper 141.
- [6] Ajzen, I. (2014): The theory of planned behaviour is alive and well, and not ready to retire: a commentary on Sniehotta, Pesseau, and Araújo-Soares, in: *Health Psychology Review*, <http://dx.doi.org/10.1080/17437199.2014.883474>.
- [7] Ajzen, I. (2012): The theory of planned behavior, in: Lange P. A. M; Kruglanski, A. W.; Higgins, E. T. (Hrsg.): *Handbook of theories of social psychology*, London, pp. 438-459.
- [8] Ajzen, I. (2008): Consumer attitudes and behaviour, in: Haugtvedt, C. P.; Herr, P. M.; Cardes, F. R. (Hrsg.): *Handbook of Consumer Psychology*, New York, pp. 525-548.
- [9] Westaby, J. (2005): Behavioral reasoning theory: Identifying new linkages underlying intentions and behavior, in: *Organizational Behavior and Human Decision Processes*, Vol. 98, No. 2, pp. 97-120.
- [10] Kruse, J. (2014): *Qualitative Interviewforschung. Ein integrativer Ansatz*, Weinheim und Basel.
- [11] Witzel, A.; Reiter, H. (2012): *The Problem-centered interview: Principles and Practice*, London.
- [12] Gläser, J.; Laudel, G. (2010): *Experteninterviews und qualitative Inhaltsanalyse als Instrumente rekonstruierender Untersuchungen*, 4. Aufl., Wiesbaden.
- [13] Jäger, U.; Reinecke, S. (2009): Expertengespräch, in: Baumgarth, C.; Eisend, M.; Evanschitzky, H. (Hrsg.): *Empirische Mastertechniken. Eine anwendungsorientierte Einführung für die Marketing- und Managementforschung*, Wiesbaden, S. 29-76.
- [14] Witzel, A. (2000): Das problemzentrierte Interview, in: *Forum Qualitative Sozialforschung – Theories, Methods, Applications* 1(1), S. 1-9.
- [15] Gassmann, O. (2013): *Crowdsourcing – Innovationsmanagement mit Schwarm-intelligenz*, 2. Aufl., München: Hanser.

Ressourcen- und umwelteffiziente Gestaltung von Fördersystemen mit Kunststoffketten

Jens Sumpf, Lynn Lüdemann, Marcus Bona, Markus Golder

TU Chemnitz, Professur Förder- und Materialflusstechnik, Reichenhainer Str. 70, 09126 Chemnitz

Kurzfassung / Abstract

Förderer mit Kunststoffketten gehören aufgrund der spezifischen Ketteneigenschaften, wie schmierungsfreier/sauberer Betrieb, geringe Kettenmasse oder effiziente Fertigung, zu den wichtigsten intralogistischen Transportsystemen in Produktions- und Distributionseinrichtungen. Während diese Systeme in der Vergangenheit primär nach rein mechanischen und tribologischen Gesichtspunkten gestaltet wurden, erweitern sich die Entwicklungsschwerpunkte heute zunehmend auf eine Verbesserung der Ressourcen- und Umwelteffizienz bei der Herstellung, im Betrieb sowie bei der Gestaltung der End-of-Life Phase der Anlagen. Der Beitrag zeigt aktuelle Entwicklungstendenzen zur Verminderung des ökologischen Fußabdrucks von Kunststoffkettenförderern, aber auch damit verbundene Herausforderungen.

Due to the specific properties of plastic chains, such as lubrication free and clean operation, low mass or efficient manufacturing, conveyors with plastic chains are one of the most important transport systems in production and distribution facilities. Initially, these systems were primarily designed according to mechanical and tribological aspects. Today, however, the improvement of resource and environmental efficiency of conveyor systems in production, in operation and at the end-of-life phase gets under scrutiny more and more. The paper shows current development trends for reducing the ecological footprint of plastic chain conveyors, but also the remaining challenges.

1 Einleitung

Seit Jahren weisen Wissenschaftler aus unterschiedlichen Fachbereichen auf den fortschreitenden Klimawandel sowie die zunehmende Umweltverschmutzung der Erde hin und mahnen nachdrücklich zur Verminderung von Schadstoff- und Treibhausgas-Emissionen sowie zum verantwortungsvollen Umgang mit verfügbaren Ressourcen. Um die die daraus resultierenden Klimaziele zu erreichen, sind die Verwendung recycelbarer Werkstoffe, Energieeffizienz in Fertigung und Produktion sowie Recycling und Wiederverwendung im Rahmen der Kreislaufwirtschaft geeignete Konzepte einer nachhaltigen Produktstrategie, die auch auf Produkte der technischen Intralogistik wie Fördersysteme angewendet werden können. [EU21] [RöS20]

Die Beurteilung der Umweltverträglichkeit eines Produktes erfolgt zumeist auf Grundlage einer sogenannten Ökobilanz, in der alle umweltrelevanten Stoff- und Energieströme über den gesamten Lebenszyklus, von der Produktion über die Nutzungsdauer bis hin zur Entsorgung, berücksichtigt werden. Innerhalb dieser Phasen werden z. B. die Entnahme von Ressourcen (Rohöl, Holz, Erze usw.) sowie die Emissionen in die Umwelt (Schadstoffe, Kohlendioxid usw.) analysiert und im Rahmen einer Wirkungsabschätzung hinsichtlich ausgewählter Kriterien, z. B. dem Beitrag des Produktes zum Treibhauseneffekt, zur Versauerung oder zur Eutrophierung, bewertet.

Der Beitrag beschreibt Entwicklungsansätze zum Aufbau einer entsprechenden Bewertungsmethode auf Grundlage einer Ökobilanzierung und zeigt am Beispiel von Stetigförderern mit Kunststoffketten Möglichkeiten und Herausforderungen zur Verbesserung der gesamtökologischen Wirkung dieser Systeme auf.

2 Stetigförderer mit Kunststoffketten

In modernen Produktions- und Distributionseinrichtungen werden innerbetriebliche Stetigfördersysteme u. a. zur Verkettung von Maschinen und Anlagen oder zum Transport von Waren zwischen Lager-, Kommissionier- und Be- bzw. Enladebereichen eingesetzt. Vermehrt werden dabei Kettenförderer mit Kunststoffketten ausgerüstet, die gegenüber den klassischen Stahlketten eine Reihe von werkstoffspezifischen Vorteilen bieten. Beispielfhaft können eine große Gestaltungsvielfalt und kostengünstige Fertigung, geringe Kettenmasse oder gutes Geräuschdämpfungsvermögen genannt werden. Besonders interessant sind die sehr guten tribologischen Eigenschaften, welche einen schmierungsfreien und damit wartungsarmen und sauberen Betrieb ermöglichen. Nachteilig sind insbesondere die gegenüber metallischen Werkstoffen deutlich geringere Festigkeit/Steifigkeit, zumal diese mechanischen Eigenschaften signifikant von den Anwendungsbedingungen wie Temperatur, Belastungsdauer oder auch Umgebungsmedien und UV-Strahlung abhängig sind.

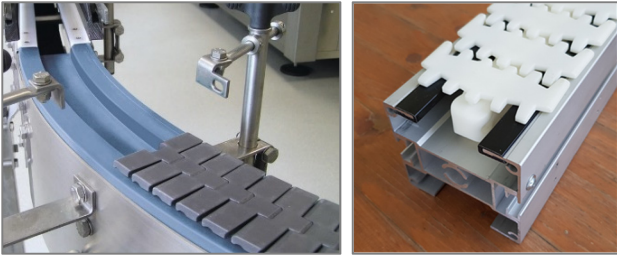


Abbildung 1: Scharnierbandketten (links) und Multiflexketten (rechts) mit Führungselementen aus PE-UHMW

Kunststoffkettenfördersysteme werden in unterschiedlichen Industriebereichen zum Transport kleiner bis mittelschwerer Fördergüter bis ca. 25 kg, aber auch für hohe Gesamtlasten von mehreren Tonnen eingesetzt. Die Leistungsfähigkeit kann insbesondere durch die Erhöhung der Zugfestigkeit der Kette sowie die Reduktion der Reibwerte zur Abstützung verbessert werden. Dementsprechend liegt der Fokus vieler Entwicklungen in erster Linie auf der Verbesserung dieser beiden Aspekte.

Als Werkstoffe für Kunststoffketten werden vorrangig Polyoxymethylen (POM) und Polybutylenterephthalat (PBT) sowie für Mattenketten zusätzlich Polyethylen (PE) und Polypropylen (PP) verwendet. Die Eigenschaften können durch Zugabe von Additiven, z. B. Glas-, Kohle- oder Aramidfasern zur mechanischen Verstärkung sowie PTFE-Pulver, Silikonöl, Ruß oder Wachse zur Verbesserung des Gleitverhaltens, modifiziert werden. Ein weiterer wichtiger Grund für die Zugabe von Additiven ist das Erreichen bestimmter anwendungsspezifischer Eigenschaften, z. B. Brandhemmung, Detektierbarkeit, Wärmeleitfähigkeit oder antistatische Ausrüstung.



Abbildung 2: Kettenglieder mit Zinkdruckgusseinleger (links) [Roh09] sowie in Sandwichstruktur (rechts)

Neben der Verwendung der Kunststoffcompounds ist ein hybrider Aufbau von Kettengliedern verbreitet, bei denen die Funktionsbereiche aus unterschiedlichen Werkstoffen bestehen. Typisch sind dabei die Verwendung von metallischen oder elastomeren Belägen, welche die Tragplatten vor Verschleiß schützen oder die Haftung zum Fördergut verbessern sollen. Zur Erhöhung der Festigkeit und Steifigkeit ist die Integration von Verstärkungsstrukturen aus Metall (Abbildung 2, links) oder Hochleistungsfasern in den Kettenkörper möglich. Grundsätzlich wird dazu eine vorgefertigte Struktur in das Spritzgießwerkzeug eingelegt und mit Kunststoff ummantelt. Der Mehrkomponenten-Spritzgießprozess ermöglicht zudem das gleichzeitige Füllen der Form mit

verschiedenen Werkstoffen, sodass durch geeignete Materialauswahl Kettenelemente mit einem Kern aus einem festen, steifen Material sowie einer dünnen, tribologisch optimierten Haut in einem einzigen Prozessschritt hergestellt werden können (Abbildung 2, rechts).

3 Bewertungsmethoden für eine nachhaltige Produktentwicklung

Zur Bewertung des Ressourcenverbrauchs und des Emissionsausstoßes wird häufig die Methode der Ökobilanzierung nach DIN EN ISO 14040 und 14044 verwendet. In der Ökobilanz werden alle umweltrelevanten Einflussfaktoren eines Produktes über seinem gesamten Lebenszyklus berücksichtigt.

Hierzu muss zunächst der Untersuchungsrahmen, d. h. Funktion des Produktes, Systemgrenzen, getroffene Einschränkungen und Annahmen usw., genau definiert werden. Um die Vergleichbarkeit verschiedener Produkte zu ermöglichen, wird eine sogenannte „funktionellen Einheit“ festgelegt, auf welche die ermittelten Umweltwirkungen bezogen werden. Für Förderer, deren Hauptfunktion der Transport von Gütern über eine bestimmte Strecke ist, eignet sich dafür der bei Transportprozessen häufig verwendete Ansatz von „transportierten Massekilometern“ [t·km].

Im Weiteren werden alle umweltrelevanten Wirkungen im gesamten Produktlebenszyklus ermittelt, der bei der Bereitstellung von Werkstoffen beginnt, welche für die Fertigung und Montage von Zwischenprodukten, einzelnen Bauelementen bis hin zu fertigen Systemen benötigt werden. Es schließt sich eine Nutzungsphase an, an deren Ende die Außerbetriebnahme und Entsorgung steht.

Aus den erfassten Daten werden verschiedene Umweltwirkungen des Produktes berechnet, wobei für eine vereinfachte Abschätzung oftmals der Ansatz des Treibhauseffektes verwendet wird. Treibhausgase, wie Kohlendioxid oder Methan, vermindern in der Atmosphäre die Rückstrahlung der Wärme ins All. Während ein natürlicher Treibhauseffekt essenziell für die Bewohnbarkeit der Erde ist, führen große Mengen Gase aus technischen Prozessen zu einem globalen Temperaturanstieg. Der Treibhauseffekt, auch als „Carbon Footprint“ oder „Global Warming Potential“ (GWP) bezeichnet, wird in sogenannten CO₂-Äquivalenten (CO₂e) angegeben, wobei die Wirkung anderer klimaschädigenden Gase relativ zum Faktor 1 von Kohlendioxid dargestellt wird. So hat z. B. die Emission von Methan eine 30-mal stärkere Wirkung in der Atmosphäre als Kohlendioxid.

Neben den in der Herstellungs- und Betriebsphase erzeugten Emissionen durch Energieverbrauch geraten zunehmend Verwertungsoptionen im Sinne einer Kreislaufwirtschaft in den Fokus. Dabei werden Produkte am Ende ihrer Nutzung nicht mehr entsorgt, sondern durch Wiederverwendung oder Recycling in den Produktions-

kreislauf zurückgebracht. Ein wesentlicher Faktor ist dabei auch die Abfallvermeidung durch längere Nutzbarkeit. [KWG12]

4 Ansätze zur energie- und ressourceneffizienten Gestaltung von Kunststoffketten

4.1 Ausgangslage

Kettenförderer werden derzeit hauptsächlich nach technischen/ technologischen und ökonomischen Gesichtspunkten ausgelegt bzw. ausgewählt. Ökologische Aspekte wurden bisher selten thematisiert und allenfalls reduziert auf den Energiebedarf in der Nutzungsphase sowie die Lebensdauer des Systems.

Der eingangs beschriebene, fortschreitende Klimawandel lässt jedoch den Druck auf Industrie und Verbraucher steigen. Viele Unternehmen werden sich dadurch zunehmend ihrer Eigenverantwortung für die Umwelt bewusst und verpflichteten sich freiwillig zum Mindesteinsatz von recycelten Werkstoffen oder zur umweltneutralen Gestaltung ihrer Prozesse. Letztlich betrifft dies auch die Intralogistik, es werden deshalb Werkzeuge zur objektiven Beurteilung der Umweltwirkung von Transportsystemen benötigt. Es bietet sich dazu die in Abschnitt 3 vorgestellte Methode der Ökobilanzierung an, die an die spezifischen Anforderungen der Förderer-technik angepasst werden muss. [Lüd19]

Langfristig sollen dabei quantitative Aussagen zur Umweltwirkung eines Transportvorgangs, zum Vergleich unterschiedlicher Förderprinzipien, zu Stellhebeln für eine mögliche Verbesserung der Umweltwirkung sowie zur Wirkung von Recyclingprozessen möglich werden.

4.2 Optimierungen in den Lebenszyklusphasen

Wie oben beschrieben, verlangt eine Ökobilanz die Einbeziehung sämtlicher umweltrelevanter Einflüsse im Rahmen des Produktlebenszyklus. Neben der in der Betriebsphase benötigten Energiemenge zur Erfüllung der Förderfunktion müssen dabei u. a. die absoluten Aufwendungen zur Herstellung und Entsorgung der Anlage berücksichtigt werden, also z. B. die Bereitstellung von Aluminium und Stahl, die Fertigung von Profilen, Motoren und Steuerungskomponenten sowie die Demontage und Entsorgung von Mechanik- und Elektronikkomponenten. Infolge der großen Komplexität wird in den folgenden Betrachtungen lediglich auf die Kunststoffkette eingegangen.

Design und Herstellung

In der Designphase wird die Umweltwirkung eines Produktes maßgeblich beeinflusst. Das Hauptpotential betrifft dabei die Optimierung des Rohstoff- und Energieverbrauches sowie die Umsetzung von Prinzipien der Recyclingfähigkeit.

Grundsätzlich ist ein geringer Materialverbrauch vorteilhaft, der durch eine optimale, belastungsgerechte Gestaltung der Ketten erreicht wird. Die Herstellung der

Kettenglieder im Spritzgießverfahren erfordert ein Mindestmaß an Energie, wobei deren Einsatz bei Mehrkomponentenspritzguss steigt. Energetisch noch aufwendiger werden die Fertigungsprozesse, wenn Einlegeteile vorgesehen sind. Dazu muss zunächst der Einleger in einem separaten Prozess gefertigt und später in das Spritzgießwerkzeug eingelegt, dort vorgewärmt und anschließend mit Kunststoff umspritzt werden.

Fest miteinander verbundene Werkstoffe lassen sich zudem in einem späteren Recyclingprozess nicht oder nur mit erheblichem Aufwand materialsortenrein trennen und werden deshalb meistens einer thermischen Verwertung zugeführt. Kettenglieder sollten deshalb aus einer Werkstoffkomponente bestehen. Sind unterschiedliche Werkstoffe aus funktionellen Gründen notwendig, sollte auf Zerlegbarkeit geachtet werden.

Einsatz umweltschonender Materialien

In Kettenförderern werden Kettenglieder aus Stahl oder Kunststoff eingesetzt, die über einen Stahlbolzen miteinander verbunden sind. Kunststoffe werden primär aus fossilen Rohstoffen hergestellt, wobei in der Herstellungsphase bis zu einem verarbeitungsgerechten Granulat klimaschädliche Emissionen freigesetzt werden, die mit dem in Datenbanken verfügbaren GWP-Wert quantifiziert werden können. Beim exemplarischen Vergleich von Scharnierbandketten gleicher Bauart fällt auf, dass die Umweltwirkung zur Produktion beider verarbeitungsfähigen Werkstoffe mit 3,84 kg CO₂e/kg für Edelstahl bzw. 3,20 kg CO₂e/kg für POM-Granulat ähnlich ist. Durch das deutlich geringere spezifische Gewicht der Kunststoffkette von 0,9 kg/m (Stahlkette 2,9 kg/m) verursacht das dafür benötigte Material jedoch etwa 70% weniger Emissionen (Kunststoffkette 3,3 kg CO₂e/m vs. Stahlkette 11,14 kg CO₂e/m).

Zu einer Verringerung des CO₂e Fußabdruckes führt die Verwendung von recycelten Kunststoffen zur Herstellung der Ketten. Dieses Closed-Loop Recycling, bei dem der Werkstoff wieder für die ursprüngliche Verwendung eingesetzt werden soll, ist im Bereich Fördertechnik bisher nicht anzutreffen, wird jedoch derzeit in einem haus-eigenen Projekt untersucht. [Bon20]

Einsparungen könnten sich außerdem ergeben, wenn der auf Rohöl basierende Kunststoff durch biologische Alternativen ersetzt wird. [Wal20] Weitere technologische Entwicklungen fokussieren auf die Verwendung von Restmüll als Ausgangsbasis für Kunststoffe oder die Einbindung von CO₂-Emissionen in das Material (Carbon Capturing) und sparen damit Emissionen ein. [Qua20], [Cov20]

Generell ist die Quantifizierung der Umweltwirkung bei der Verwendung von Rezyklaten aus gebrauchten Kunststoffprodukten als schwierig einzuschätzen. Die Gründe dafür liegen in der Notwendigkeit von Detektions-, Separierungs- und Reinigungsprozessen, die je nach Produkt-

gestaltung und Einsatzbedingung unterschiedlich aufwendig sind, sowie in der Reduktion der physikalischen Gebrauchseigenschaften durch den Recyclingprozess (vgl. Punkt Recycling/ End-of-Life-Phase).

Energieverbrauch in der Betriebsphase

Neben Kettendesign und Materialverbrauch bildet der Energieverbrauch in der Nutzungsphase des Fördersystems den wichtigsten Angriffspunkt zur Reduktion der Emissionen. Dieser lässt sich maßgeblich auf Reibungsverluste, den Energiebedarf beim Beschleunigen und Anheben von Gütern sowie elektrische und mechanische Verluste des Antriebssystems zurückführen.

Tabelle 1: CO₂-Emission unterschiedlicher Ketten- und Layoutvarianten; **Basis:** Kunststoffkette, ebene, gerade Strecke, L = 10 m, v = 0,5 m/s, $\mu = 0,4$ (weitere Reibeinflüsse vernachlässigt), $\eta_{Antr} = 0,6$, Nutzungsdauer t = 10⁴ Stunden, CO₂-Faktor für Strommix Deutschland 2019 = 0,401 kg CO₂/kWh [BMU20]

| Einflussparameter | Motorleistung [W] | Umweltwirkung [kg CO ₂] | Änd. bez. auf Basis [%] |
|--|-------------------|-------------------------------------|-------------------------|
| Basis, Kunststoffkette | 29,4 | 118,0 | 100% |
| Edelstahlkette | 94,8 | 380,3 | 322% |
| geringerer Reibwert ($\mu = 0,2$) | 14,7 | 167,4 | 50% |
| höherer Antriebswirkungsgrad ($\eta_{Antr} = 0,8$) | 22,1 | 121,6 | 75% |
| ansteigende Strecke ($\beta = 10^\circ$) | 41,8 | 419,0 | 142% |
| 90°-Gleitkurve am Streckenanfang | 30,3 | 59,0 | 103% |
| 90°-Gleitkurve am Streckenende | 104,5 | 114,6 | 355% |
| Kurve am Streckenende, $\mu = 0,2$) | 28,6 | 88,5 | 97% |

Der Einfluss verschiedener Parameter soll anhand des einfachen Beispiels aus Tabelle 1 verdeutlicht werden, in dem ein vergleichbar langer Förderabschnitt eine bestimmte Zeit betrieben und als Basis eine horizontale, gerade Förderstrecke mit Kunststoffkette angenommen wird. Die Umweltwirkung bei Variation der entsprechenden Einflussgröße wird auf die Kohlendioxid-Emission reduziert, welche sich aus der insgesamt benötigten Energiemenge und dem spezifischen CO₂-Emissionsfaktor der genutzten Stromsorte, im konkreten Beispiel der Strommix Deutschland, berechnen lässt.

Es wird deutlich, dass sich die CO₂-Emission direkt proportional zur Antriebsleistung, und damit bei gerader

Streckenführung und Vernachlässigung von Nebeneffekten u. a. auch zur bewegten Masse, dem Reibwert oder dem Antriebswirkungsgrad, verhält. Bei Verwendung von Kunststoffketten wirkt sich somit die gegenüber vergleichbaren Stahlketten geringere Masse positiv aus, wobei sich dieser Effekt bei hoher Fördergutmasse deutlich verringert. Eine proportionale Wirkung hat auch die Reduktion von Reibwerten. In diesem Zusammenhang spielt auch die Effizienz des Antriebssystems eine wesentliche Rolle. Hierbei ist zu beachten, dass die Motor- und Getriebekomponenten bei Nennleistung betrieben werden sollten, weil der Wirkungsgrad der meist verwendeten Asynchronmotoren im unteren Teillastbereich massiv sinkt und damit andere Verbesserungen an Wirksamkeit verlieren.

Die Umweltwirkung hängt bei Stetigförderern maßgeblich vom Förderlayout ab. Wie Tabelle 1 zeigt, entsteht bei ansteigender Transportstrecke ein höherer Energiebedarf. Besonders zu beachten ist die Verwendung von horizontalen Gleitkurven, bei denen zusätzliche Reibungsverluste in der radialen Kettenabstützung auftreten. Die Verluste steigen deshalb exponentiell mit dem Reibwert und dem Kurvenwinkel und sind zudem von der Kettenzugkraft am Kurveneingang abhängig, sodass auch die Position der Kurve im Gesamtsystem eine entscheidende Bedeutung hat. Tabelle 1 verdeutlicht, dass bei gleicher Gesamtstrecke die Anordnung der Kurve am Beginn oder am Ende (nahe Motor) der Förderstrecke zu einer Vervielfachung der Antriebsleistung und damit auch der Umweltwirkung führen kann. Aus dem Beispiel geht ebenfalls hervor, dass sich eine Reibwertreduktion im Kurvenbereich überproportional positiv auswirkt.

Verschleiß und Lebensdauer

Generell hat die Nutzungsdauer der untersuchten Produkte einen signifikanten Einfluss auf die Ökobilanz, da sich bei längerer Nutzung die relative Umweltwirkung der Herstellung sowie der End-of-Life-Phase verringert. Schäden an den Ketten, die einen Austausch erforderlich machen, sind z. B. Verschleiß in Gelenken oder an den Kontaktstellen zu den Gleitschienen, thermische Schädigung, Versprödung durch ungünstige Umgebungseinflüsse oder UV-Strahlung bis hin zum Kettenbruch.

Grundsätzlich kann die Umweltbelastung also durch verschleißfeste und mechanisch stabile Ketten verbessert werden. Bezogen auf den Verschleiß ist vor allem eine geringe reibungsbedingte Erwärmung anzustreben, weil die verwendeten thermoplastischen Kunststoffe dadurch signifikant erweichen und an Festigkeit verlieren.

Recycling/ End-of-Life-Phase

Eine Förderanlage enthält auch nach Beendigung der Nutzungsphase noch wertvolle Komponenten und Werkstoffe, die einer weiteren Nutzung zugeführt werden können. Während metallische Gestellkomponenten

wiederverwendet oder recycelt werden, werden gebrauchte Kunststoffketten, welche z. B. beim verschleißbedingten Austausch oder beim Umbau eines Förderers anfallen, derzeit im Restmüll entsorgt und stehen damit lediglich als Energieträger in Müllverbrennungsanlagen zur Verfügung.

Eine werkstoffliche Wiederaufbereitung von Kunststoffketten würde dem Gedanke der Kreislaufwirtschaft entsprechen. Hierbei entfällt der erste Schritt der Kunststoffherstellung aus einem fossilen Rohstoff bis zum verarbeitungsfähigen Material, der Ressourcenverbrauch wird gesenkt und Emissionen durch Verbrennung vermieden. Recycling wird derzeit aus wirtschaftlichen Gründen vorrangig bei Verpackungsmaterialien wie PE, PP oder PET praktiziert, die in großen Mengen als Abfallprodukte vorliegen. Das Recycling von thermoplastischen Konstruktionswerkstoffen wie PA, PBT oder POM erfolgt zumeist nur in Form von Produktionsabfällen (post-industrial recycling), bei dem z. B. Angüsse und Zerspanungsabfälle direkt in den Produktionskreislauf zurückgeführt oder sortenrein gesammelt und regranuliert werden. Dagegen wird eine Rückführung von Altteilen von Endverbrauchern (post-consumer-recycling) in der Regel nicht durchgeführt.



Abbildung 3: Förderketten nach mehrjähriger Anwendung

Die nach dem Gebrauch aus dem Förderer entfernten Kettenstränge sind für eine direkte Weiterverarbeitung ungeeignet. Die genaue Materialzusammensetzung incl. der verwendeten Additive ist dem Anwender in der Regel nicht bekannt und an den Ketten sind nach mehrjährigem Betrieb in unterschiedlichsten Anwendungen oft massive Verunreinigungen festzustellen. Diese reichen von leicht löslichen Verstaubungen bis hin zu klebrigen oder fettigen, stark anhaftenden Belägen, die äußerlich schwer zugänglich sind (Abbildung 3). Der eigentlichen Wiederaufbereitung müssen demnach aufwendige Erkennungs-, Trenn- und Reinigungsprozesse vorgelagert werden. Hierbei spielt das Design der Kette eine wesentliche Rolle, weil z. B. Verstärkungseinleger, mehrkomponentiger Aufbau oder schwer zu detektierende Werkstoffe, die zuverlässige Materialtrennung erschweren oder sogar unmöglich machen. Im Fokus aktueller Untersuchungen stehen deshalb Demontierbarkeit, effektive Entfernung anhaftender Fremdstoffe sowie Untersuchungen, wie sich geringe Mengen Verunreinigungen oder die Vermischung unterschiedlicher Compoundvarianten auf die Eigenschaften der Rezyklate auswirken.

Neben den Verunreinigungen ist die Degradation der Kunststoffe zu beachten. Diese molekularen Zerfallsprozesse können durch physikalische, chemische oder biologische Einflüsse verursacht werden und führen neben optischen Veränderungen zu einer Verschlechterung der mechanischen Eigenschaften, insbesondere durch Versprödung. Im Zusammenhang mit den Einsatzbedingungen der Kunststoffketten sind hierbei vor allem der Abbau durch die mechanische Zug-Schwell-Beanspruchung, hohe oder wechselnde Temperaturen (z. B. in Trockenöfen), bestimmte Umgebungsmedien (z. B. Reinigungsmittel, Säuren, Öle) oder UV-Strahlung zu nennen. Unabhängig von der Anwendung wirkt sich das mehrfache Aufschmelzen bei der Verarbeitung auf die Eigenschaften aus.

Die genannten Aspekte der Verunreinigungen und Degradation führen gegenüber Neumaterialien zwangsläufig zu Rezyklaten mit verminderten Eigenschaften, was jedoch nicht zwingend mit einer schlechten Materialqualität in Verbindung einhergeht. Ein großer Teil der Fördersysteme wird mit sehr geringer Beladung und damit weit unter der zulässigen Belastungsgrenze der Ketten betrieben, in solchen Fällen können problemlos Rezyklate verwendet werden.

Die Bereitstellung von „zertifizierten“ Rezyklaten mit reproduzierbarem Eigenschaftsprofil und entsprechenden Datenblättern, welche zudem explizit die verbesserte Umweltwirkung aufzeigen, ist als vorteilhaft einzuschätzen. Auf diese Weise könnten Bedenken und Vorurteile von potentiellen Anwendern ausgeräumt und der Einsatz von Rezyklaten in industriellen Anwendungen forciert werden.

5 Zusammenfassung und Ausblick

Die in Produktion und Distribution häufig verwendeten Kunststoffkettenförderer werden in erster Linie nach der zulässigen Kettenzugkraft dimensioniert und deshalb häufig auf hohe Zugfestigkeit und geringe Reibung zum Führungssystem ausgerichtet. Ökologische Aspekte spielten bei der Entwicklung bisher eine untergeordnete Rolle, obwohl mit dem Design über einen ressourcen- und umweltschonenden Materialeinsatz entschieden wird und wichtige Grundlagen zur Beeinflussung der Umweltwirkung sowohl in der Nutzungs- als auch der End-of-Life-Phase geschaffen werden.

Die ökologisch höchste Bedeutung bei Stetigförderern hat die Nutzungsphase. Hierbei sind eine hohe Verschleißfestigkeit der Ketten und ein energiesparender Betrieb anzustreben. Der Energiebedarf ist neben den Systemreibwerten auch vom Wirkungsgrad des Antriebssystems sowie vom Förderverlauf abhängig und erfordert deshalb eine vorausschauende Layoutgestaltung und leistungsgerechte Motorauslegung.

Ein wesentlicher Einflussfaktor ist auch die Verwertung gebrauchter Kunststoffketten im Sinne einer Kreislauf-

wirtschaft, d. h. die Werkstoffe werden nach ihrer Nutzung wieder in den Produktionskreislauf zurückgeführt und sparen deshalb Ressourcen und herstellungsbedingte Emissionen. Ein ökologischer und gleichzeitig ökonomischer Recyclingprozess von Gebrauchtteilen kann jedoch nur mit effizienten, automatisierbaren Technologien zur Demontage, Reinigung und sortenreinen Trennung der Bauteile funktionieren. Wichtige Grundvoraussetzungen werden bereits bei der Gestaltung gelegt und können u. a. durch Demontierbarkeit, Vermeidung von Mehrkomponentenbauteilen oder Materialkennzeichnung positiv beeinflusst werden. Vorbehalte gegen die Verwendung von Rezyklaten betreffen häufig die gegenüber Neuware verringerten Eigenschaften aufgrund von polymeren Abbauprozessen sowie fehlender Kennwerte. Hier ist die Deklaration reproduzierbarer Materialeigenschaften in Verbindung mit der klaren Ausweisung der Umweltwirkung erforderlich, um die Marktakzeptanz recycelter Werkstoff weiter zu verbessern.

Zur Forcierung der Umsetzung einer Kreislaufwirtschaft bei technischen Produkten wie Förderketten fehlen derzeit noch wirksame Konzepte und Anreize. So sind Rezyklate bei vergleichbarer Qualität oft deutlich teurer als Neuware und es stehen recyclingfähige Kunststoffe nicht in ausreichendem Maß zur Verfügung, weil gebrauchte Produkte nicht einer Wiederaufbereitung zugeführt werden. Gute Ansätze sind gesetzliche Regelungen, z. B. die bereits in der EU beschlossene Steuer für Neuware oder eine Kennzeichnungspflicht der Umweltwirkung, sowie auch die zunehmende Bereitschaft von Unternehmen zu einer Selbstverpflichtung zur CO₂-neutralen Produktion oder zum Mindesteinsatz von Rezyklaten in ihren Produkten.

Danksagung

Die Autoren bedanken sich bei der Röchling Stiftung Mannheim für die finanzielle Unterstützung der Forschungen.

Literaturverzeichnis

- [BMU20] Umweltbundesamt: Energiebedingte Emissionen, <https://www.umweltbundesamt.de/daten/energie/energiebedingte-emissionen#energiebedingte-treibhausgas-emissionen>, Abruf 09.02.2021.
- [Bon29] Bona, M. et al.: Recycling von Kunststoffgleitketten. Logistik Werkstatt Graz Tagungsband, S. 43-64, Graz, November 2020, DOI 10.3217/978-3-85125-742-7.
- [Cov20] Covestro: Neuer Rohstoff CO₂ auf dem Weg zum Alleskönner. <https://www.covestro.com/de/sustainability/lighthouse-projects/co2-dreams>, Abruf 09.02.2021.

- [EU21] Europäische Kommission: 2030 climate & energy framework. https://ec.europa.eu/clima/policies/strategies/2030_en, Abruf 06.02.2021.
- [KWG12] Kreislaufwirtschaftsgesetz vom 24. Februar 2012 (BGBl. I S. 212), letzte Änderung 09.12.2020 (BGBl. I S. 2873).
- [Lüd19] Lüdemann, L. et al.: Development of a Product Category Rule (PCR) for Environmental Product Declaration (EPD) of Conveyor Chain Systems. Proceedings of XXII International Conference MHCL 2019, S. 127-134, Vienna, September 2019, ISBN 978-86-6060-020-4.
- [Qua20] Quantis SA: UBQ Materials. www.quantis-intl.com/casestudy/ubq-materials, Abruf 09.02.2021
- [Roh09] Rohne, C. et al.: Hybrid Conveyor Chains - Calculation, Design and Manufacturing. International Symposium Plastic-Slide-Chains and Tribology in Conveyor Systems: Proceedings, Vol. 3, April 2017, Chemnitz. ISBN 978-3-945479-08-7, p. 65-77, URN: urn:nbn:de:bsz:ch1-qucosa-231781.
- [RöS20] Röchling-Stiftung: PolyProblem-Report 2020. www.polyproblem.org, Abruf 09.02.2021.
- [Wal20] Walker, S., Rothman R.: Life Cycle Assessment of bio-based and fossil-based plastic: A review. Journal of Cleaner Production 261 (2020) 121158, <https://doi.org/10.1016/j.jclepro.2020.121158#>.

Impedimetrische Sensoren zur Überwachung von Deichbauwerken

Wolfgang Fichtner¹, Michael Mertig^{1,2}, Dirk Fleischer³, Ulf Helbig³

1) Kurt-Schwabe-Institut für Mess- und Sensortechnik Meinsberg e.V.

Kurt-Schwabe-Straße 4, 04736 Waldheim

2) Institut für Physikalische Chemie, Technische Universität Dresden

01062 Dresden

3) Institut für Wasserbau und Technische Hydromechanik, Technische Universität Dresden

01062 Dresden

Deichbauwerke sind bedingt durch Auswirkungen des Klimawandels erhöhten Anforderungen hinsichtlich ihrer Standfestigkeit ausgesetzt. Beispielsweise können lange niederschlagsfreie Phasen zu irreversiblen Rissbildungen führen und von zunehmend volatilen Starkregenereignissen flankiert werden. Preiswert herzustellende Sensoren, eingebettet in repräsentativen Bereichen des Deiches, ermöglichen die Online-Überwachung von Durchfeuchtungszuständen. Als Elektrodenmaterialien kommen Karbonfasermatten in spezieller Konfektionierung zum Einsatz, die für den Praxiseinsatz bei der Sanierung oder den Neuaufbau von Deichen geeignet sind. Die Entwicklung der Messtechnik am KSI Meinsberg wird von Untersuchungen begleitet, bei der die Sensoren in Experimenten an einem Versuchsdeich an der TU Dresden mit vorgegebenen Einstauszenarien getestet werden. Die Ergebnisse zeigen, dass impedimetrische Sensoren den Wassergehalt von Deichbaumaterialien realistisch abbilden können.

Due to the effects of climate change, dykes will be exposed to increasing demands on their stability. For example, long rainless periods accompanied by irreversible erosion effects might be followed by increasing volatile heavy rain events. Low-cost sensors, embedded in representative areas of a dyke, enable online monitoring of moisture penetration in the dyke. Woven carbon fabric mats in special packaging are used as electrode materials in form of measuring devices suitable for practical use when installed during renovation or rebuilding of dykes. The development of the measurement technology at the KSI Meinsberg accompanied by experiments on a test dyke with specified irrigation at TU Dresden. The obtained results show that the impedimetric sensor devices can display the water content of dyke construction materials in realistic manner.

1. Einleitung

Die Standsicherheit von Deich- und Dammbauwerken hängt wesentlich von den hydrologischen Verhältnissen in deren Inneren ab (Abb. 1).

Aufgrund eines hydraulischen Gradienten an der Deichoberfläche (Einstau, Niederschlag) wird der Deich infiltriert, wobei sich eine wassergesättigte Zone ausbildet, nach oben begrenzt durch die Sickerlinie. Deren Verlauf ist von dem Gradienten, geohydraulischen Materialparametern, der Verteilung der Permeabilität (K_f -Wert), den Wasserverhältnissen im Untergrund sowie der Einstauhöhe und -zeit abhängig. Dynamisch veränderliche Gradienten führen zu instationären Sickerlinienerläufen mit starken zeitlichen Schwankungen. Während unterhalb der Sickerlinie die Sättigung des Materials gegen 100 % geht, ist der Bereich oberhalb der Sickerlinie (= Spiegellinie) nur teilgesättigt. Ein für die Standsicherheit eines Deichbauwerkes meist kritischer Zustand liegt vor, wenn die Sickerlinie auf der Luftseite aus der Böschung an die Oberfläche tritt. Dann strömt dort Wasser aus dem Deichkörper, trägt Erdmaterial mit sich, was zu rückschreitender Erosion und in der Folge zum Deichversagen führt. Besteht die Möglichkeit, durch eine in den Deich eingebrachte Sen-

sorik den Anstieg der Feuchte im teilgesättigten Bereich zu messen, können ein gefährdender Anstieg der Sickerlinie hin zu einem kritischen Zustand frühzeitig erkannt und somit rechtzeitig Gegenmaßnahmen ergriffen werden.

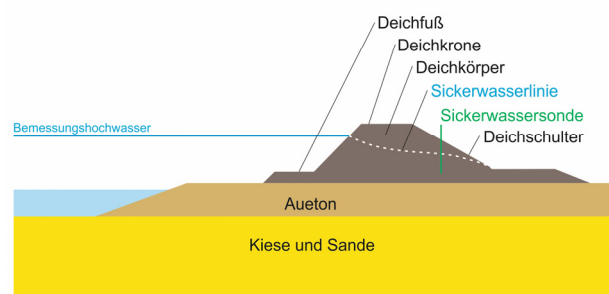


Abbildung 1: Querschnitt durch ein Deichbauwerk mit Verlauf der Sickerlinie zwischen Wasser- und Luftseite (schematisch).

2. Experimentelles

Die Erfassung der Bodenfeuchte des Deichmaterials soll durch einen impedimetrischen Sensor erfolgen. Die zu entwickelnden Sensorelemente sollen in der Fertigung preiswert, langzeitstabil und robust hinsichtlich

Einbau und Betrieb sein. Sie sollen nur umweltneutrale Konstruktionsmaterialien beinhalten. Weiter ist zu berücksichtigen, dass im künftigen Praxiseinsatz eine Tiefbaufirma die Einbringung in den Deich vornehmen wird und die Konstruktion dabei eine mechanische Verdichtung unbeschadet überstehen muss (Baustellentauglichkeit).

2.1. Impedanzmessung

Als Messprinzip wird die Impedanzmessung [1, 2] realisiert, bei der sich das hinsichtlich des Wassergehaltes zu charakterisierende Bodenmaterial zwischen zwei Elektroden befindet (Abb. 2).

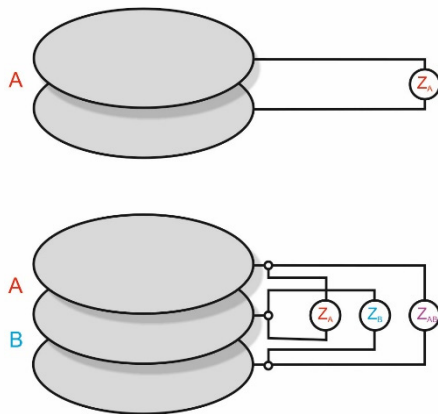


Abbildung 2: Elektrodenanordnung zur Impedanzmessung in Deichbaumaterialien. Oben: 2-Elektrodenanordnung A. Unten: 3-Elektrodenanordnung AB.

2.2. Sensoraufbau

Die Sensorelemente sind als scheibenförmige Elektrodenanordnung ausgeführt, die in horizontaler Einbaulage am Messort eingebracht werden. Als Elektrodenmaterial kommen Kohlenstoffasermatten zum Einsatz. Diese zeichnen sich durch hohe elektrische Leitfähigkeit, mechanische Flexibilität und Robustheit sowie gute Durchlassfähigkeit für Bodenfeuchte aus. Die textilen Matten werden kommerziell in den Gewebearten Leinwand und Körper angeboten. Die Kohlenstoffmatten in Leinwandbindung weisen im Vergleich zu den in Körperbindung eine höhere Festigkeit bei ähnlicher Flexibilität auf und wurden für die weiteren Versuche eingesetzt.

Stützendes Element des Elektrodenaufbaus ist ein PVC-Ring der Nennweiten 100 mm bzw. 125 mm, dessen Höhe den Elektrodenabstand vorgibt (Abb. 4). Die Ringe werden aus einem Brunnenfilterrohr (PVC-U) entsprechender Nennweite geschnitten. Die so gefertigten Ringe weisen 0,3 mm breite Schlitz in einem vertikalen Abstand von 5 mm auf. Damit ist auch der horizontale Transport von Wasser durch den Elektrodenaufbau gewährleistet.

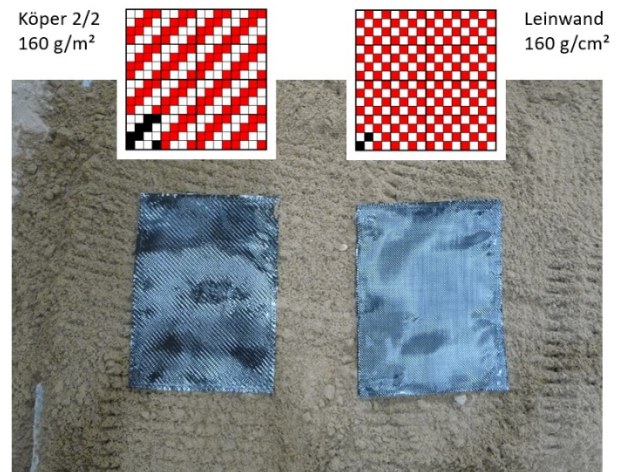


Abbildung 3: Kohlenstoffasergewebe des Untersuchungsprogrammes.

Die Elektrodenmatten werden mit den Stirnflächen des Ringes verklebt und mit geschirmtem Doppelkoaxialkabel kontaktiert. Der Übergang der Kohlenstofffaser auf die Kupferlitze des Messkabels wird durch eine Klemmverbindung mit integrierter Kabelzugentlastung hergestellt. Diese ist gegenüber Einwirkung der Erdfeuchte geschützt angeordnet. Im Inneren ist der Sensor mit dem Mineralgemisch des vorgesehenen Einsatzortes gefüllt.

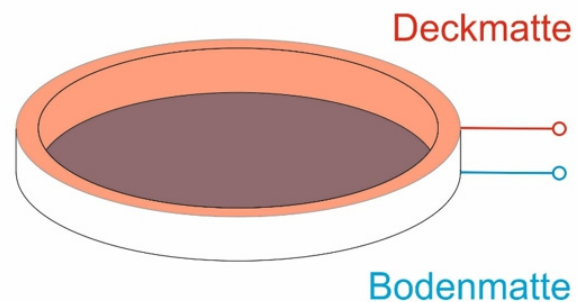


Abbildung 4: Konfektionierung eines Elektrodenpaares.

2.3. Labormessungen

Die Messung der Impedanzspektren erfolgte mit einem Potentiostaten *Interface 1010 - 25019* im Frequenzbereich von 0,1 Hz bis 100 kHz. Die Verwendung eines elektrochemischen Multiplexer *ECM-8* (beide *Gamry Instruments, USA*) ermöglichte die quasisimultane Messung an 8 Kanälen. Somit konnten in einer Laboranordnung bei Vorversuchen Langzeitmessungen an mehreren Sensoren gleichzeitig durchgeführt werden. Im Rahmen der Untersuchungen wurden Aufbauten mit Ringdurchmesser 100 mm und 125 mm, einer Ringhöhe von 10 mm und 20 mm, Zwei- und Drei-Elektrodenanordnungen sowie Aufbauten mit zusätzlichen seitlichen Bohrungen im Ring für verbesserten horizontalen Feuchtigkeitsaustausch erprobt (Abb. 5). Die Sensoren werden vor Verschließen mit angefeuch-

tetem Mineralgemisch gefüllt. Dabei wird die Proctordichte gemäß (DIN 18127) angestrebt (hier: 10–11 % Wassergehalt). Bei diesem Wassergehalt ist eine optimale mechanische Verdichtung möglich, die dem Deichmaterial am nächsten kommt. Die Sensoren wurden für die Laborversuche in einer Messbox in Mineralgemisch (Rossendorfer Sand) eingebettet. Das Behältnis war geschlossen und der Wassergehalt der Füllung konnte durch Zugabe bzw. gezielte Verdunstung von Wasser verändert werden. Die Bestimmung der Feuchte erfolgte dabei gravimetrisch durch Ofentrocknung.

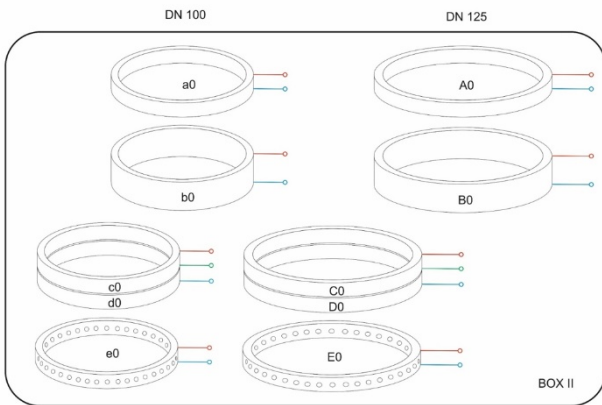


Abb. 5: Anordnung von acht Sensoren in einer Versuchsanordnung mit Mineralgemisch (schematisch).

2.4 Messungen am Versuchsdeich

Zum Test der Funktionstüchtigkeit der Sensoren wurden 10 Exemplare in einen Versuchsdeich der Technischen Universität Dresden (Hubert-Engels-Labor) eingebaut [3]. Die Positionierung erfolgte in drei verschiedenen Ebenen (I, II und III) gemäß Abb. 6. Einige Sensoren wurden paarweise nebeneinander angeordnet, bspw. A13 und A4, um eine Dreidimensionalität der Deichdurchströmung untersuchen zu können. Die Einbaupositionen waren unterschiedlich weit von der Wasserseite entfernt (A13 und A14: 1 m, A1, A2, A11, A12 und A15: 2,25 m sowie A3, A4 und A5: 4 m). Die Deichhöhe beträgt mitsamt Bodenplatte 2 m bei einer Gesamtbreite am Deichfuß von 6 m. Es wurden verschiedene Einstauszenarien von bis zu 96 h Dauer untersucht mit Scheitelhöhen des Wasserstandes von 95 cm (Abb. 7) und 120 cm. Die Impedanzspektren sind während der Einstauvorgänge an 8 Kanälen mit einer Abtastrate von 1 h erfasst worden.

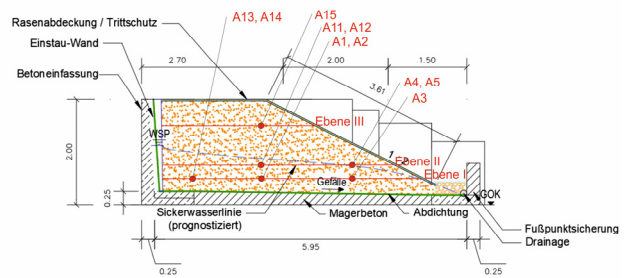


Abb. 6: Querschnitt durch einen Versuchsdeich mit Lage der Sensoren A1 bis A15 in den Ebenen I, II und III.

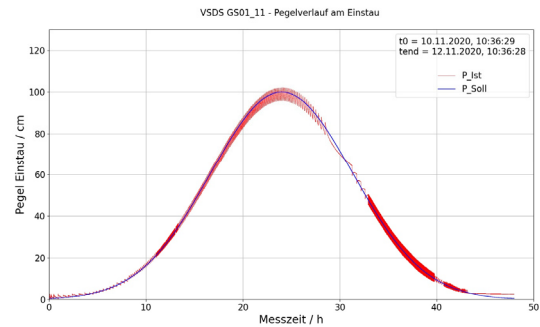


Abb. 7: Pegelverlauf eines Einstauvorgangs.

3. Ergebnisse

3.1. Labor

Wie die Ergebnisse der Laborversuche zeigen, weist das Messverfahren der Impedanzmessung mit Kohlenstoffmatten eine hohe Empfindlichkeit gegenüber der Änderung des Wassergehalts in Bodenmaterialien auf (Abb. 8). Bereits die Simulation von 1 mm Niederschlag (1 l/m²) führt nach einer Einsickerzeit von 4 h zu einer Verringerung der Impedanzwerte um ca. 10 %. Nach 16 h Ausbreitung der zugegebenen Wassermenge sinkt der Wert um weitere 20 % ab.

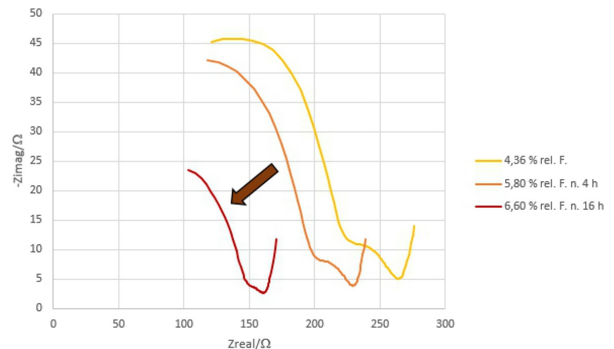


Abb. 8: Änderung der Impedanzwerte (Nyquist-Darstellung) gemessen zwischen zwei Elektrodenmatten im Abstand von 2 cm bei Zugabe von 1 mm Niederschlag.

3.2 Versuchsdeich

Die Sensoren in der unteren Einbauebene I werden stark von den Einstauvorgängen beeinflusst. Zunächst sinken nach 14 h die Impedanzwerte (Modulus der komplexen Impedanz in Abb. 9) rasch auf ein Tiefplateau ab. Zu diesem Zeitpunkt erreicht die Sättigung

100 %; die Sensoren befinden sich nun unterhalb der Sickerlinie. Nach ca. 44 h steigen die Impedanzwerte langsam wieder an, der Sensor befindet sich nun oberhalb der Sickerlinie im teilgesättigten Bereich.

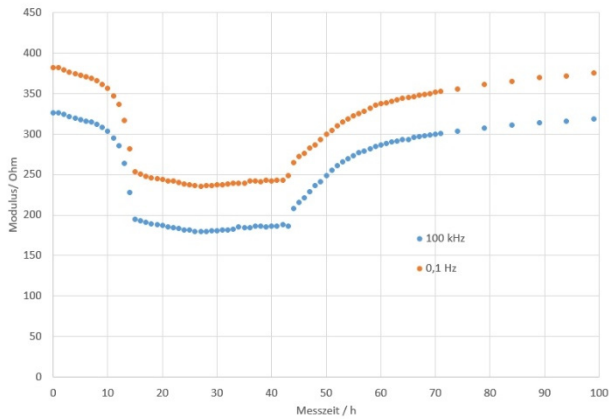


Abb. 9: Zeitlicher Verlauf des Modulus der komplexen Impedanz, Sensor A1 bei Einstaupegel 95 cm.

Im Verlaufsdigramm sind aus Gründen der Übersichtlichkeit nur die Moduluswerte für die kleinste und die größte Messfrequenz dargestellt. Das komplette Impedanzspektrum zwischen 0,1 Hz und 100 kHz zeigt ein analoges Verhalten. Der in Ebene II höher als Sensor A1 angeordnete Sensor A11 wird in geringerem Maße von dem Einstau beeinflusst. Hier tangiert die Sickerlinie die Einbaulage nur im Zeitraum von 24 h bis 29 h (Abb. 10). Es kann daher wie erwartet im Kurvenverlauf nur ein Tiefplateau in abgeschwächter Form ausgebildet werden.

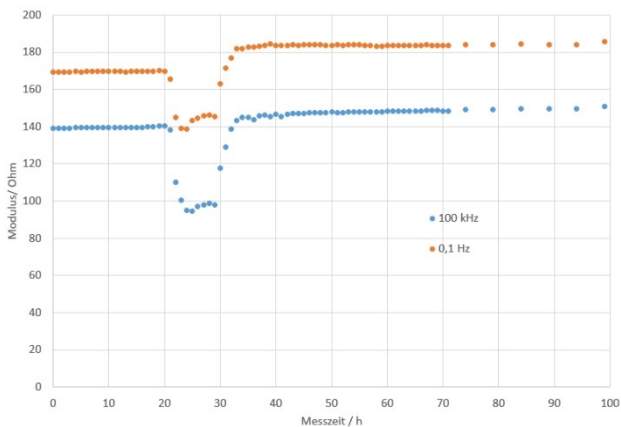


Abb. 10: Zeitlicher Verlauf des Modulus der komplexen Impedanz, Sensor A11 bei Einstaupegel 95 cm.

4. Zusammenfassung

Die Impedanzmessung mittels Kohlenstoffelektrodenmatten zur Bestimmung der Feuchte von Deichbaumaterialien führt zu Ergebnissen, die den Einsatz an realen Bauwerken aussichtsreich erscheinen lassen. Untersuchungen an einem Versuchsdeich zeigen, dass sich mittels der neuentwickelten Sensoren punktuelle Veränderungen des Wassergehaltes in der ungesättigten Zone früh qualitativ erkennen lassen.

Danksagung

Die Untersuchungen wurden im Rahmen eines geförderten SAB-Projektes VSDS (Förderkennzeichen: 100362350) durchgeführt. Die Autoren bedanken sich bei der Sächsischen Aufbaubank – Förderbank (SAB) für die finanzielle Unterstützung. Diese Maßnahme wird mitfinanziert durch Steuermittel auf der Grundlage des vom Sächsischen Landtag beschlossenen Haushaltes. Die Autoren danken weiterhin den Projektpartnern M&S Umweltprojekt GmbH, Institut für Grundwasserwirtschaft der TU Dresden und der Weischlitzer Tiefbau und Umweltschutz GmbH für die gute Zusammenarbeit.

Literaturverzeichnis

- [1] W. Fichtner, U. Guth, D. Becker, K.-K. Jüttner, Orts- und zeitaufgelöste Impedanzmessung zum Online-Monitoring in Deponien mit Tonabdichtung. In: G. Gerlach, K. Kaden (Hrsg.) Dresdner Beiträge zur Sensorik. Band 24 (2005) 307.
- [2] J. Posseckardt, C. Schirmer, W. Fichtner, M. Mertig, Impedimetrische Bestimmung der metabolischen Aktivität von *Saccharomyces cerevisiae* mit Platindickschichtelektroden. Tagungsband 18. Heiligenstädter Kolloquium, 19.9.-21.9.2016, Heilbad Heiligenstadt.
- [3] D. Fleischer, U. Helbig, V. Kostkanová, G. Tintelnot, T. Hohlfeld, Digitale Chancen in der Messtechnik – LowCost-Ansätze für komplexe Messaufgaben am Beispiel des Modelldeichs der TU Dresden. 43. Dresdner Wasserbaukolloquium, in: Dresdner Wasserbauliche Mitteilungen, Heft 63 (2020) 245-254.

Kunststoffgleitlager mit intelligenter Verschleißüberwachung

Norman Katzer, Hagen Bankwitz, Jörg Hübler
Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

Die Zustandsüberwachung (Condition Monitoring) und die vorausschauende Wartung (Predicted Maintenance) gelten als Schlüsselinnovationen der Industrie 4.0. Im Zuge dessen arbeiten Forscher der Professur Intelligente Maschinensysteme an einem Kunststoffgleitlager, welches eine integrierte Sensorik besitzt, die dem Anwender die Überwachung von Betriebsdaten ermöglichen soll. Die aus einem elektrisch leitfähigen Kunststoff hergestellte Sensorik trägt dazu bei, dass in Echtzeit sowohl eine Aussage zur Lagertemperatur als auch zum Verschleißgrad des Lagers getroffen werden kann. Damit lassen sich Wartungsintervalle besser planen und ein prophylaktischer Austausch von noch gebrauchsfähigen Lagern kann vermieden und damit Kosten gesenkt werden. Dafür sind umfangreiche numerische Untersuchungen zum mechanischen, thermischen und elektrischen Verhalten des Gleitlagers durchgeführt worden. Außerdem sind die Auswertelektronik und Werkstoffe entsprechend den Anforderungen der Sensorik entwickelt bzw. ausgewählt worden.

1. Einleitung

Kunststoffgleitlager (Abb. 1) besitzen ein riesiges branchenübergreifendes Anwendungsgebiet wie zum Beispiel in der Agrartechnik, der Automobilindustrie, der Baumaschinenbranche und der Luftfahrtindustrie. Selbst in Küchenmaschinen oder im Fitnessbereich sind Kunststoffgleitlager ein alltäglicher Begleiter. Ihre Popularität verdanken sie dabei der kostengünstigen Herstellung, der Beständigkeit gegenüber einer Vielzahl von Medien, dem schmierungsfreien Betrieb und der Toleranz gegenüber ungleichmäßiger Lasteinleitung. Aufgrund der Optimierung der tribologischen Eigenschaften der Kunststoffe, verdrängen sie dabei auch immer häufiger metallische Lager. Die Entwicklung hin zu höherer Temperaturbeständigkeit, höherer Festigkeit und besserem Reibungsverhalten führt dazu, dass Kunststoffe mittlerweile auch in hochbelasteten Reibkontakten Anwendung finden. Als hochbelastet wird dabei die Kombination aus Gleitgeschwindigkeit, Flächenpressung und Umgebungstemperatur bezeichnet. [1]

Problematisch dabei ist die Tatsache, dass das Reibungs- und Verschleißverhalten der Kunststoffe unter Einsatzbedingungen immer noch nur ungenau vorherzusagen ist. Dies begründet sich in der starken Abhängigkeit der mechanischen Eigenschaften von der Temperatur, welche sich wiederum auf den Reibungswert und den Verschleiß auswirken. Da die Reibung im Anwendungsfall zu einem Wärmeeintrag in der Kontaktstelle führt, verändern sich aufgrund der Temperaturerhöhung auch die mechanischen Eigenschaften der Kontaktbereiche. Dies führt zu einem dynamischen Reibungs- und Verschleißverhalten, welches durch die schlechte Wärmeleitfähigkeit der Kunststoffe noch verstärkt wird. [2]

Infolgedessen lassen sich die Ergebnisse aus Laborversuchen nur selten auf den realen Anwendungsfall übertragen. Eine Übertragbarkeit ist nur dann gegeben, wenn die Beanspruchung und die Geometrie im Wesentlichen mit dem realen Anwendungsfall übereinstimmen.

Eine exakte Lagerdimensionierung ist somit nicht möglich. Daraus resultiert, dass die Lebensdauer des Lagers nicht bestimmt werden kann und Austauschintervalle damit nicht auf einen konkreten Zeitpunkt festgesetzt werden können. Solche Intervalle müssen bei hochproduktiven Maschinen und Anlagen jedoch planbar sein, um Stillstandzeiten und damit verbundene hohe Kosten zu minimieren.

Im Zuge dessen soll ein sensorintegriertes Kunststoffgleitlager entwickelt werden, welches eine zustandsorientierte Instandhaltung ermöglicht. Im Gegensatz zu den bisher üblichen reaktiven (Austausch nach Defekt) und präventiven (Austausch nach festgelegtem Intervall) Instandhaltungsmethoden, können damit Folgekosten durch Lagerschäden vermieden und vorhandene Restlaufzeiten optimal ausgenutzt werden. Die zusätzlich gewonnene Betriebszeit führt zu einer gesteigerten technischen Verfügbarkeit des gesamten technischen Systems und damit letztlich auch zu einer Produktivitätssteigerung. [3]



Abbildung 1: Kunststoffgleitlager [4]

2. Funktionsprinzip

Das Kunststoffgleitlager besteht in seinem Grundaufbau (Abb. 2) aus zwei unterschiedlichen Materialien. Eines der beiden Materialien (in der Abbildung rot dargestellt)

besitzt elektrisch leitfähige Eigenschaften und wird daher als Sensorelement benutzt. Dieses Sensorelement wird in einem Spritzgussprozess in das zweite elektrisch nicht leitfähige Material eingebettet. Das zweite Material (grün) dient als Verschleißschicht mit guten tribologischen Eigenschaften.

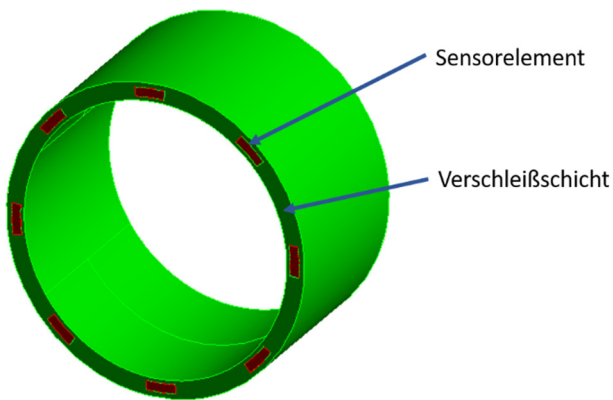


Abbildung 2: Schematischer Aufbau

Die Verschleißschicht, die in direktem Kontakt zur Welle steht, wird mit zunehmender Betriebsdauer verschleiben. Bei ausreichend großem Materialverlust der Verschleißschicht, werden die elektrisch leitfähigen Sensorelemente mit der Welle in Kontakt treten. Die mit einer Spannung beaufschlagten Sensorelemente werden durch die elektrisch leitfähige Welle kurzgeschlossen (Abb. 3). Damit kann ein Signal erzeugt werden, welches dem Anwender signalisiert, dass das Lager einen bestimmten Verschleißgrad überschritten hat und dementsprechend ausgetauscht werden sollte. Da das elektrisch leitfähige Material ähnliche tribologische Eigenschaften wie die elektrisch nichtleitende Verschleißschicht hat, kann das Lager noch eine gewisse Zeit ohne Probleme weiter in Betrieb bleiben und die Wartung zu einem günstigen Zeitpunkt geplant und durchgeführt werden.

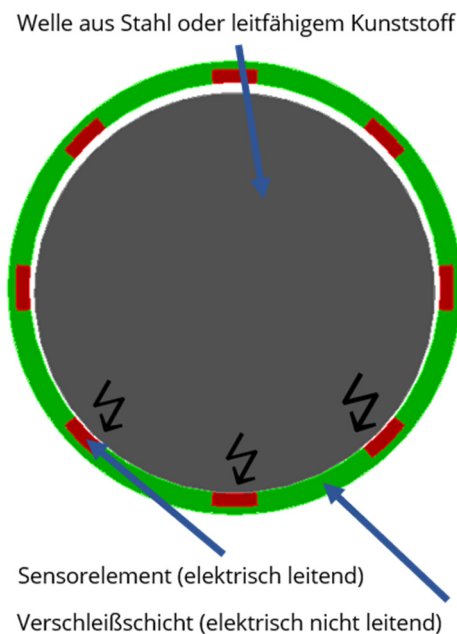


Abbildung 3: Kurzschluss im Verschleißfall

Weiterhin kann im Betrieb bei noch intakter Verschleißschicht die Temperatur mittels einer Widerstandsmessung des elektrisch leitfähigen Materials bestimmt werden. Dazu wird der physikalische Effekt genutzt, dass sich die elektrische Leitfähigkeit mit der Temperatur ändert. Damit können unvorhersehbare Ausfälle, die aus einer Überhitzung des Materials resultieren, vermieden werden. Die temperaturabhängigen Belastungsgrenzen sind materialabhängig und können ermittelt werden. Wird das Kunststoffgleitlager unterhalb dieser Grenze betrieben, ist es zuverlässig sehr lange einsetzbar. Wird die thermische Belastungsgrenze hingegen überschritten, fällt es mit einer bestimmten Wahrscheinlichkeit abhängig von der Last und der Dauer der Temperaturüberschreitung zeitnah aus.

3. Numerische Untersuchungen

Um ein besseres Verständnis der mechanischen, thermischen und elektrischen Eigenschaften des Systems Gleitlager-Welle zu bekommen, wurden umfangreiche numerische Untersuchungen durchgeführt. Dazu wurde ein Simulationsmodell aus einer Welle, dem Kunststoffgleitlager mit integrierter Verschleißsensorik und einem Lagergehäuse aufgebaut (Abb. 4). Die Randbedingungen wurden so gewählt, dass das Gehäuse auf der Unterseite fixiert ist. Außerdem wurde die rotatorische und axiale Bewegung der Welle blockiert, um eine stabile Konvergenz der Berechnung zu gewährleisten. In der Mitte der Welle wurde eine Kraft von 60 kN appliziert, welche beispielhaft dem Anwendungsfall entspricht. Um Rechenzeit zu sparen, wurde die Symmetrie des Modellaufbaus ausgenutzt und die Berechnung auf ein Halbmodell beschränkt.

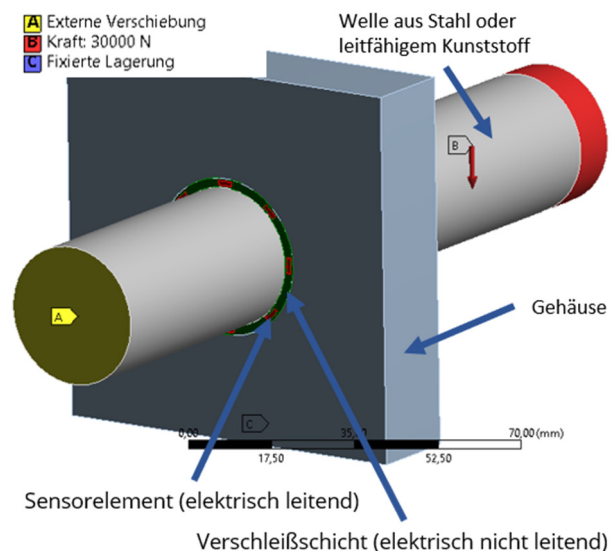


Abbildung 4: Simulationsmodell mit Randbedingungen

Mithilfe der statisch-mechanischen Berechnung konnten erste Erkenntnisse über das Deformationsverhalten des Systems Gleitlager-Welle gewonnen werden. Dabei wurde ersichtlich, dass es auf der Lagerober- und -unterseite hälftig zur Ausprägung einer Flächenpressung kommt (Abb. 5 links).

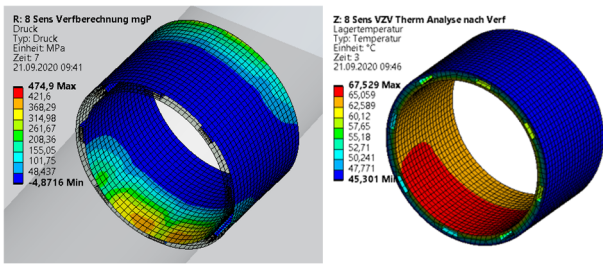


Abbildung 5: Flächenpressung im Lagergrund (links) und Temperaturverteilung im Lager (rechts)

Diese Ergebnisse halfen bei der sich anschließenden thermischen Simulation zur Bestimmung der Gleitlager-temperaturen. Die Flächenpressung definierte in der thermischen Simulation eine Wärmequelle. Außerdem wurden die Oberflächen des gesamten Modells mit einer Konvektionsrandbedingung beaufschlagt, um eine Abfuhr der generierten Wärmeleistung zu gewährleisten. Durch die Variation der Wärmeleistung der Wärmequelle konnten verschiedene thermische Belastungszustände simuliert werden.

Die Simulationsergebnisse zeigen eine Temperaturverteilung mit den beiden Maxima in den Gebieten der Flächenpressungen (Abb. 5 rechts). Die Lagermaxima und die Sensormittelwerte wurden dazu genutzt, um einen Zusammenhang zwischen messbarem Sensorwert und der realen maximalen Lagertemperatur herzustellen. Dieser Zusammenhang wurde mittels einer linearen Regression für verschiedene Verschleißgrade des Lagers in Abb. 6 dargestellt.

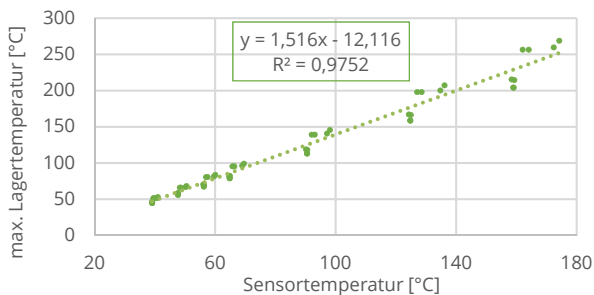


Abbildung 6: Zuordnung der messbaren Sensortemperaturen zu den realen maximalen Lagertemperaturen

Weiterhin erfolgte eine numerische Untersuchung des elektrischen Strömungsfeldes. Hierbei sollte der Stromfluss untersucht werden, der im Verschleißfall dem Anwender das Signal zum Austausch des Lagers geben soll. In Abb. 7 ist dargestellt, dass nach der Erosion der Verschleißschicht, ein elektrischer Strom zwischen mindestens zwei Sensoren über die Welle fließt. Damit konnte simulativ gezeigt werden, dass das Funktionsprinzip gewährleistet ist.

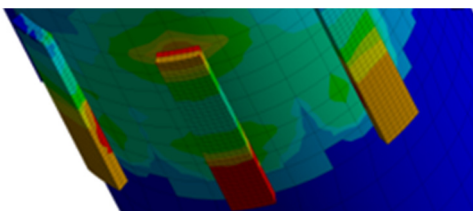


Abbildung 7: Kurzschluss mehrerer Sensoren durch Wellenkontakt im Verschleißfall

4. Experimentelle Untersuchungen

Um einen Bezug zwischen der Temperatur und dem elektrischen Messsignal der Sensorelemente herstellen zu können, bedurfte es präziser Untersuchungen des Durchgangswiderstands der Sensorelemente. Dabei wurden die Sensorelemente aus vier verschiedenen elektrisch modifizierten Materialien in einer Klimakammer untersucht. Der Messaufbau ist in Abb. 8 dargestellt.

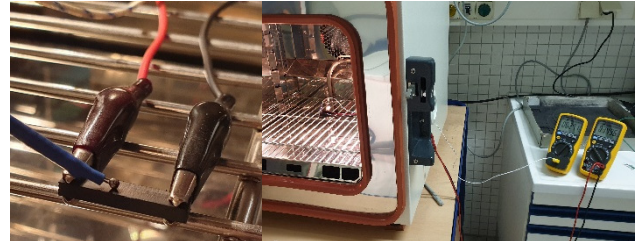


Abbildung 8: Versuchsaufbau (links: Sensorelementprobe mit Klemmen und Thermoelement; rechts: Klimakammer und Messgeräte für Temperatur und Widerstand)

Die Temperatur wurde dabei mittels eines Thermoelementes gemessen und im Betriebsbereich des Kunststoffgleitlagers variiert. Parallel dazu wurden die Durchgangswiderstände der Sensorelemente bestimmt. Die Messungen zeigten, dass bei einer Temperaturerhöhung auch der elektrische Widerstand der Sensorelemente eine Änderung aufweist. Dieser Zusammenhang ist in Abb. 9 dargestellt und liefert die Grundlage, um später in Echtzeit Aussagen zur Temperatur des Lagers treffen zu können. Besonders positiv ist dabei das Verhalten des Materials PA66 hervorzuheben, da es eine lineare Erhöhung des Widerstands bei einer Temperatur ab 50 °C aufweist. Auch der Werkstoff PBT würde sich trotz seines exponentiellen Verhaltens eignen. Das Material PEEK zeigt eine nahezu lineare negative Widerstandsänderung bei Temperaturerhöhung. Die Widerstandsänderung ist allerdings sehr gering und damit im Anwendungsfall schwer aufzulösen. Der Werkstoff PPS zeigt hingegen eine hohe Temperaturstabilität hinsichtlich seiner Widerstandsänderung und eignet sich damit nicht für den hier beschriebenen Anwendungsfall.

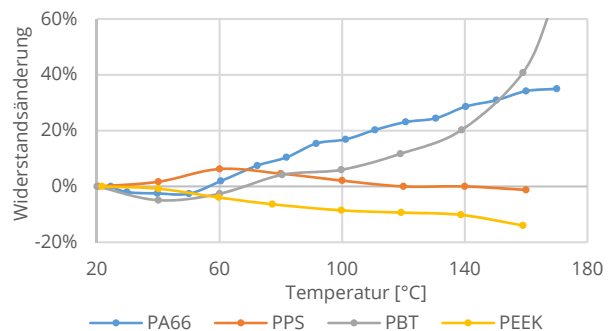


Abbildung 9: Messergebnisse des Durchgangswiderstand der Sensorelemente für verschiedene elektrisch modifizierte Kunststoffe

5. Auswerteelektronik

In einem weiteren Arbeitsschritt wurde die Auswerteelektronik an der Professur IMS ausgelegt (Abb. 10). Das kompakte Modul (Platinenmaße: 35 mm x 70 mm) beinhaltet neben den Messverstärkern auch einen Mikrocontroller, der die Messsignale in die physikalischen Größen umwandelt. Weiterhin ist ein WLAN Modul integriert, mit dem die Messdaten an eine Empfängereinheit gesendet werden können. Dies kann z. B. ein Computer oder ein Smartphone sein. Später im Entwicklungsprozess soll noch eine App entwickelt werden, um die Daten graphisch ansprechend auf einem Smartphone oder Tablet darzustellen.

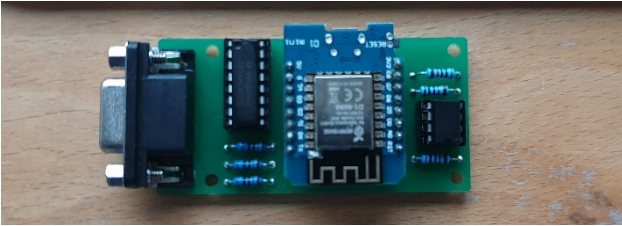


Abbildung 10: Auswerteeinheit mit D-Sub-Stecker, Multiplexer, USB-Anschluss, Mikrocontroller und WLAN-Modul

6. Herstellung eines Funktionsmodells mittels FDM-Druck

Ein erster Prototyp wurde mittels FDM-Technologie hergestellt. Dabei wurden zunächst Kunststoffe gewählt, die leicht beschaffbar und für den 3D-Druck gut geeignet sind. In der Abb. 11 ist der funktionsfähige Prototyp dargestellt. Er besteht aus zwei Kunststoffen (einem leitfähigen und einem nicht leitfähigen), die in einem Arbeitsgang aufgetragen werden. Die elektrische Funktion ist bei diesem Lager bereits voll gegeben, die tribologische Funktion aber nur mit Einschränkungen. Deshalb wird in einem weiteren Schritt ein spezieller tribologisch optimierter Kunststoff für die Gleitflächen und ein leitfähiger Kunststoff, der höhere Temperaturen verträgt, verwendet. Somit sollte dieses Lager aus dem 3D-Drucker nahezu die gleichen Eigenschaften wie ein konventionell hergestelltes Lager aufweisen. Es kann somit auch speziell auf Kundenwünsche ab einer Stückzahl von eins kostengünstig eingegangen werden.



Abbildung 11: FDM gedruckter Prototyp

7. Zusammenfassung

Das im Forschungsprojekt vorgestellte Kunststoffgleitlager mit intelligenter Verschleißsensorik bietet die Grundlage, Wartungsintervalle planbarer zu machen und Stillstandzeiten oder gar Folgeschäden zu vermeiden. Infolgedessen besitzt es ein enorm hohes Potenzial Kosten zu senken und eine nachhaltigere Instandhaltung zu gewährleisten, welche den vorzeitigen Austausch von noch gebrauchsfähigen Lagern vermeidet. Es zeichnet sich daher als ökonomischere und ökologischere Alternative zum herkömmlichen Kunststoffgleitlager aus.

Zusammenfassend konnte mithilfe simulativer und messtechnischer Untersuchungen gezeigt werden, dass das grundlegende Funktionsprinzip zur Messung des Verschleißgrades und der Temperatur gewährleistet ist. Dabei wurde das elektrische leitfähige Material PA66 als besonders geeignetes Sensorelement herausgearbeitet. Ebenso konnte eine Kennlinie zum Verhalten zwischen gemessener Sensortemperatur und realer maximaler Lagertemperatur simulativ bestimmt werden. Weiterhin wurde die Auswerteelektronik entwickelt und ein Prototyp mittels FDM-Druck hergestellt.

Die Ergebnisse können in der weiteren Entwicklung dazu verwendet werden, das Kunststoffgleitlager in einem Spritzgussverfahren serienmäßig herzustellen und damit in Massenproduktion dem Markt zugänglich zu machen. Dafür sind umfangreiche Funktionstest des Sensorgleitlagers notwendig. Weiterhin ist ein Verständnis über das temperaturbedingte Verschleißverhalten der verwendeten Kunststoffe nötig und diese sind hinsichtlich ihrer Zuverlässigkeit und Lebensdauer zu optimieren.

Danksagung

Das ZIM-Vorhaben ZF 4516707LP9 wurde im Rahmen des Zentralen Innovationsprogramms Mittelstand vom Bundesministerium für Wirtschaft und Energie aufgrund eines Beschlusses des Deutschen Bundestages gefördert. Wir danken den genannten Institutionen für die Bereitstellung der finanziellen Mittel.

Literaturverzeichnis

- [1] <https://www.ggbearings.com/de/gleitlager-produkte/solid-polymer> Stand: 18.02.2021
- [2] Valentin L. Popov: Kontaktmechanik und Reibung, Berlin 2009
- [3] VTH Verband Technischer Handel e.V.: Predicted Maintenance. In: VTH Ausblicke #1, page 6 (VTH Verband Technischer Handel e.V., 2018)
- [4] <https://cdn1.vogel.de/unsafe/540x0/smart/images.vogel.de/vogelon-line/bdb/1372900/1372979/original.jpg> Stand: 18.02.2021

Wissenschaft und Forschung zur additiven Fertigung an der Hochschule Mittweida

Hagen Bankwitz, Jörg Hübler, Claudius Petzold
Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

Additive Fertigung, auch generische Fertigung oder 3D-Druck, ist eine der Schlüsseltechnologien der nächsten Jahre bzw. Jahrzehnte. Sie findet dort Anwendung, wo konventionelle Fertigungstechnologien an ihre Grenzen stoßen, wo nur kleine Stückzahlen oder kostenintensive Materialien benötigt werden. Diesem Thema widmet sich auch die Hochschule Mittweida seit vielen Jahren in Lehre und Forschung. Einen kurzen Überblick über die erzielten Forschungsergebnisse in den Bereichen Lasermikrosintern, Selective Laser (Micro) Melting, (Hochrate) Micro Cladding, Laserauftragschweißen, FDM-Druck (leitfähiger) Strukturen auf Textilien und Entwicklung eines Hochgeschwindigkeits-Rotationsdruckverfahrens soll dieser Beitrag geben.

1. Einleitung

Additive Fertigung, auch generische Fertigung oder 3D-Druck, ist eine der Schlüsseltechnologien der nächsten Jahre bzw. Jahrzehnte. Sie findet dort Anwendung, wo konventionelle Fertigungstechnologien an ihre Grenzen stoßen, wo nur kleine Stückzahlen oder kostenintensive Materialien benötigt werden. Vorteile dieser Technologie sind u. a. der ressourcen-schonende Einsatz des Materials, eine Gestaltungsfreiheit der zu fertigenden Bauteile mit nahezu keinen Einschränkungen, ein hochautomatisiertes Verfahren und eine hohe Flexibilität der Fertigung. Bei einigen Technologien sind auch unterschiedliche Materialien in einem Fertigungsschritt herstellbar. Die Herausforderungen die es noch zu lösen gilt sind u. a. die Erhöhung der Fertigungsgeschwindigkeit, Reduktion der Materialkosten und die Eliminierung der Nacharbeit bei manchen Technologien. Die Hochschule Mittweida widmet sich bereits seit Jahren diesem Thema in Forschung und Lehre. Durch Entwicklung neuer Technologien werden neue Anwendungen entwickelt und Probleme reduziert. Einige Beispiele der erzielten Forschungsergebnisse der letzten Jahre in diesen Bereichen soll dieser Beitrag geben.

2. FDM-Druck auf Textilien

Mit dem an der Hochschule entwickelten Verfahren, können mittels eines FDM-Druckers leitfähige thermoplastische Kunststoffe auf Textilien gedruckt werden. Mit diesen gedruckten Strukturen können z. B. Sensoren und einfache Schaltkreise hergestellt werden. Aufgrund der Elastizität der Textilien, ist die Fixierung des Textils auf dem Druckbett eine der Herausforderungen des Verfahrens, welche in einem Forschungsprojekt untersucht und eine optimale Lösung hinsichtlich der Druckqualität und der Unversehrtheit der Struktur des Textils gefunden wurde.

Weiterhin wurde die Haftung zwischen Textil und gedrucktem Kunststoff optimiert. Dies erfolgt durch Parameteroptimierung und aufbringen eines Haftvermittlers, welcher als zweite Komponente im 3D-Druck in das

Verfahren integriert wurde. Die Ablösekräfte der unterschiedlichen Varianten wurden mit dem Rollschälversuch ermittelt, vgl. Abbildung 1.

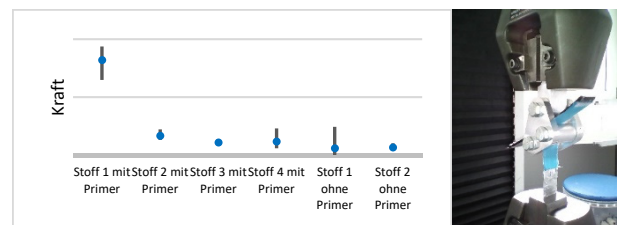


Abbildung 1: Haftkraftermittlung mittels Rollschälversuch

Für eine fehlerfreie Übertragung der Signale der gedruckten Sensoren oder Schaltkreise ist die Verbindung der im Textil befindlichen elektrischen Leiter mit den gedruckten Strukturen von großer Bedeutung. Sie beeinflusst die Störanfälligkeit, die Vibrationsfestigkeit und den Übergangswiderstand. Hierfür wurde ein optimaler Leiter im Projekt gefunden, mit dem es möglich ist Signale von gedruckten Sensoren zu Auswerteeinheiten zu übermitteln.

Im Weiteren wurden zwei Sensoren exemplarisch für das 3D-Druck-Verfahren entwickelt. Es können so z. B. einfache resistive Sensoren in einem Arbeitsgang hergestellt und mit den im Textil liegenden elektrischen Leitern kontaktiert werden. So wurde beispielsweise ein Temperatursensor entwickelt, der in einem Messbereich von bis zu 60°C funktionsfähig ist. Die Abhängigkeit des Widerstandes von der Temperatur ist in der Abbildung 2 dargestellt. Es ist ein nichtlinearer Verlauf zu sehen. Dieser kann allerdings sehr gut mit einem Polynom approximiert werden, sodass ein Bestimmtheitsmaß von über 99% erreicht werden kann.

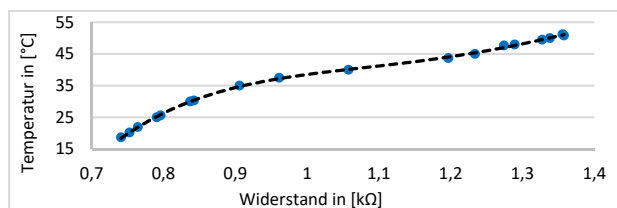


Abbildung 2: Temperatur-Widerstandsverlauf des gedruckten Sensors

Weiterhin wurde ein resistiver Berührungssensor, siehe Abbildung 3, für das Verfahren entwickelt und getestet, der als Schalter auf Textilien Anwendung finden kann.



Abbildung 3: gedruckter Berührungssensor [2]

In einem weiteren Projekt wird durch lokales aufbringen einer Schuppenstruktur auf ein Textil die Schutzwirkung gegen Stich und Schneideeinwirkung verbessert. Durch die FDM-Technologie kann diese Schutzstruktur individuell, je nach Kundenwunsch, auf das Textil aufgebracht werden. Während des Forschungsprojektes werden die Parameter der Fertigung optimiert und mögliche Materialien hinsichtlich ihrer Funktion untersucht.

Lasermikrosintern

Das Lasermikrosintern ist ein am Laserinstitut Hochschule Mittweida entwickeltes additives Fertigungsverfahren, was sich gegenüber den SLM-Verfahren durch einige Besonderheiten auszeichnet: Um sehr hohe Auflösungen von bis zu 15 µm zu erreichen, wird sowohl mit gepulster Laserstrahlung als auch mit sehr feinen Pulvern (Korngröße < 10 µm) gearbeitet, die mit einer neuartigen Ringrakel in verschiedenen Modifikationen aufgetragen werden. Dabei werden sehr kleine Schichtdicken bis unter 1 µm erreicht. Durch verschiedene Bearbeitungsstrategien kann die Materialdichte je nach Anwendungszweck im Bereich 40 % bis 99 % variiert werden, auch selektiv im Bauprozess. Es können sowohl Metalle als auch verschiedene keramische Materialien verarbeitet werden. Gemische in Form von Cermets sind ebenfalls in großer Vielfalt erzeugbar. Dadurch ergeben sich vielfältige Möglichkeiten zur Herstellung von Mikro-3D-Bauteilen, die anderweitig nicht herstellbar sind.

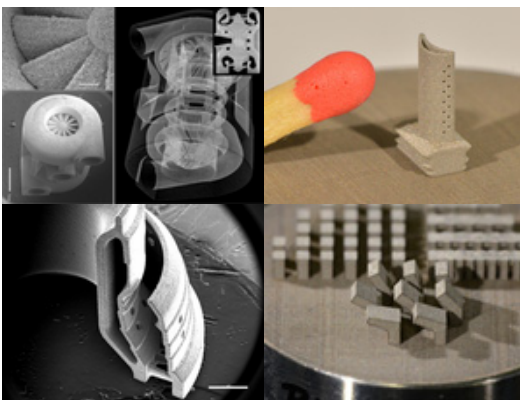


Abbildung 4: Anwendungsbeispiele Mikrosintern [3]
Selective Laser Micro Melting

Mikro-SLM ist ein pulverbettbasiertes additives Fertigungsverfahren für die Herstellung von Mikrobauteilen

aus Metall. Das Verfahren basiert auf dem Prinzip des selektiven Laserstrahlschmelzens (engl. Selective Laser Melting – SLM oder Laser Powder Bed Fusion – L-PBF), jedoch mit einigen Besonderheiten. Dazu gehören die Verwendung von sehr feinen Pulvern mit Korngrößen < 10 µm, Fokusdurchmessern < 30 µm und einer Verdichtung des Pulvers mit einer eigens entwickelten Rakeleinheit. Durch die Anpassungen lassen sich maximale Strukturauflösungen lateral bis zu 40 µm und vertikal bis zu 5 µm erzielen. Des Weiteren weisen die Bauteile eine geringe Oberflächenrauheit von $Ra \leq 3 \mu m$ auf. Durch Nachbehandlungsschritte wären sogar Ra-Werte kleiner 1 µm möglich. Die bisher getesteten Metalle sind Edelstahl, Molybdän und Wolfram. Die Aufbauraten und damit auch die Produktivität sind aufgrund der höheren Strukturauflösung im Vergleich zum kommerziellen SLM jedoch deutlich geringer. Daher ist das Verfahren derzeit eher für Mikrobauteile geeignet, was sich im kleineren Bauraum widerspiegelt. Neuste Forschungen befassen sich daher auch mit der Steigerung der Produktivität. Hierzu wird die Laserleistung von wenigen 10 W auf mehrere 100 W erhöht. Um jedoch die hohen Laserleistungen im Mikrometerbereich umsetzen zu können bedarf es auch einer Erhöhung der Scangeschwindigkeit. Dafür kommt ein am Laserinstitut Hochschule Mittweida entwickelter Polygonspiegelscanner zum Einsatz. Während herkömmliche Scansysteme den Laserstrahl nur mit wenigen Metern pro Sekunde ablenken können, schafft der Polygonspiegelscanner über 1 Kilometer pro Sekunde. Erste Untersuchungen für die Verfahren Hochrate Mikro-SLM zeigten bereits, dass die Laserleistung ohne Verluste der Detailauflösung um das 10-fache gesteigert werden konnte. Mit der Steigerung der Laserleistung erhöht sich im fast gleichen Verhältnis auch die Produktivität, was zukünftig die Herstellung von Mikrobauteilen deutlich kostengünstiger macht. Durch die Steigerung der Produktivität wäre erstmals auch die Fertigung von größeren Bauteilen mit gesteigerter Detailauflösung wirtschaftlich realisierbar.

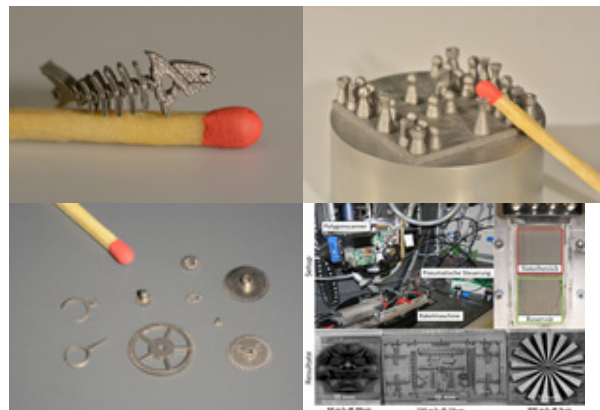


Abbildung 5: Anwendungsbeispiele Mikro-SLM [4]

(Hochrate) Micro Cladding

Micro Cladding ist eine Modifikation des klassischen Laser-Pulver-Auftragschweißens zum flexiblen Generieren

von Mikrostrukturen auf beliebig gekrümmten Oberflächen. Möglichkeiten zum Einsatz des Rapid-Micro-Verfahrens bieten sich vor allem in der Mikrosystemtechnik, z. B. zur Mikrostrukturierung und Texturierung von Oberflächen, bei der Reparatur bzw. Maßkorrektur von Mikroteilen oder beim Rapid-Microtooling.

Die Besonderheit der speziellen Verfahrensvariante Micro Cladding am Laserinstitut Hochschule Mittweida liegt in der Nutzung eines gepulsten Faserlasers in Verbindung mit einer schnellen Strahlablenkung. Das bietet den Vorteil den Energieeintrag im Mikrometerbereich gezielt zu steuern und die thermische Belastung des Substrates sowie der aufzutragenden Struktur gering zu halten. Für die Umsetzung der hohen Ablenkgeschwindigkeiten des Laserstrahls kann die Pulverdüse nicht wie im klassischen Laserpulverauftragsschweißen relativ zum Substrat mitgeführt werden, sondern muss während des Bauprozesses statisch bleiben. Mittels Pulverstrahl wird ein kontinuierliches Pulverbett auf dem Substrat erzeugt in dem sich der Laserstrahl bewegt. Durch Wechselwirkung mit der fokussierten Laserstrahlung werden die Pulverpartikel am Bearbeitungsort aufgeschmolzen und ein Materialauftrag erzielt. Eine weitere Besonderheit stellt die gezielte Mikrostrukturierung während des Aufbauprozesses dar. Damit konnte ein erheblicher Sprung in der Erhöhung der Baurate erzielt werden. Mit der Möglichkeit der Generierung von 2,5D-Strukturen bietet das Laser Micro Cladding eine Alternative zum Lasermikrosintern. Einen entscheidenden Vorteil stellt das kontinuierlich erneuerte Pulverbett dar. Im Gegensatz zum Lasermikrosintern, fällt dadurch die Belichtungspause bei jeder neuen Bauteilschicht für den Pulverbeschichtungsprozess weg und spart viel Zeit. Hochrate Micro Cladding ist eine Weiterentwicklung des Micro Cladding Verfahrens vom Laserinstitut Hochschule Mittweida. Hierfür kommen derzeit zwei verschiedene Strahlquellen und Strahlablenksysteme zum Einsatz. Zum einen werden ein kurzgepulster Hochleistungsfaserlaser und ein Hochleistungs-Galvanometer-Scanner verwendet, zum anderen ein cw-Hochleistungsfaserlaser in Verbindung mit einem akustooptischen Modulator (AOM) und einem Polygonspiegelscanner. Damit ließen sich bisher extrem hohe Bauraten von bis zu 3.900 mm³/h (Steigerung: Faktor 20 gegenüber Micro Cladding) bei Ablenkgeschwindigkeiten bis 100 m/s erzielen. Für die Generierung von großflächigen Mikrostrukturen stehen außerdem ein großer Bearbeitungsraum mit 3-Achssystem und ein Zyklonabscheider zur Pulverrückgewinnung zur Verfügung. Weiterhin wurde eine Flachstrahldüse entwickelt, die einen homogenen Pulverstrahl mit einem Querschnitt von 0,5 x 200 mm² erzeugt. Damit lassen sich ohne Achsenbewegung 200 mm lange lineare Mikrostrukturen generieren. Mit dem Verfahren sind auch Mikrostrukturen mit variablem Wandquerschnitt erzeugbar, wie zum Beispiel Mikro-Schneiden. Zudem können durch eine gezielte Umschmelzung der Oberflächen der generierten Mikrostrukturen auch Hinterschneidungen realisiert werden.

Das kann zum Beispiel genutzt werden, um eine mechanische Verankerung für schlecht haftende Schichten zu erzielen. Die derzeitigen Untersuchungen befassen sich, neben der Überführung des Verfahrens in industrielle Anwendungen, mit der Möglichkeit zur schnellen Flächenbeschichtung mit Schichtdicken im Mikrometerbereich. Erste Ergebnisse zeigen, dass homogene Schichtdicken zwischen 1 µm und 10 µm gezielt eingestellt werden können. Die maximale Flächenrate liegt derzeit bei 242 cm²/min bei einer Ablenkgeschwindigkeit des Laserstrahls von 200 m/s. Mit der neuartigen Methode können auch temperaturempfindliche Substrate beschichtet werden.

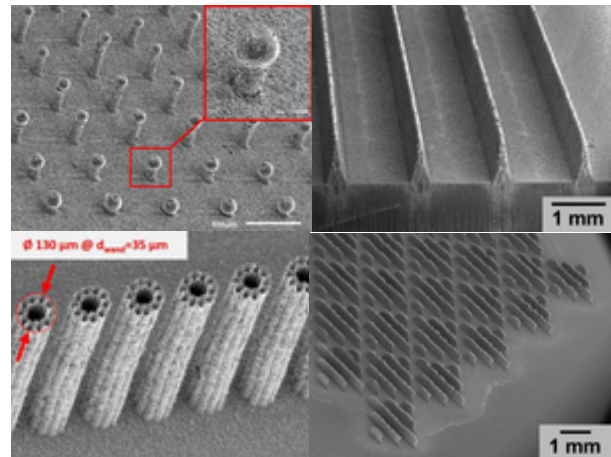


Abbildung 6: Beispiele Micro Cladding [5]

Laserauftragsschweißen

Das Laserauftragsschweißen findet schon in Reparaturprozessen, für die additive Fertigung sowie für den Verschleiß- und Korrosionsschutz vielfältige wirtschaftlich nutzbare Anwendungen. Am Laserinstitut Hochschule Mittweida erfolgt dabei die Spezialisierung im Bereich des Aufbringens hochwertiger Gleit- und Verschleißschichten auf beliebig geformten funktionellen Bauteilen. Anwendungsfelder finden sich in der Automobilindustrie zur Steigerung der Haltbarkeit von Getriebebauteilen oder auch in der Mäh- und Zerkleinerungstechnik zur Verstärkung von Industrieschneidmessern sowie der Optimierung von Brems scheiben. Spezielle Anlagentechnik ermöglicht es weiterhin, anwendungsbezogene Schichtsysteme zu generieren, um z. B. den Verschleißvorgang durch Gradientenschichtsysteme gezielt zu steuern.

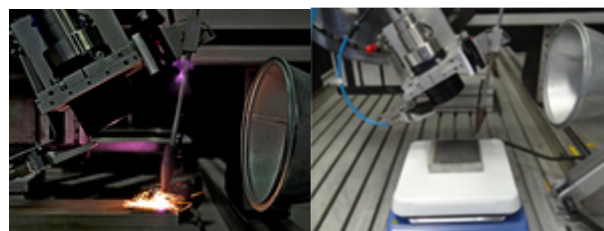


Abbildung 7: Beispiele Laserauftragsschweißen [6]

Entwicklung eines Hochgeschwindigkeits-Rotationsdruckverfahrens

Zielstellung des Projekts war es einen Hochgeschwindigkeits-3D-Drucker, basierend auf dem FDM-Verfahren, zu entwickeln. Als Haupteinsatzgebiet eines solchen Druckers ist die Orthopädietechnik angedacht. Speziell der Bereich Prothesen- und Orthesenbau. Diese orthopädischen Hilfsmittel müssen passgenau, für die individuellen Bedürfnisse des jeweiligen Patienten, gefertigt werden. Um eine optimale Hilfsstellung im Alltag und einen möglichst hohen Tragekomfort zu gewährleisten. Es handelt sich hierbei um eine Einzelfertigung bzw. Kleinstserienfertigung, wodurch sich der 3D-Druck als optimales Fertigungsverfahren darstellt.

Aktuell nimmt die Herstellung einer Orthese oder Prothese 2-3 Werkzeuge in Anspruch. Der Herstellungsprozess soll auf mindestens einen Arbeitstag reduziert werden. Von der Abnahme der Maße bis zum fertigen orthopädischen Hilfsmittel. Um dies zu realisieren wird die aufwendige Modellerzeugung mittels Gipsabdrücken und anschließender Modellierung der Gipsabdrücke durch 3D-Scan, mittels Handscanner ersetzt. Mit den Scandaten des Patienten werden dann in der Modellierungssoftware die orthopädischen Hilfsmittel passgenau erzeugt und für den 3D-Druck vorbereitet. Neben einer sauberen und schnelleren Modellerzeugung könnten so Lagerkapazitäten und Arbeitsraum freigemacht werden.

Die wichtigsten Merkmale des Druckers sind eine rotierende, beheizte Bauplattform und 4 separat verfahrbare, aber simultan arbeitende Druckköpfe. Zusätzlich werden Extruder verwendet die einen hohen Volumenstrom aufweisen. Diese Kombination gewährleistet einen hohen und konstanten Materialfluss. Durch die rotierende Bauplattform eignet sich der Drucker eher für runde, hohlzylindrische Objekte. Es wurde ein Rahmengestell konstruiert an dem die horizontalen Achsen mit den Extrudern befestigt sind. Die Extruder sind jeweils 90° versetzt und fahren auf den Mittelpunkt der Bauplattform zu. Es ist jedoch nur einem Extruder möglich über den Mittelpunkt der Bauplattform hinweg zu fahren. Die Bauplattform rotiert, angepasst an die langsamste Achse des gesamten Druckers, entsprechend schneller oder langsamer. Zu Beginn des Druckprozesses steht die Bauplattform direkt unter den Extrudern und senkt sich im Prozess um die jeweiligen Schichthöhen ab. So ist ein stabiler Fertigungsprozess möglich.

Da dieser Aufbau eines 3D-Druckers unüblich ist, musste neben der Hardware auch eine eigene Software entwickelt werden. Sogenannte Slicingprogramme erzeugen mit den Objektdaten des Druckmodelles die Bewegungsbahnen der Extruder, die Geschwindigkeiten der Achsen und diverse andere Einstellungen für den Druckprozess. Slicingprogramme für herkömmliche 3D-Drucker, die mit kartesischen Koordinatensystemen arbei-

ten, gibt es auch als Freeware zum kostenlosen Download. Der Hochgeschwindigkeits-Rotations-3D-Drucker basiert jedoch auf einem Zylinderkoordinatensystem, was eine eigene neue Slicingsoftware unumgänglich macht.

Im Laufe des Projektes wurden zwei Demonstratoren gefertigt. Der erste Demonstrator diente der Untersuchung des Druckprozesses auf eine rotierende Bauplattform. Druckparameter und Steuerungslösungen wurden hiermit erarbeitet. Der Zweite Demonstrator ist ein Komplettaufbau mit allen eingangs erwähnten Spezifikationen. Mit diesem Demonstrator wurden erste Hilfsmittel gedruckt. Bei den Hilfsmitteln handelt es sich um orthopädische Helmschalen, welche bei Patienten mit Schädeldeformationen oder im Zuge der Rehabilitation nach einer Operation zum Einsatz kommen. Die Geometrie der Helmschalen ist relativ einfach und eignet sich somit gut um den Druckprozess und die Slicingsoftware zu erproben und diese entsprechend anzupassen.

Das Projekt endete am 31.12.2020. Es wurde ein funktionsfähiger Hochgeschwindigkeits-Rotations-3D-Drucker entwickelt, welcher die orthopädischen Helmschalen innerhalb von 2 Stunden druckt. Handelsübliche 3D-Drucker benötigen für das gleiche Modell 13 Stunden.

Gefördert wurde das Projekt vom BMWi im Rahmen des „zentralen Innovationsprogramm Mittelstand“. Neben der Hochschule Mittweida waren folgende Firmen am Projekt beteiligt. Das Kunststoff-Zentrum in Leipzig gGmbH, saXcare GmbH, Jenaer Antriebstechnik GmbH, Ehnert Maschinenbau GmbH & Co. Kg, ATZ GmbH Industrieautomation und CADsys Vertriebs- und Entwicklungsgesellschaft mbH. Organisiert und koordiniert wurde das Forschungsprojekt von der FGMD GmbH.

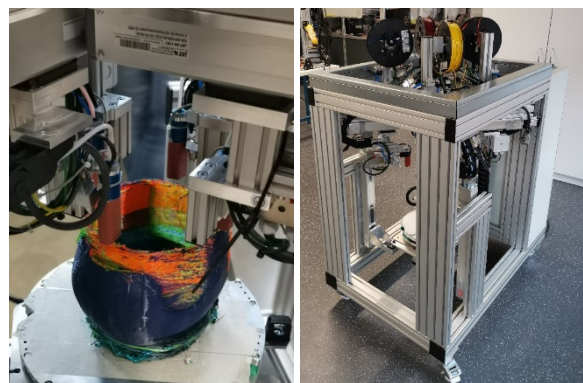


Abbildung 8: Hochgeschwindigkeits-Rotationsdrucker

Danksagung

Diese Maßnahme wird mitfinanziert durch Steuermittel auf der Grundlage des von den Abgeordneten des Sächsischen Landtags beschlossenen Haushaltes.



Europäische Union

Europa fördert Sachsen.

EFRE

Europäischer Fonds für
regionale Entwicklung



Literaturverzeichnis

- [1] H. Bankwitz: Additiv gefertigte Sensoren auf technischen In: Textilien. 7. Mitteldeutsches Forum Rapid Technologien, 26. Oktober 2020, Merseburg. (2020)
- [2] H. Bankwitz, J. Hübler: Herstellung elektrisch leitfähiger Strukturen auf technischen Textilien mittels Freiformtechnologie In: 25. Interdisziplinäre Wissenschaftliche Konferenz Mittweida. ISSN 1437-7624 (2018)
- [3] M. Erler, S. Gronau, F. Heinrich, R. Ebert, A. Streek: 3D-Druck von metallischen Mikrobauteilen mittels Lasermikrosintern, In: Innovative Verfahren der Lasermaterialbearbeitung, 12. Jenaer Lasertagung, DVS-Berichte Band 367, Düsseldorf: DVS Media, ISBN: 978-3-96144-132-7, S. 151-160 (2021)
- [4] J. Drechsel, M. Erler, R. Ebert, H. Exner: Pulver- und Prozessuntersuchung für die generative Fertigung von Edelstahlbauteilen mittels Mikro-SLM, In: Scientific Reports, 11. Mittweidaer Lasertagung, S. 125-129 (2019)
- [5] M. Erler, R. Ebert, M. Horn, S. Gronau, S. Klötzer, H. Exner: High Speed Drucken von Mikrostrukturen aus Metall mit Micro Cladding, In: 3D-Druck in der Anwendung: 3. Mitteldeutsches Forum 3D-Druck in der Anwendung, Merseburg: Hochschule Merseburg, Tagungsband, ISBN: 978-3-942703-51-2, S. 380-405 (2016)
- [6] M. Erler, S. Klötzer, M. Heinrich, L. Hartwig, R. Ebert, B. Steiger, H. Exner: Untersuchungen zum Mikropulverauftragschweißen mittels gepulster Faserlaserstrahlung, In: Tagungsband der 8. WTK TU Chemnitz, Oberflächen- und Wärmebehandlungstechnik, S. 222-228 (2009)
- [7] J. Matthes, C. Petzold: Hochgeschwindigkeits 3D-Druck zur Herstellung von Prothesenelementen, In: Scientific Reports 2018, Digitalisierung in Industrie und Gesellschaft - Arbeiten und Leben im Umfeld der 4. Industriellen Revolution (2018)

EMOTION – Electric Mobility in L-Category Vehicles for all Generations

Nicole Perterer, Martin Perterer

KTM E-Technologies, St. Leonharderstraße 10, 5081 Anif/Salzburg

Die ambitionierten Klimaziele, sowohl in Europa als auch weltweit, erfordern ein Umdecken der aktuellen Mobilitätskonzepte. Das EMOTION Projekt stellt einen möglichen Lösungsansatz vor, der die Lücke zwischen elektrischen Mopeds und Motorrädern durch klimafreundliche und kostenattraktive Fahrzeugkonzepte schließen und gleichzeitig die Akzeptanz für E-Fahrzeuge durch innovative HMI Ansätze steigern soll. Dabei kommen effiziente Plattform- und Leichtbaustrategien zum Einsatz, um sowohl die Fahrzeugstruktur als auch den Antriebsstrang gleichsam effizient wie kostengünstig zu gestalten. In einer Pilotphase werden zwei im Projekt aufgebaute Fahrzeuge im Realbetrieb getestet und sowohl die Fahrzeugplattform, als auch das neuartige HMI Konzept validiert.

1. Einleitung

Die leichte E-Mobilität in Europa unterliegt aktuell einem starken Wachstum [1, 2, 3]. Neben den Produkten aus Fernost, die in vielen Fällen den europäischen Qualitäts- und Sicherheitsstandards nicht oder nur zum Teil entsprechen (vgl. hier z.B.: Funktionale Sicherheit) finden auch immer mehr hochqualitative Premiumprodukte den Weg in den Markt, wie zum Beispiel die Vespa Elettrica von Piaggio [4]. Der hohe Preis verhindert aber aktuell eine hohe Marktdurchdringung.

Die Benutzerschnittstellen der am Markt existierenden Produkte bei leichten elektrischen zweirädrigen Fahrzeugen (E-PTW) sind zwar sehr divers, haben aber Anpassungspotential [5] für einen sicheren, nutzergerechten und effizienten Umgang im urbanen und ländlichen Bereich. Anschaffungskosten als auch die Batteriereichweite stellen hier die größte Barriere für den Umstieg der Verbraucher hin zur nachhaltigen Mobilitätsnutzung dar. Das HMI System hätte hier das Potenzial, die Nutzer dabei zu unterstützen.

Insbesondere ältere Personen finden nur schwer den Zugang zur E-Mobilität, da einerseits die Akzeptanz hier nicht gegeben und andererseits auch die Angst der Überforderung präsent ist. Um dem entgegen zu wirken, gibt es bereits einige Initiativen [6], die Ältere ermutigen sollen, verstärkt klimaverträgliche Verkehrsmittel als Alternative zum Auto zu nutzen.



Abbildung 1: Visualisierung Fahrzeugplattform EMOTION

Emotion soll hier einen zeitgemäßen Lösungsansatz für neue leichte Mobilitätskonzepte darstellen, der einerseits den europäischen Sicherheits- und Qualitätsstandards entspricht und andererseits den Kostenanspruch der Zielgruppen genügt. Eine Visualisierung dieses Ansatzes ist in Abbildung 1 dargestellt. Die Benutzerschnittstelle soll als zentrales Bindeglied dienen, um unterschiedliche Zielgruppen (Alter: hier speziell 16 bis 18 jährige sowie die Generation 50+, unterschiedliche Erfahrungsstände im Umgang mit E-Fahrzeugen bzw. generelle 2-rad Erfahrung) zum Umstieg auf leichte E-Fahrzeuge zu motivieren und sie in ihrer täglichen Nutzung beim Pendeln zu unterstützen.

2. Projektsteckbrief

Das Projekt EMOTION (**E**lectric **M**obility in L-Category Vehicles for all Generation**ions**) ist ein 3-jähriges österreichisches Leitprojekt, wird mit Mitteln aus dem Klima- und Energiefonds (KLIEN) unterstützt und im Rahmen der Zero Emission Mobility" (<https://www.klimafonds.gv.at/>) durchgeführt.

Ziel der Ausschreibung ist eine 100% Elektrifizierung von Fahrzeugen sowie die Entwicklung und Erprobung von intelligenten E-Mobilitäts- und Wasserstofflösungen bzw. deren Infrastruktur und darauf aufbauender Logistiklösungen und Mobilitätsservices.

2.1. Projektedaten

Das Projekt wurde im März 2020 gestartet und hat eine Projektlaufzeit von 3 Jahren. Insgesamt entwickeln 10 Partner aus Industrie und Wissenschaft gemeinsam eine Plattform für Elektrozweiräder für die L Kategorien, speziell für die junge (16-18 Jahre) und ältere (50+) Generation.

2.2 Konsortium

Im Konsortium sind 2 Forschungs- und Entwicklungsorganisationen, 2 Universitäten und 7 Industriepartner vertreten. Unter den Industriepartnern sind 4 Klein- und Mittelunternehmen, darunter auch 1 Start-up, enthalten.

- Austrian Institute of Technologies (Projektkoordinator)
- KTM E-Technologies GmbH
- KISKA GmbH
- WIVW Würzburger Institut für Verkehrswissenschaften GmbH
- Salzburg Research Forschungsgesellschaft m.b.H.
- FH OÖ Forschungs und Entwicklungs GmbH
- NUMERICA GmbH & Co KG
- TU Graz, Institut für Betriebsfestigkeit und Schienenfahrzeugtechnik
- Daxner & Merl GmbH
- Kobleder GmbH

3. Projektziele

Die Projektziele können in 3 Bereiche bzw. Areas eingeteilt werden, die in Abbildung 1 zu sehen sind und folgend beschrieben werden.

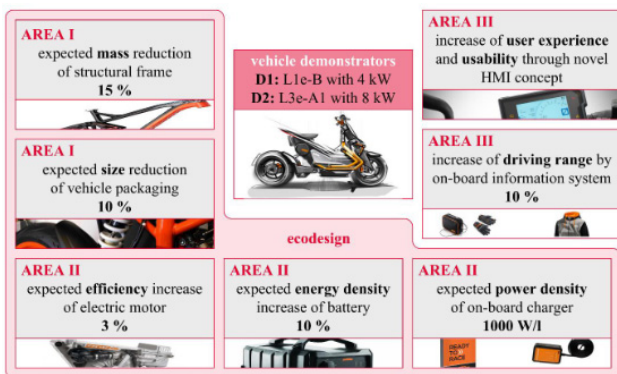


Abbildung 2: Projektziele EMOTION

3.1. Area I: Fahrzeugstruktur und Konstruktionsdesign

Der erste Bereich befasst sich mit der Fahrzeugstruktur. Durch eine Gewichtsreduktion um 15% einerseits und einer Reduktion der Fahrzeugabmaße um 10% in Relation zu Vergleichsfahrzeugen am Markt soll ein Mehrwert für das Fahrzeughandling geschaffen werden. Hier stehen die 16-18 Jährigen als auch die 50+ Generation im Vordergrund, um speziell diesen Zielgruppen die Einstiegschürde zu erleichtern.

Um diese Ziele zu erreichen wird in EMOTION der Fokus auf den kosteneffektiven Leichtbau unter Berücksichtigung einer virtuellen Entwicklungsstrategie gelegt (z.B.: Topologie-Optimierung, Funktionsintegration, Modulare und skalierbare Ansätze). Ein wesentlicher Aspekt hier ist nicht das Nutzen neuester Werkstoffe und Technologien, sondern das Zusammenbringen existierender und validierter Methoden, um die Fahrzeugplattform nicht nur leicht, sondern auch kostenattraktiv darstellen zu können.

Wie auch bei den folgenden Bereichen, Area II und Area III, spielt hier auch der Nachhaltigkeitsgedanke eine

wichtige Rolle und beeinflusst maßgeblich die Auswahl der Materialien und Prozesse. Um dieses Ziel zu erreichen werden laufend im Projekt Eco-Design Evaluierungen durchgeführt, um schon in frühen Konzeptphasen die notwendigen Entscheidungen beeinflussen zu können.

3.2 Area II: Innovativer Antriebsstrang

Der zweite Bereich umfasst den Antriebsstrang (E-Motor, Controller, Batterie, Leistungselektronik, etc.) sowie alle notwendigen elektrischen Zusatzkomponenten wie z.B. den Charger. Da diese Komponenten in der Regel den größten Kostenanteil an einem zweirädrigen E-Fahrzeug darstellen, wird versucht durch einen innovativen, und gleichsam modularen wie skalierbaren Ansatz unterschiedlichste Fahrzeugkategorien mit einem Motor-konzept abzudecken. Durch diesen Ansatz kann die Stückzahl fahrzeugplattformübergreifend erhöht und damit die anteiligen Entwicklungskosten wie auch die Einzelteilpreise der Motorkomponenten gesenkt werden. Im Detail wird untersucht, ob es möglich ist, mit einem konstanten Motordurchmesser durch Änderung der Motorlänge unterschiedliche Leistungsklassen darzustellen.

Zusätzlich soll die Größe des Antriebsstrangs durch Integration der einzelnen Antriebsstrangkomponenten reduziert und auch die Effizienz auf mindestens 90% gesteigert werden. Der E-Motor soll hier um mindestens 3% effizienter als vergleichbare Lösungen am Markt sein, bei gleichzeitiger Reduktion der Herstellungskosten um 20%.

Die Traktionsbatterie, als weitere kosten- sowie baurelevante Komponente soll in diesem Projekt ebenfalls optimiert werden. Da hier auf Komponentenseite schon eine Vielzahl an Lösungsmöglichkeiten am Markt verfügbar sind und der Energieinhalt maßgeblich von den Neuentwicklungen der Zellhersteller abhängt und somit wenig beeinflusst werden kann, soll in EMOTION der Fokus auf ein effizientes Batteriekonzept an sich gelegt werden.

Durch geeignete Materialwahl, Zellanordnungen, innovative Kühlsysteme und Batteriemanagement soll in EMOTION die Energiedichte um 10% gegenüber marktüblichen Batterielösungen gesteigert werden. Durch einen analog zum Motor hohen Modularitätsgrad sowie Skalierungseffekten sollen auch hier die Systemkosten (Herstellungskosten, Kosten der Einzelkomponenten, etc.) um mindestens 20% gesenkt werden.

Eine weitere in der Regel weniger beachtete Komponente bei E-Fahrzeugen ist das Ladegerät. Obwohl es für die Fahrzeugfunktion eine untergeordnete Rolle spielt, so beeinflusst es doch maßgeblich die Nutzung des Fahrzeuges. Der Zielkonflikt hier ist einerseits der Wunsch nach einer möglichst schnellen Aufladung des Fahrzeu-

ges, andererseits die Notwendigkeit nach kleiner Bauweise und wenig Gewicht. Beides ist mit den am Markt verfügbaren Technologien in der Regel nicht abbildbar.

Daher soll in EMOTION ein neuer Ansatz verfolgt werden. Hocheffiziente Gallium-Nitrid-Halbleiter können deutlich schneller als siliziumbasierte Feld-Effekt-Transistoren (FET) schalten und erlauben es somit die Größe der passiven Komponenten und Filter zu reduzieren und die globale Effizienz des Ladegeräts zu steigern. Somit kann hier die Energiedichte drastisch reduziert und dieser Zielkonflikt aufgelöst werden. Das Ziel hier sind 1000W/l bei gleichzeitig akzeptablen Kosten, was in etwa das Fünffache von aktuell am Markt verfügbaren Lösungen darstellt.

3.3. Area III: Nutzer zentriertes HMI Design und On-Board Informationssystem

Die Mensch-Maschine-Schnittstelle oder englisch *human machine interaction* (HMI) ist das zentrale Bindeglied zwischen dem Fahrer und dem E-Fahrzeug und beeinflusst direkt das Verhalten aber auch die Einstellung und das Wohlbefinden der Nutzer am Fahrzeug. Somit kann es auch als zentrales Instrument verwendet werden, um z.B. unerfahrene Nutzer den Einstieg in die E-Mobilität zu erleichtern als auch versierte Fahrer die Möglichkeit geben, ihren Fahrstil hin zu mehr Nachhaltigkeit zu optimieren.

In EMOTION gehen wir beim nutzerzentrierten Ansatz sogar einen Schritt weiter und wollen ein generationsübergreifendes HMI und On-Board Informationssystem konzipieren, das sich speziell den Bedürfnissen unterschiedlicher Altersgruppen anpasst. Dieses HMI System soll im Kern auf bestehende Interaktionsmöglichkeiten auf Zweirädern (z.B.: Lenkerschalter) zurückgreifen, um einerseits die Fahrer nicht zu überfordern und andererseits auch Umsteigern von Verbrennungsfahrzeugen den Einstieg zu erleichtern. Zusätzlich werden in EMOTION aber auch neue Interaktionsmöglichkeiten untersucht, welche die konventionellen Konzepte erweitern oder auch ergänzen können.

Hier soll speziell die Interaktion basierend auf textilen Sensoren für funktionalisierte Bekleidung untersucht werden, um ggf. einzelne Funktionen aus dem konventionellen Bedienkonzept herauszulösen und somit die Größe oder Komplexität dieser Bedienelemente (z.B. Lenkerschalter) zu reduzieren. Der zweite Ansatz hier ist das Darstellen von Zusatz- und Komfortfunktionen, um das Fahrerlebnis zu steigern.

Abschließend werden spezielle Coaching- und Gamingstrategien untersucht, um den Nutzer dazu zu bewegen, ökonomischer und umweltschonender zu fahren, um eine, um bis zu 10% höhere Reichweite dadurch erlangen zu können.

4. Projektablauf

Das EMOTION Projekt beinhaltet acht Arbeitspakete. Eine Übersicht ist in Abbildung 3 zu sehen.

Arbeitspaket 1 (WP1) umfasst das Projektmanagement und erstreckt sich auf die gesamte Projektlaufzeit von 3 Jahren. Neben dem technischen Projektmanagement erfolgt in diesem Arbeitspaket auch das Risikomanagement.

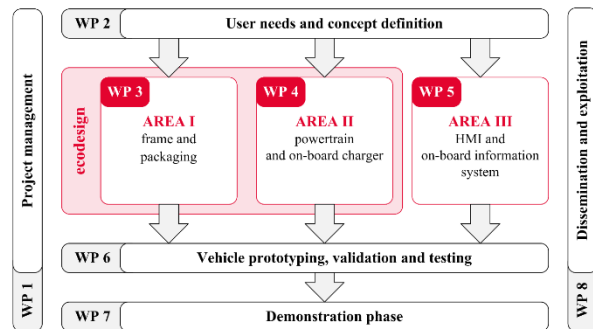


Abbildung 3: Arbeitspakete EMOTION

Das Arbeitspaket 2 (WP 2) steht am Projektanfang und beinhaltet unterschiedliche Fragestellungen. Um sicherzustellen, dass die anvisierte Fahrzeugplattform den Nutzeranforderungen der Zielgruppen genügt, wird eine europaweite Befragung zu Projektbeginn durchgeführt, welche sowohl das aktuelle Mobilitätsverhalten unterschiedlicher Generationen sammelt, als auch mögliche aktuelle Konflikte und Schwächen der am Markt erhältlichen E-Mobilitätslösungen bündelt. An der Umfrage sind sowohl aktuelle Nutzer von E-Fahrzeugen als auch potenzielle zukünftige Nutzer beteiligt. Mit Hilfe der ODI Methode (englisch für *outcome driven innovation*) werden aus den Feedbacks spezielle Kundenbedürfnisse extrahiert die als Grundlage für die weitere Entwicklung dienen und als konkrete Vorgaben in die Fahrzeugplattformspezifikation übernommen werden.

Diese erste Fahrzeugspezifikation wird mit Hilfe von Konzept- und Machbarkeitsbetrachtungen hinsichtlich der technischen Umsetzbarkeit einerseits und den damit verbundenen Kosten sowie des ökologischen Fußabdrucks bewertet.

Zusätzlich wird schon zu Projektbeginn ein Versuchsträger aufgebaut, um die im Betrieb auftretenden Lasten dieser neuen Fahrzeugplattform im Realversuch zu ermitteln und somit die Auslegungsgüte entsprechend steigern zu können. Die Ermittlung dieser Lasten erfolgt mit Hilfe einer Lastdatenermittlung.

Schließlich wird die Spezifikation der Fahrzeugplattform mit Hilfe dieser Inputs weiterentwickelt und stellt die Basis für die weiteren Arbeitspakete dar.

Die Arbeitspakete 3,4 und 5 (WP 3, WP4 und WP 5) setzen auf diesen prinzipiellen Ergebnissen auf und beinhaltet die detaillierten Untersuchungen in diesen drei Bereichen.

Arbeitspaket 3 umfasst die Entwicklung der Fahrzeugstruktur. Hier wird zunächst ein modularer Fahrzeugaufbau auf Basis der Spezifikationen entwickelt, das es ermöglicht, durch geringfügige Änderungen unterschiedliche Fahrzeugleistungsklassen und Fahrzeuggeometrien darstellen zu können. Danach wird unter Zuhilfenahme von virtuellen Methoden sowie Optimierungsalgorithmen die modulare und skalierbare Fahrzeughauptstruktur sowie auch alle notwendigen Anbau- und Verkleidungsteile entwickelt. Abschließend werden die zwei anvisierten Prototypenfahrzeuge in der L1e und L3e Klasse daraus abgeleitet.

Arbeitspaket 4 bündelt alle Tätigkeiten zur Entwicklung des Antriebsstrangs und des Ladegeräts. Hier dienen Systemsimulationen als Startpunkt, um, von den Fahrzeuganforderungen kommend, geeignete Anforderungen für die Einzelkomponenten definieren zu können. Spezieller Fokus wird hier, wie schon in Kapitel 2 beschrieben, auf eine möglichst große Anzahl an Gleichteilen für unterschiedliche Fahrzeugklassen gelegt.

Aus diesem Grund werden bei allen Komponenten geeignete Einheiten bzw. Module definiert, welche die Basis für die kleinen und leistungsschwachen Fahrzeuge bildet, dann in weiterer Folge aber für die größeren und leistungsstärkeren Fahrzeugklassen auch verwendet werden kann. Neben dem Motorkonzept mit gleichbleibendem Durchmesser und variierender Länge ist hier speziell ein Grundbatteriemodul zu nennen, das dann beliebig oft kombiniert werden kann, um die geforderten Energieinhalte darstellen zu können.

Da in Arbeitspaket 4 und Arbeitspaket 5 das Gesamtfahrzeug an sich entwickelt wird, liegt auch hier der Fokus für alle Eco-Design Betrachtungen.

In Arbeitspaket 5 sind alle Tätigkeiten zur Entwicklung des Neuartigen HMI Systems gebündelt. Hier werden Methoden zum Rapid-Prototyping von unterschiedlichen HMI and On-Board Informationssystemen verwendet, um möglichst früh schon ein Feedback von den anvisierten Nutzern als auch von HMI Experten zu bekommen und dieses wiederum in die Entwicklung einfließen lassen zu können. Da sich die Beanspruchung des Fahrers im Zweiradbereich grundlegend von dem im Vierradbereich unterscheidet, werden auch Fahrsimulatoren in der Entwicklung genutzt und adaptiert, um eben diese erhöhte Beanspruchung im Fahren nachstellen zu können.

Als Ergebnis stehen in allen 3 Arbeitspaketen, Arbeitspaket 3,4 und 5, jeweils validierte Lösungen, die anschließend in Arbeitspaket 6 (WP 6) zusammengeführt werden. In Arbeitspaket 6 erfolgt auch der Aufbau zweier Demonstratoren mit 4kW bzw. 8kW, um das Potential dieses neuartigen Fahrzeugkonzeptes im Realversuch nachweisen zu können.

Um ein hohes Maß an Sicherheit im Realversuch gewährleisten zu können, erfolgt in Arbeitspaket 6 auch

eine Validierung des Gesamtsystems an unterschiedlichen Prüfständen sowie in Realfahrten. Die gegebenenfalls notwendigen Änderungen können hier dann noch in die zwei Versuchsträger einfließen.

Die Demonstrationsphase wird in Arbeitspaket 7 (WP 7) durchgeführt. Hier werden die Demonstratoren in der Zielumgebung bewegt, um Rückschlüsse auf den Mehrwert dieser neuartigen Fahrzeugplattform hinsichtlich Fahrperformance und Akzeptanz ziehen zu können. Die Demonstrationsphase erstreckt sich auf 8 Monate. Die Fahrzeug- und Telemetriedaten werden hier gesammelt und über die Demonstrationsphase ausgewertet, um das tatsächliche Fahr- und Nutzerverhalten bewerten und mit den in WP 2 gesammelten Nutzeranforderungen korrelieren zu können. Gezielte Nutzerbefragungen werden hier ergänzend durchgeführt, um die Usability als auch die User Experience gezielt bewerten zu können.

Das Arbeitspaket 8 (WP 8) umfasst die Öffentlichkeitsarbeit und alle Publikationen im Projekt und erstreckt sich genauso wie das Arbeitspaket 1 über die gesamte Projektlaufzeit von 3 Jahren. Hier wird eine Projektwebsite zu Projektbeginn erstellt (<https://www.emotion-project.at/>) auf welcher alle relevanten Informationen zum Projekt öffentlich zugänglich sind. Mit Projektfortschritt werden hier auch die Projektergebnisse sowie Publikationen gesammelt.

Zusammenfassung und Ausblick

Die leichte E-Mobilität ist nicht nur in Europa am Vormarsch, sondern mittlerweile ein weltweiter Trend. Speziell die großen fernöstlichen Industrienationen sind hier Vorreiter, was die Marktdurchdringung anbelangt. Da sich die Anforderung des asiatischen Marktes nicht gänzlich mit denen des europäischen Marktes decken, speziell was das Thema Qualität, Sicherheit aber auch das Nutzerverhalten angeht, ist es notwendig, hier Fahrzeuge und Plattformen zu entwickeln, die eben diesen Anforderungen entsprechen.

Das EMOTION Projekt stellt hier einen möglichen Lösungsweg dar, um nicht nur eine kostengünstige, qualitativ hochwertige und funktional sichere E-Fahrzeugplattform darzustellen, sondern auch den Paradigmenwechsel hin zur E-Mobilität zu fördern und somit höhere Akzeptanz innerhalb der Bevölkerung zu schaffen.

Die Beteiligung von KTM E-Technologies als führende Firma innerhalb der Pierer Mobility AG im Bereich E-Mobilität stellt sicher, dass die Ergebnisse des Forschungsprojektes rasch Einzug in entsprechende Produkte finden.

Danksagung

Die Autoren bedanken sich bei allen Projektpartnern für die tatkräftige Unterstützung und insbesondere beim Klima- und Energiefonds Österreich für Förderung des

EMOTION Projekts im Zuge des „Zero Emission Mobility“ Programms.

Literaturverzeichnis

- [1] P. Hezel. „Elektrofahrzeuge der Klasse L in Österreich bis 2020“, in: Statista (2021), unter: <https://de.statista.com/statistik/daten/studie/827578/umfrage/elektrofahrzeuge-der-klasse-l-in-oesterreich/> (abgerufen am 14.03.2021)
- [2] ACEM. “Motorcycles registrations in the European Union grew by 8% in 2019”, in: ACEM statistical release (2020), unter: <https://www.acem.eu/motorcycles-registrations-in-the-european-union-grew-by-8-in-2019#/> (abgerufen am 14.03.2021)
- [3] SafetyNet. “Powered Two Wheelers”, in: SafeyNet (2009), unter: https://ec.europa.eu/transport/road_safety/sites/roadsafety/files/specialist/knowledge/pdf/powered_two_wheelers.pdf (abgerufen am 14.03.2021).
- [4] Piaggio & C. SpA. „Vespa Ellectrica 45 km/h“, unter: https://www.vespa.com/de_DE/vespa-modelle/vespa-elettrica/vespa-elettrica.html (abgerufen am 14.03.2021)
- [5] Eccarius, T.; Lu, C.-C., Powered two-wheelers for sustainable mobility: A review of consumer adoption of electric motorcycles. *International Journal of Sustainable Transportation* , Nr. 14, (2020), 215-231.
- [6] Verkehrsclub Deutschland VDC. „Klimavertraglich mobil 60+“, unter: <https://www.in-form.de/netzwerk/projekte/klimavertraeglich-mobil-60/> (abgerufen am 14.03.2021)

Recovery of electronic waste (lithium battery) for the industrial competitiveness of electronic recycling

Dra. María Luisa Valenzuela¹, Dr. Rene Rojas², Dra. Marjorie Segovia,
Sebastián Quintana,

¹Universidad Autónoma, Instituto de Ciencias Químicas Aplicadas, Facultad de Ingeniería.
maria.valenzuela@uautonoma.cl

²Universidad Católica, Facultad de Química y de Farmacia.

Abstract

Lithium batteries are one of the few elements that cannot pass to recovery processes in Chile since their physical-chemical characteristics do not allow it and consequently they are taken to final disposal in security deposits, devaluing valuable and limited resources such as lithium and other metals such as nickel, cobalt and manganese. This work generates the capacity to make the Chilean electronic recycling industry more competitive, developing a hydrometallurgical extractive methodologies, based on low-cost chemical reagents, and currently in use in mineral extraction processes in Chile, and whose application will allow the recovery of most of the metallic elements contained in batteries. The metallic elements are collected as soluble (stable) salts of high purity and in this way complete the sustainable virtuous circle of this type of device, so important for our society.

1. Introduction

In September 2015, China published a regulation on the recycling of lithium ion batteries (BIL). Manufacturers of automobiles, motorcycles, scooters and others will be responsible for recycling, but they expect battery manufacturers to provide the service. The EU is 100% dependent on imports of all potentially scarce battery materials. This justifies its policy, defined in 2006 and ratified in 2014, focused on the recovery of more than 45% of the batteries, processing them and recovering more than 50% of the weight of the battery materials. The EU has much stricter laws on battery recycling than the US. However, batteries cannot be landfilled in either region, and recycling companies / networks exist in North America as well. Lastly, Japan has a legislative requirement for batteries like the EU. The Japan Portable Rechargeable Battery Recycling Center provides collection of used batteries throughout the country. Companies like Sony and Nippon Recycle Center Corporation are contracted to receive these batteries for recycling.

In Chile, the entry into force of Law 20,920 on extended producer responsibility and promotion of recycling (REP), promulgated in 2016, seeks to reduce waste generation and promote reuse. Among other provisions, it establishes that producers or importers of elements that have been defined as priority products, which include this type of device, have the obligation to manage the waste originating from their products. For this work, it was defined that the study objects would correspond to cell phone batteries and notebooks manufactured from 2010 onwards. This considering that the composition of these batteries are the bases from which they are currently manufactured and from which they will be manufactured in the coming years for application in Electromobility. The work does not include LiFePO₄ (LFP) batteries because, due to their chemical composition, they

must be treated independently, so as not to complicate the matrix object of this project and reduce the recovery performance of elements of interest (Li, Co, Ni, Mn, Cu). [1-21].

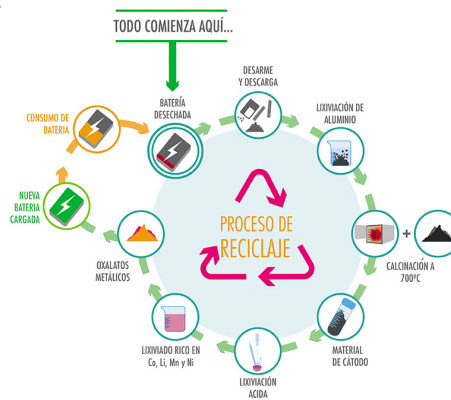


Figure 1: Scheme of the recycling process developed through this research

2. Results

2.1 Collection, classification and disassembly and neutralization of lithium rechargeable battery components.

For this work, it was defined that the study objects would correspond to cell phone batteries and notebooks manufactured from 2010 onwards. This considering that the composition of these batteries are the bases from which they are currently manufactured and from which they will be manufactured in the coming years for application in Electromobility.

Other important aspects to take into account is that their state is diverse, highlighting batteries with varying degrees of corrosion, especially in the connectors, inflated batteries, showing the degree of decomposition of

solvents and / or generation of gases within the battery. There are also bent and / or broken batteries.

Following this, the residual charge of the batteries was verified using a tester. Considering that the number of batteries of each type available was reduced (cell phones 26 and notebook 6), the charge of each of these was verified. As expected, considering that the batteries incorporate a system that limits total discharge during their useful life, all batteries have a residual charge of less than 1.3 eV and in some cases reaching 9.5 eV (computers). This information is very relevant and corroborates the danger of this type of waste, since, due to its state, heterogeneous distribution within any container, it is highly probable that a short circuit, increases in temperature or explosion may occur, which may be a possible source of ignition. during storage, transportation or processing without prior unloading. The batteries are immersed in a 1 M NaCl (common salt) saline solution for 24 hours, producing the discharge by generating an electric current between the positive and negative terminals of each submerged battery, Figure 2.

The methodology is applied in parallel to LCO and NCM batteries. This considering that they are the most commonly found batteries in the sample received.



Fig. 2. Battery discharge in 1M NaCl solution.

The disassembly of the batteries was done manually, with tools such as pliers and scissors. In this way, the container is separated from the cells (aluminum and plastic).

It was found that notebook batteries are composed of a cylindrical type battery pack that form a single unit, cell phone batteries are prismatic type batteries. Figure 3 shows the separation of the components of a cell phone battery, the procedure is similar for cell phone and notebook batteries.

It is important to remember that:

- Anode, is composed of a Cu sheet coated with graphite.
- Separator, porous plastic sheet, usually polyethylene.
- Cathode, it is composed of an aluminum foil covered by metallic oxides that contain (Co, Li for LCO and Co Ni, Mn and Li for NCM). Additionally, all cathodes contain approximately 6% graphite relative to the cathode material.

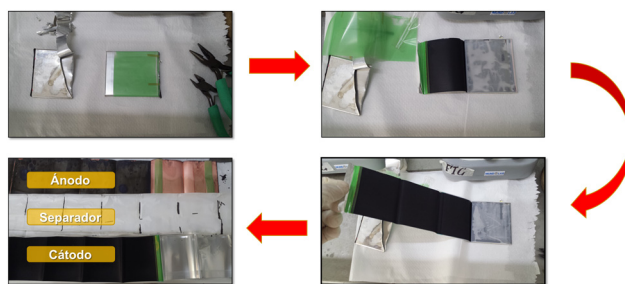


Fig. 3. Cell battery disassembly.

2.2 Separation of metallic and non-metallic components by chemical treatment.

2.2.1 Recovery of elements of interest from the anode:

The anode is immersed in water to separate the copper from the graphite. This process is very fast (5 hrs) and 100% Cu is recovered by removing it from the "reactor". The graphite is recovered by filtration, dried and stored. The water is also recovered and used for a next set of anodes, Figure 4.



Fig. 4. Recovery of copper and graphite.

2.2.2. Recovery of the separator.

The separator recovers mostly clean and does not require additional treatment, then it is stored for later application, for example as fuel in pyrolysis to generate energy.

2.2.3 Recovery of elements of interest from the cathode

The cathode is immersed in a 2M NaOH solution to dissolve the metallic aluminum and free the insoluble cathode material (LCO or NCM). The aluminum dissolution process takes only a few minutes (15 minutes). From the resulting suspension, the metal oxides are recovered by filtration, washed with distilled water and dried at 100 oC for 24 hours (100% recovery). Then it is calcined at 700 ° C for 5 hrs to eliminate organic residues (Binder, separator residues and the electrolyte). The solution rich in Aluminum is concentrated and the Al(OH)₃ precipitates (Figure 5).

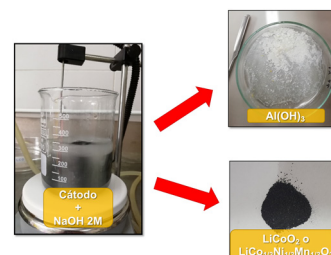


Fig. 5: Dissolution of aluminum foil and recovery of cathode material.

In this stage, the compositions of the cathode materials of different batteries were analyzed, finding, as expected, that the batteries considered are composed of LiCoO₂ or LiMn₂O₄ or LiCo_{0.2}Ni_{0.5}Mn_{0.3}O₂. Table 1 summarizes the results of the analyzes taken on the black solids, shown in Figure 5, lower right side.

A quick approximation to the composition of the anode and cathode materials is achieved by X-ray fluorescence, a technique that does not require digestion or chemical treatment of the sample, it is rather a "scanner" directly on the electrode or from the solid collected when separating aluminum and / or copper. Table 1 summarizes the results collected from this analysis on the anode and cathode electrodes.

Table 1.- Results of the composition of the cathodes and anodes of the disassembled batteries obtained from an X-ray fluorescence analysis. *

| Material | Co (%) | Ni (%) | Mn (%) | Fe (%) | Cu (%) |
|------------------------|--------|--------|--------|--------|--------|
| Cell cathode (Samsung) | 75,60 | --- | 0,12 | 0,11 | --- |
| Cell anode (Samsung) | 2,83 | --- | --- | --- | 0,77 |
| Cell cathode (Nokia) | 72,54 | 0,14 | 0,33 | 0,19 | 1,14 |
| Cell anode (Nokia) | 0,22 | 0,13 | --- | --- | 5,79 |
| Cell cathode (Alcatel) | 1,02 | --- | 56,2 | 0,22 | 0,13 |
| Cell anode (Alcatel) | --- | --- | --- | --- | 5,73 |
| Notebook Cathode (HP) | 12,49 | 30,27 | 18,88 | 0,19 | --- |
| Notebook Anode(HP) | --- | --- | --- | --- | 0,82 |

The cathodes scanned by X-ray Fluorescence, (Table 1), were "digested" and prepared for quantitative analysis of the composition by atomic absorption spectroscopy (AAEs). Results are shown in table 2.

Table 2.- Composition of uncalcined cathode material obtained through AAEs analysis.

| Material | Co (%) | Li (%) | Mn (%) | Ni (%) | Total (%) |
|------------------------|---------------|--------------|---------------|---------------|-----------|
| Cell cathode (Samsung) | 53,89 +/- 1,2 | 5,29 +/- 0,0 | 0,05 +/- 0,0 | < 0,006 | 59 |
| Cell cathode (Nokia) | 52,54 +/- 0,4 | 6,17 +/- 0,0 | 0,1 +/- 0,0 | < 0,006 | 59 |
| Cell cathode (Alcatel) | 0,87 +/- 0,0 | 3,84 +/- 0,2 | 55,71 +/- 2,2 | < 0,006 | 60 |
| Notebook Cathode (HP) | 13,24 +/- 0,0 | 6,03 +/- 0,1 | 16,48 +/- 0,0 | 29,53 +/- 0,1 | 65 |

2.3 Recovery and recovery of metallic and non-metallic components (salts) for their reuse in the production processes of new batteries.

2.3.1 Cathode material leaching

Based on the literature reports, inorganic acids such as sulfuric acid (H₂SO₄), hydrochloric (HCl) and nitric (HNO₃), are par excellence, powerful leaching agents. However, for this type of materials, organic acids such as citric acid (C₆H₈O₇), oxalic (C₂H₂O₄), formic (CH₂O₂) and acetic (C₃H₄O₂) are also good leaching agents. Considering the above, LCO and NCM type materials were subjected to leaching in the presence of H₂O₂ as a reducing agent. The results of complete screening and optimization of reaction parameters (concentration, temperature, reaction time). From this study it was determined that the best leaching agents for both materials (LCO and LCM) are H₂SO₄ and citric acid. , with almost quantitative yields under the same reaction conditions. Considering the above and that citric acid is less dangerous and consequently easy to handle, in addition to being more environmentally friendly, this acid was chosen as a leaching agent.

The following results correspond to those obtained from citric acid leaching in the presence of H₂O₂.

For the leaching process, 20 g of cathode material were taken (both in the case of LCO and NCM) and immersed in a 1M Citric Acid solution with 3% v / v of Hydrogen Peroxide, at a temperature of 90 ° C and a constant stirring of 450 rpm for 60 min. After the leaching process is finished, the solution is filtered to separate the unreacted material from the metal-rich mother liquor, which in the case of LCO corresponds to Lithium and Cobalt (Figure 6), while for NCM the metals are Lithium, Nickel, Manganese and Cobalt, the latter is characterized by being a darker solution than that shown in Figure 6.

What happens in the leaching procedure is explained chemically in the following equation, taking LCO material as an example.

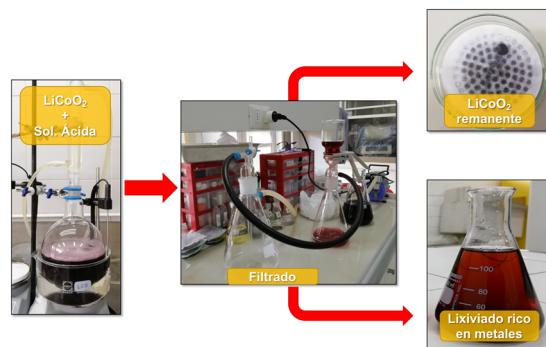
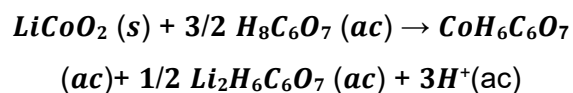


Fig. 6. Leaching of LCO type material with citric acid.

Table 3: shows the results of the characteristic analyzes for this type of solutions.

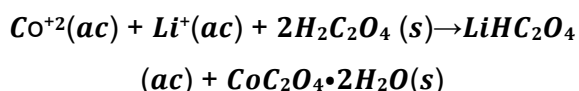
| ID | Co [%] | Li [%] | Mn [%] | Ni [%] |
|----------|--------|--------|--------|--------|
| LCO | 86,15 | 10,02 | 1,420 | 2,407 |
| LCO (pp) | 1,011 | 98,68 | 0,221 | 0,085 |
| NCM | 18,53 | 10,60 | 24,67 | 46,15 |
| NCM (pp) | 0,379 | 96,25 | 2,936 | 0,855 |

* pp indicates that the measurement is after the transition metals precipitate.

In all cases (LCO and NCM), the leaching performance determined by quantifying the solids recovered by filtration (not leached) was over 93%. Additionally, these solids can be reincorporated to new leaching processes. The beginning of the leaching generates gases and foam for which it is necessary constant stirring, during the heating of the solution and also, a reflux system or foam containment system so as not to lose solution with elements of interest. It is necessary to wash the non-leached solid, because it can be reincorporated into a new leaching batch. The filtered solution goes directly to the next stage of precipitation of the metals of interest.

2.3.2 Recovery of Cobalt and Lithium from LCO leachate by precipitation with oxalic acid.

Solid oxalic acid (7.34 g) of are added directly to the leaching solution (Figure 6) and it is maintained with constant stirring of 500 rpm at 60 ° C for 3h. The mass of oxalic acid added is calculated by the following equilibrium:



This pink solid is recovered by filtration and washing with distilled water, in a 95% yield. Under these conditions the lithium remains in solution. To recover, it is necessary to concentrate the solution, evaporating 80% of the water and then cooling to 0 ° C. In this condition, lithium oxalate (white solid) precipitates, which is recovered by filtration and washed with cold distilled water. The recovery efficiency of lithium as oxalate reaches 88%.

To recover Ni, Co, Mn and Li from materials of the NCM type, the same procedure and the same reaction conditions are followed. It is important to note that the transition metals Ni, Co and Mn are obtained as a mixture of oxalates and it is not necessary to separate them since the leaching is close to 100% and, therefore, the ion ratio is maintained in the precipitate. of oxalate, so these materials (mixture) can be directly reused for the manufacture of new batteries and it is not necessary to spend energy and time separating them to mix them again. It will only continue to adjust the metal ratio so that the final material contains the ratio of interest according to

the user's requirement. Figure 13 shows the appearance and performance of the recovery of oxalates of the metals of interest from LCO and NCM. Figure 7.

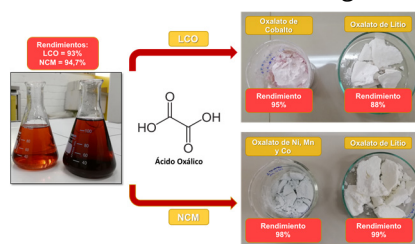


Fig. 7. Recovery of elements of interest such as oxalates.

The analysis of the metal content in the oxalates of Co, Ni, Mn, and Li are summarized in Table 4.

Table 4: Average composition of Oxalates obtained from LCO and NCM material

| ID | Co [%] | Li [%] | Mn [%] | Ni [%] |
|--------|--------|--------|--------|--------|
| Ox-LCO | 28,032 | 0,089 | 0,750 | 1,323 |
| Ox-Li | 0,042 | 7,157 | 0,005 | 0,003 |
| Ox-NCM | 6,295 | 0,211 | 8,184 | 15,574 |
| Ox-Li | 0,058 | 4,259 | 0,155 | 0,064 |

The results show the selectivity of the precipitation process under these conditions (optimal) and its high efficiency.

Conclusions

The use of an organic acid to carry out the acid leaching process was valued as the best alternative for a project of these characteristics, because it does not have strict safety regulations that are required for its application and transport, in addition to the fact that it does not generate hazardous waste after application and has a performance very similar to the inorganic acids most used in the industry.

Obtaining different products in yields over 88% was demonstrated, of economic interest from discarded batteries, products that can be sold in different markets and for different applications. One of the most relevant points that has been achieved is to take charge of each and every one of the components in an effective way, giving an exit that only requires a more industrial vision for its improvement.

Acknowledgements

FIC-2017 Proyect no BIP: 40002621-0

Reference:

- [1] The Boston Consulting Group, Batteries for Electric Cars: Challenges, opportunities and the Outlook to 2020. (2018)
- [2] K.T. Tran et al. / Hydrometallurgy 160 (2016) 106–114
- [3] Preliminary Assessment and Economic Evaluation

- of The Minera Salar Blanco Project Atacama Region, Chile, 2017
- [4] Diekmann et al, Recycling of Lithium Ion Batteries, 2018
- [5] Kesler, S.E.; Gruber., P.W.; Medina., P.A.; Keoleian., G.A.; Everson., M.P.; Wallington., T.J. Global lithium resources: Relative importance of pegmatite, brine and other deposits. *Ore Geol. Rev.* 2012, 48, 55–69
- [6] *Minerales* 2019, 9, 334; doi: 10.3390 / min9060334
- [7] Cobalt - British Geological Survey (www.MineralsUK.com) and CAILTEUX, J L H, KAMPUNZU, A B, LEROUGE, C, KAPUTO, A K and MILESI, J P. 2004. Genesis of sediment-hosted stratiform copper-cobalt deposits, central African Copperbelt. *Journal of African Earth Sciences*, Vol 42, 134–158.
- [8] <https://www.guiaminera.cl/nueve-prospectos-de-cobalto-en-chile-crean-expectativas-para-el-mayor-desarrollo-de-la-electromovilidad/>
- [9] a) Lithium-Ion Batteries: Advances and Applications. <http://dx.doi.org/10.1016/B978-0-444-59513-3.00021-2>, Linda L. Gaines, de Laboratorio Nacional Argonn. b) <https://cen.acs.org/materials/energy-storage/time-serious-recycling-lithium/97/i28>
- [10] M. Endo, C. Kim, K. Nishimura, T. Fujino, K. Miyashita, *Carbon* 38 (2000) 183.
- [11] T. Ohzuku, A. Ueda, *Solid State Ionics* 69 (1994) 201.
- [12] D. Morgan, A. Van der Ven, G. Ceder, *Electrochem. Solid State Lett.* 7 (2004) A30.
- [13] K. Ozawa, *Solid State Ionics* 69 (1994) 212.
- [14] M.M. Thackeray, A. Dekock, M.H. Rossouw, D. Liles, R. Bittihn, D. Hoge, *J. Electrochem. Soc.* 139 (1992) 363.
- [15] A.K. Padhi, K.S. Nanjundaswamy, J.B. Goodenough, *J. Electrochem. Soc.* 144 (1997) 1188.
- [16] Scrosati, B.; Hassoun, J.; Sun, Y.-K. Lithium-ion batteries. A look into the future. *Energy Environ. Sci.* 2011, 4 (9), 3287.
- [17] Gratz, E., Sa, Q., Apelian, D., Wang, Y., 2014. A closed loop process for recycling spent lithium ion batteries. *J. Power Sources* 262, 255–262.
- [18] a) Li, L., Ge, J., Wu, F., Chen, R., Chen, R., Wu, B., 2010. Recovery of cobalt and lithium from spent lithium ion batteries using organic citric acid as leachant. *J. Hazard. Mater.* 176, 288–293. b) Li, J., Wang, G., Xu, Z., 2016. An overview on the processes and technologies for recycling cathodic active materials from spent lithium-ion batteries. *Waste Manage.* 52, 221–227.
- [19] a) Lain, M.J., 2001. Recycling of lithium ion cells and batteries. *J. Power Sources* 97–98, 736–738; b) US 20070196725 A1. 2007, Method for the mixed recycling of lithium-based anode batteries and cells, Tedjar F., Foudraz, c) Sonoc, A., Jeswiet, J., Soo, V.K., 2015. Opportunities to improve recycling of automotive lithium-ion batteries. *Proc. CIRP* 29, 752–757, d) S. Ojanen et al. / *Waste Management* 76 (2018) 242–249.
- [20] Zhang, X., Cao, H., Xie, Y., Ning, P., An, H., You, H., Nawaz, F., 2015. A closed-loop process for recycling LiNi_{1/3}Co_{1/3}Mn_{1/3}O₂ from the cathode scraps of lithium-ion batteries: process optimization and kinetics analysis. *Sep. Purif. Technol.* 150, 186–195.)
- [21] J. Diekmann et al, Recycling of Lithium Ion Batteries, 2018.

Untersuchung von Kraftverlusten in der Antriebskinematik von Umformmaschinen

Armin Schleinitz, Henry Kirchner, André Sewohl, Holger Schlegel, Martin Dix
Technische Universität Chemnitz, Institut für Werkzeugmaschinen und Produktionsprozesse (IWP),
Professur Produktionssysteme und -prozesse, Reichenhainer Straße 70, D-09126 Chemnitz

Servopressen erlauben die Beeinflussung der Antriebskinematik im Werkstückkontakt. Zu berücksichtigen sind dabei die Genauigkeitseigenschaften, welche maßgebend durch Spiele der im Kraftfluss liegenden Komponenten beeinflusst werden und sich daher auf die Kinetik des Stößels auswirken. In dem statisch überbestimmten System entstehen durch Verkippung des Stößels sowie Fertigungs- und Montageungenauigkeiten der Führungen hohe Verspannungen. Folglich werden Teile der Antriebskräfte über zusätzliche Lastpfade in das Maschinengestell abgeleitet. Die Verspannungen führen zu erhöhten Kräften, welche von den Antrieben zusätzlich aufzubringen sind. In dieser Arbeit werden die Auswirkungen des Führungsspiels bei unterschiedlichen Lastfällen auf die resultierenden Leistungsverluste am Beispiel einer servoexzentrischen Presse untersucht. Die gewonnenen Erkenntnisse sollen Ansätze liefern, um einen sicheren Prozess zu gewährleisten sowie die energetischen Verluste zu beschreiben und zu minimieren.

1. Einleitung

Unter steigenden ökonomischen, technischen, ökologischen und gesellschaftlichen Anforderungen sind Unternehmen gefordert, kundengerechte und qualitativ hochwertige Leistungen zeit- und kosteneffizient zu erbringen. Um dennoch mit einer stetigen Intensivierung des Wettbewerbs bei gleichzeitig steigenden Kundenanforderungen und kürzer werdenden Entwicklungs- und Produktionszeiten Schritt zu halten, stellen Prozess- und Kostenoptimierungen einen Ausweg dar [1]. Dazu sind jedoch tiefgreifende Kenntnisse über Abläufe und Prozesse zwingend notwendig. Aus diesem Spannungsfeld zwischen steigendem gesellschaftlichen sowie politischen Druck, ökologische Aspekte in der Fertigung stärker zu berücksichtigen und erschwerten wirtschaftlichem Umfeld, resultieren eine Vielzahl von Arbeiten, welche eine ressourcenschonende Produktion zum Ziel haben [2].

Im Bereich der Umformtechnik ergeben sich derartige Ansätze durch eine Erweiterung der Umformgrenzen bzw. eine Erhöhung der Prozesssicherheit mittels neuartiger innovativer Fertigungstechnologien und angepassten Regelungskonzepten. Dabei sind auch die Genauigkeitseigenschaften von Pressen zu berücksichtigen. Diese Genauigkeitseigenschaften werden maßgeblich durch Montage- und Fertigungsungenauigkeiten der im Kraftfluss liegenden Bauteile beeinflusst. Insbesondere bei schnellen Lastrichtungswechseln im Presswerk können diese zu Verlagerungen der werkzeugtragenden Flächen führen und damit die erreichbaren Werkstückgenauigkeiten limitieren. Die Führungen in derartigen Systemen haben demnach die Aufgabe, den Stößel in gewissen Grenzen parallel zum Tisch zu halten und Verkippungen zu vermeiden bzw. zu verringern. Infolge dessen werden Teile der Antriebskräfte über zusätzliche Lastpfade als die eigentliche Wirkstelle am Werkzeug in

das Gestell abgeleitet (vgl. Abbildung 1). Außerdem treten bei der Montage der Führungen grundsätzlich gewisse Ungenauigkeiten auf, was bei einer konventionellen, indirekten Regelungsstrategie auf Motorebene zwangsläufig zu einer Verkippung des Stößels führt. Zusätzliche, unvermeidbare Montageabweichungen verstärken diesen Effekt weiter. Mit Hilfe der direkten Regelung unter Einsatz von Lineargebern am Ende der kinematischen Kette, dem Stößel selbst, können auch diese Verkippungen ausgeregelt werden. In dem statisch überbestimmten System entstehen dadurch jedoch hohe Verspannungen, welche in Folge der Antriebsregelung zu erhöhten Spindelkräften führen.

In dieser Arbeit sollen deshalb die Auswirkungen des Führungsspiels bei unterschiedlichen Lastfällen auf die

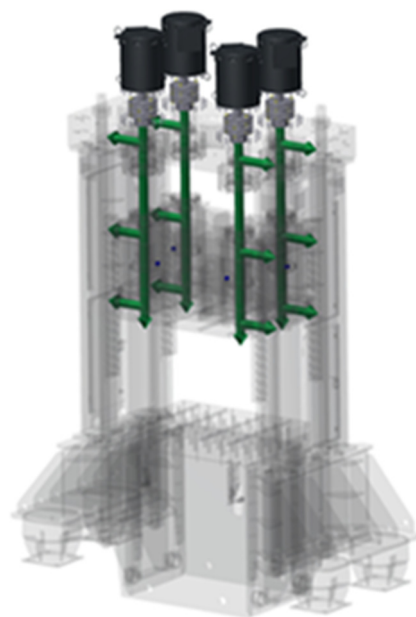


Abbildung 1: Lastpfade am Beispiel einer Servospindelmaschine

resultierenden Leistungsverluste am Beispiel einer servoexzentrischen Presse untersucht werden. Die gewonnenen Erkenntnisse sollen Ansätze liefern, um einerseits einen sicheren Prozess zu gewährleisten und andererseits die energetischen Verluste zu beschreiben und zu minimieren.

2. Stand der Technik

Das Fertigungsverfahren Umformen ordnet sich in die 2. Hauptgruppe der DIN 8580 ein. Darin werden Pressmaschinen nach ihren charakteristischen Kenngrößen in energiegebundene Maschinen (Fallhammer), kraftgebundene Maschinen (Hydraulik- und Servopressen) oder weggebundene Maschinen (Exzenter- und Kniehebelpressen) untergliedert. Auch Kombinationen dieser Funktionsprinzipien sind möglich (Servoexzenterpresse: kraft- und weggebunden) [3]. Eine für Pressen herausragende Kenngröße ist die erzielbare Genauigkeit, welche von der Steifigkeit der Maschine abhängig ist und direkte Auswirkung auf die zu fertigende Bauteilqualität besitzt [4]. Darüber hinaus zählt die Kraft bzw. die Kraftverteilung zu den wichtigsten Prozessparametern bei der Blechumformung. Dieser Kennwert eignet sich besonders für eine Regelung. Er besitzt ebenfalls Einfluss auf die Bauteilqualität und dient als Indikator für die Prozessstabilität [5]. Die Kraftverteilung hängt von verschiedenen Faktoren ab, wie Werkzeuggeometrie, Werkzeugverschleiß oder Materialparametern. Neben prozessimmanenten Faktoren im Kontaktbereich von Werkstück und Werkzeug, haben auch konstruktive und strukturelle maschinenseitige Randbedingungen sowie die Steuerung und Antriebskomponenten einen nennenswerten Einfluss, insbesondere bei mehrachsigen Maschinen wie Servospindelpressen (SSP). Aufgrund der Elastizität der Systemmaschine / des Werkzeugs können die Position der resultierenden Kräfteinleitung auf Werkstückebene und die Kraftverteilung stark variieren. Das Spiel in den Führungselementen sowie die Steifigkeit im Antriebsstrang der Presse beeinflussen die Verschiebung der werkzeugtragenden Flächen während des Prozesses erheblich. Die erreichbare Werkstückgenauigkeit ist damit limitiert.

Aus der beschriebenen Problemstellung konnten bereits Ergebnisse hinsichtlich der Kraftverteilung, speziell in der Wirkebene erarbeitet werden [6]. Hierbei stellte sich u.a. heraus, dass die Führungselemente einen wesentlichen Einfluss auf den Kraftfluss im Pressensystem haben. Es zeigte sich, dass in Abhängigkeit der Spindelkraft zwischen 4% und 25% der durch die Antriebe eingebrachten Kräfte nicht in die Wirkebene gelangten. Der dafür verwendete Versuchstand ist in seiner Nennkraft jedoch begrenzt. Um Rückschlüsse für Pressen mit größeren Nennkräften zu erhalten, erfolgten die aktuellen Versuche an einer Servoexzenterpresse am Fraunhofer IWU Chemnitz. Diese Presse zeichnet sich besonders durch die individuelle Einstellbarkeit der Rollenführungselemente am Stößel aus (vgl. Abbildung 2).

Der Fokus bei der Einrichtung eines Pressensystems und der damit verbundenen Einstellung des Führungsspiels zwischen Stößel und Führung wird in aller Regel auf die Genauigkeitsabweichung gelegt. Dies ist gleichbedeutend mit der Verhinderung/Minderung der Verkippung zwischen Stößel und Tisch. Im Allgemeinen wird deshalb ein möglichst kleines Führungsspiel empfohlen [4], [7], [8].

Darüber hinaus erzwingt eine außermittige Belastung ebenfalls eine Kippung zwischen Tisch und Stößel [9]. Diese horizontale Verlagerung zur Arbeitsrichtung (Versatz) besitzt hohe Relevanz auf die Genauigkeit einer Presse, dient als Kenngröße und ermittelt sich aus dem Abstand der Mittelsenkrechten des Stößels gegenüber der Mittelsenkrechten des Tisches und sollte in halber Distanz zwischen der Tischaufspannfläche und der Stößelfläche gemessen werden. Dieser Effekt ist unabhängig von der Gestellbauart.

Die aus der Verkippung resultierende Verlagerung und entstehende Verspannungen führen zu enormen Belastungen der Führungen [10], [11]. Um diese Verspannungen zu überwinden, erscheint es plausibel, dass auch erhöhte Antriebskräfte benötigt werden. Deshalb wird die Nennkraft F_n beispielsweise bei der Auslegung von Hochleistungsstanzautomaten stets höher gewählt, als die notwendige Stanzkraft F_s (kN) bzw. F_{dyn} (kN) [12].

3. Versuchsaufbau und -durchführung

Für die Untersuchungen wurde eine Servoexzenterpresse des IWU-Chemnitz mit einer Nennkraft von 1000kN verwendet, die einen Arbeitsweg (Hub) von 250mm besitzt. Deren grundlegender Aufbau ist in Abbildung 2 dargestellt. Der linke Antrieb ist hinsichtlich der Regelung der Maschine als Master definiert. Der rechte Antrieb folgt diesem als Slave (motorgeberseitige Lagekoppelung). Darüber hinaus besitzt die Anlage ein nicht in der Zeichnung abgebildetes direktes Wegmesssystem zur Kontrolle der Verkippung des Stößels. Außerdem sind am Übergang der beiden Pleuel zum Stößel jeweils sechs Kraftmesssensoren (Dehnmessstreifen) verbaut. Diese DMS arbeiten im Verbund und ergeben die Gesamtkraft je Pleuel.

An dieser Presse sind 16 Rollenführungselemente, insbesondere die einzelnen Führungsrollen am Stößel, individuell einstellbar. Diese am Stößel montierten Führungselemente besitzen in X-Z-Ebene drei und in Y-Z-Ebene zwei Führungsrollen, welche durch exzentrische Führungsrollenaufnahmen individuell eingestellt werden können (vgl. Abbildung 2 mitte).

Die Durchführung der Messungen erfolgte an drei unterschiedlichen Positionen P1, P2 und P3 (vgl. Abbildung 2 rechts). Hierzu wurden jeweils mehrere Durchläufe als Einzelhub realisiert. Außerdem wurde die Stößelgeschwindigkeit in vier Stufen (4, 8, 16 und 32 Hübe/min) gestaffelt.

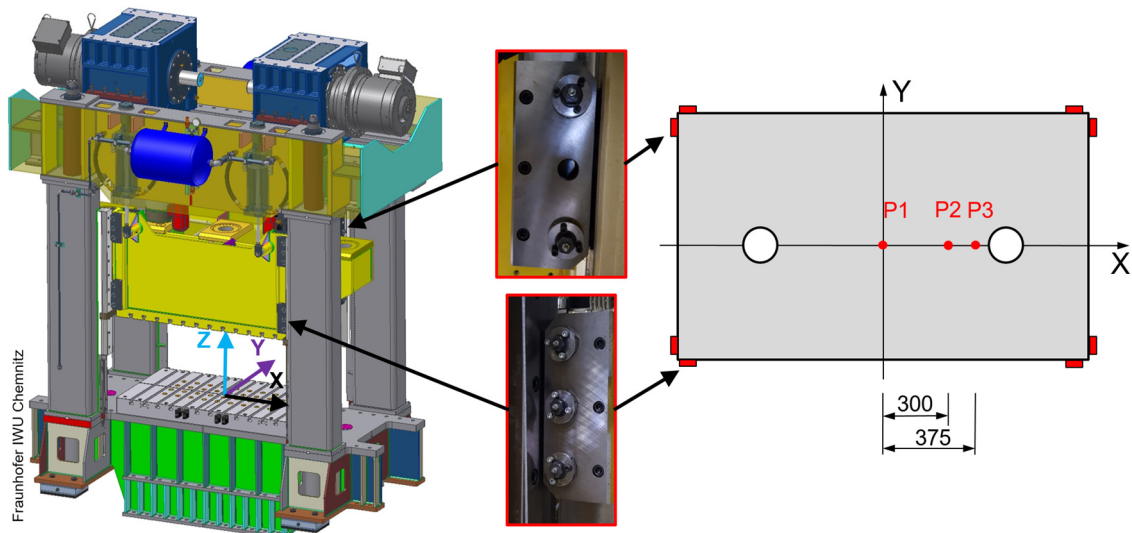


Abbildung 2: links: CAD-Modell einer Servoexzenterpresse; mitte: Ausführung der Führungselemente am Stößel; rechts: Position der Belastungsfälle in der Stößeldraufsicht

Um nun die Auswirkungen des Führungsspiels auf die Kraftverluste zu untersuchen, wurden zunächst die Führungsrollen mit einem Anzugsmoment von 30Nm an die Führungsschienen angestellt und somit **spielfrei** vorgespannt. Im zweiten Versuchsdurchlauf erfolgte die gleiche Vorgehensweise, jedoch mit dem Unterschied, dass zwischen den Führungsrollen und den Führungsschienen ein Abstand von 0,1mm eingestellt wurde. Dies wurde unter Zuhilfenahme entsprechender Stahllehren und gleichem Spannmoment arretiert. Die Ergebnisse dieses zweiten Versuchsdurchlaufes sind im Folgenden mit „Zehntel“ bezeichnet.

Um die Kraftverluste in den Führungselementen untersuchen zu können, wurden die zeitgleich aufgenommenen Ergebnisse aus zwei Messsystemen miteinander verglichen. Die Kraftmessung innerhalb des Kraftflusses der Presse erfolgte am Übergang vom Pleuel zum Stößel mit den beschriebenen 12 DMS und wurde von dem Engineeringtool SimotionScout® aufgezeichnet. Die Messung in der Werkzeugebene erfolgte mit einer DMS-Kraftmessdose (KMD). Diese besitzt eine Nennkraft von

2000kN, was der doppelten Pressennennkraft entspricht. Um die notwendige Gegenkraft erzeugen zu können und Beschädigung der Presse entgegenzuwirken wurden elastische Kunststoffplatten zwischen Stößel und KMD positioniert.

Die Daten wurden für die Auswertung zeitlich synchronisiert und verglichen. Außerdem wurde der Einfluss des Rauschens durch mehrere Messungen begrenzt.

4. Ergebnisse

In Tabelle 1 sind die Ergebnisse der Messungen aufgeführt. Hierbei handelt es sich um die gemittelten, maximalen Kraftwerte in kN, welche unter den entsprechenden Einstellungen gemessen wurden.

Sie zeigen zum einen, dass ein signifikanter Teil der aufgebrauchten Antriebskräfte über die Führungen des Stößels in das Maschinengestell abgeleitet werden. Dies wird ersichtlich, da die Werte am Stößel (Scout) über denen in der Werkzeugebene (KMD) liegen, unabhängig von der Führungseinstellung.

Tabelle 1: Messergebnisse der maximalen und gemittelten Kraftwerte in kN

| Stößelgeschwindigkeit | | mittig | | r_300 | | r_375 | |
|-----------------------|-------|-----------|---------|-----------|---------|-----------|---------|
| | | spielfrei | Zehntel | spielfrei | Zehntel | spielfrei | Zehntel |
| 4 | KMD | 163,52 | 164,25 | 164,11 | 148,13 | 173,06 | 160,32 |
| | Scout | 167,71 | 166,76 | 167,70 | 149,97 | 177,75 | 162,87 |
| 8 | KMD | 165,06 | 163,22 | 165,02 | 151,57 | 169,99 | 161,84 |
| | Scout | 170,52 | 168,93 | 170,60 | 155,41 | 176,98 | 166,16 |
| 16 | KMD | 166,08 | 164,66 | 165,90 | 155,79 | 173,79 | 163,54 |
| | Scout | 172,55 | 169,14 | 172,40 | 160,34 | 181,51 | 167,64 |
| 32 | KMD | 168,88 | 167,85 | 168,88 | 163,14 | 178,35 | 171,55 |
| | Scout | 170,36 | 166,59 | 173,72 | 163,18 | 181,48 | 170,70 |

Bei der **spielfreien Einstellung** liegen die Kraft-verluste zwischen 2% bis 4,5%. Insbesondere bei einer hohen Hubzahl (16 Hübe/min) steigen die Kraftverluste deutlich an. Dass diese Unterschiede bei 32 Hüben/min geringer sind, ist mit der Mess-rate der Kraftmessdose zu erklären, welche bei 200Hz liegt. Dadurch ist es möglich, dass die Spitzenwerte nicht ausreichend genau detektiert wurden. Außerdem zeigt sich, dass im exzentrischen Lastfall die Kraftverluste höher sind als im zentrischen. Dies deutet darauf hin, dass mit steigender exzentrischer Belastung die Kraftverluste steigen.

Bei der Versuchsdurchführung mit dem durch die Lehren eingestellten Führungsspiel von 0,1mm („Zehntel“) liegen die Stößelkräfte ebenfalls höher, als die der Kraftmessdose. Jedoch nur zwischen 1,5% und 3%. Auch hier lassen die Daten der Messung mit der Geschwindigkeit von 32 Hüben/min auf eine unzureichende Messfrequenz der Kraftmessdose schließen. Es ist überdies ersichtlich, dass auch bei exzentrischen Lastfällen die Kraftverluste in der gleichen Größenordnung im Vergleich zum zentrischen Lastfall bleiben.

Nun werden die Kraftverluste zwischen „spielfrei“ und „Scout“ mit „Zehntel“ und „KMD“ verglichen. Im zentrischen Lastfall weisen die Messdaten eine Abweichung zwischen 0,5% und 2,2% auf, wobei hier die höchsten Abweichungen bei hohen Hubzahlen auftreten. Diese Kraftverluste sind entsprechend als relativ gering zu bewerten. In den untersuchten exzentrischen Lastfällen werden die Kraftdifferenzen deutlich größer. Sie liegen in diesen Szenarien zwischen 3,5% und 10,6%. Daraus kann geschlossen werden, dass mehr als 10% der durch die Antriebe aufgewendeten Kräfte die Wirkebene nicht erreichen, sofern von einer spielfreien Führungseinstellung ausgegangen wird.

5. Zusammenfassung

Aus den Ergebnissen wird ersichtlich, dass die Erkenntnisse aus vorangegangenen Untersuchungen (vgl. [6]) auf Produktionssysteme übertragbar sind. Die Ergebnisse zeigen, dass Spielfreiheit oder gar ein Verspannen der Stößelführung zu vermeidbaren Kraftverlusten führt. In letzter Konsequenz folgern sich vermeidbare Betriebskosten durch den stärkeren Verschleiß der höher beanspruchten unteren Stößelführungselemente und im erhöhten energetischen Einsatz. Dies kommt insbesondere bei exzentrische Lastfällen zum Tragen.

Hieraus lassen sich folgenden Empfehlungen für einen wirtschaftlichen und qualifizierten Prozess ableiten:

- kleines Spiel zwischen den Führungselementen am Stößel und den Führungsleisten ist empfehlenswert und
- ein zentrischer Lastfall innerhalb der Werkzeug-ebene ist anzustreben.

Um die vorgestellten Ergebnisse weiter zu verfeinern, sollten zukünftig die Motorströme in weitere Analysen einbezogen werden.

Danksagung

Diese Forschung wird gefördert aus Landesmitteln des Freistaats Sachsen durch die „Sächsische Aufbaubank“ (SAB) sowie durch die Europäischen Union aus Mitteln des Europäischen Fonds für regionale Entwicklung (EFRE) als Bestandteil des Projektes „Sensorplatte zur großflächigen Erfassung und Verwertung der Kraftverteilung in Pressen“.



European Union



Quellen

- [1] B. Schächli, M. M. Andreasen, M. Kirchgeorg und F.-J. Radermacher, Handbuch Produktentwicklung, München Wien, Carl Hanser Verlag (2005).
- [2] R. Neugebauer ed. all, Ressourceneffizienz: Schlüsseltechnologien für Wirtschaft & Gesellschaft, Springer-Verlag (2016).
- [3] A. Hirsch, Werkzeugmaschinen, 2. Auflage, S. 347ff, Springer, Wiesbaden (2012).
- [4] E. Doege und B.-A. Behrens, Handbuch Umformtechnik: Grundlagen, Technologien, Maschinen, Springer Berlin Heidelberg (2007).
- [5] S. Rittmeier, Systemunterstützte Umformung, Dissertation, Technische Universität Dresden (2007).
- [6] A. Schleinitz, A. Sewohl, H. Schlegel and M. Putz, "Detection of the Force Distribution Close to the Effective Site in Forming Machines for a Force Control," 7th International Conference on Control, S. 216-220, Mechatronics and Automation (ICMA), Delft, Netherlands (2019).
- [7] H. Tschätsch, Praxis Der Umformtechnik: Arbeitsverfahren, Maschinen, Werkzeuge, 10. überarb. und erw. Aufl. Wiesbaden, Vieweg + Teubner (2010).
- [8] Ch. Brecher, Werkzeugmaschinen Fertigungssysteme 1: Maschinenarten Und Anwendungsbereiche. 9. Auflage, Berlin, Springer Vieweg (2019).
- [9] B. Bender und D. Göhlich. Dubbel Taschenbuch für den Maschinenbau 2: Anwendungen, Springer Berlin Heidelberg (2020).
- [10] S. Teichrib, V. Salfeld, R. Krimm, B.-A. Behrens, Untersuchung der Stößelführungsbelastung an Pressen, Verlag Meisenbach GmbH, Bamberg (2013).
- [11] M. Ahrens, Ermittlung der Belastungs-kräfte zur Auslegung von Stößelführungen mit Profilschienen, Dissertation, Universität Hannover, IFUM (2009).
- [12] M. Kolbe, Spanlose Fertigung Stanzen: Stanzteile - Feinstanzteile - Hochleistungsstanzenwerkzeuge - Hochleistungs- und Feinstanzpressen, Wiesbaden, Springer Vieweg (2018).

Adaptive Produktion von Holzbauteilen mittels Part-to-Tool

Marten Stepputat, Florian Beuß, Uwe Pflötscher, Jan Sender, Wilko Flügge
Fraunhofer-Institut für Großstrukturen in der Produktionstechnik IGP, Albert-Einstein-Str. 30, 18059 Rostock

Der Einsatz von Automatisierungstechnik zur Produktion von Einzelstücken ist in kleinen und mittleren Unternehmen derzeit häufig nicht wirtschaftlich. Gründe hierfür sind einerseits die hohen Investitionskosten und zusätzlich die notwendigen Umrüst- sowie Adaptionsvorgänge, welche in der Regel manuell vorgenommen werden müssen. Dagegen steht der aktuelle Fachkräftemangel, welcher eine effiziente Nutzung der vorhandenen Personalressourcen erforderlich macht. Durch einen Part-to-Tool-Ansatz wird im vorliegenden Beitrag ein Ansatz aufgezeigt, um die maschinelle Produktion von Einzelteilen für kleine Unternehmen im Holzbau praktikabel zu machen. Dabei wird ein 6-Achs-Knickarm-Roboter so ausgerüstet, dass dieser alle Produktionsschritte ausführen und gleichzeitig das Bauteilhandling zum Be- und Entladen übernehmen kann. Durch die Verbindung zu einer übergeordneten Steuerungsebene und der Anbindung an eine Mensch-Maschine-Schnittstelle können ungelernete Mitarbeiter Einzelteile konfigurieren und automatisiert fertigen lassen.

The use of manufacturing machines for the production of items in one-off is currently often not economical for small and medium enterprises. The reasons for this are, on the one hand, the high investment costs and, in addition, the necessary retooling and adaptation processes, which usually have to be carried out manually. On the other hand, there is the current shortage of skilled workers, which makes efficient use of the available human resources necessary. Using an innovative part-to-tool concept this article develops an approach to make the machine production of individual parts practicable for small companies in wood construction. A 6-axis articulated arm robot is equipped so that it can carry out all production steps and at the same time take over the component handling for loading and unloading. Through the connection to a higher-level control level and the connection to a user interface, unskilled workers can configure individual parts and have them manufactured automatically.

1. Einleitung

In Deutschland sind Handwerksbetriebe besonders durch den fortschreitenden Fachkräftemangel betroffen. [1,2] Dagegen steht der hohe Personalbedarf durch die Spezialisierung auf Kleinserien oder Sonderanfertigungen. Insbesondere in kleinen Betrieben, wie Tischlereien müssen somit vorhandene Ressourcen so effektiv wie möglich eingesetzt werden.

Ein vielversprechender Ansatz ist hierbei der Einsatz von CNC-Maschinen zur automatisierten Fertigung von Frästeilen. Dabei werden hauptsächlich 3-Achs-Maschinen eingesetzt. Hierbei überwiegen die geringeren Investitionskosten sowie der geringere Platz- und Schulungsbedarf ggü. 5-Achs-Fräszentren und deren signifikant höheren Fertigungsoptionen. Somit sind für komplexe dreidimensionale Bauteile nach wie vor Umspannvorgänge notwendig, welche derzeit überwiegend manuell durchgeführt werden und somit personelle Arbeitszeit binden. Zusätzlich dazu ist, bedingt durch die Auftragsstruktur, der Anteil der Programmier- und Einrichtungszeit von der gesamten Fertigungszeit sehr hoch. Hierdurch wird der Einsatz von 5-Achs-Fräsmaschinen für kleine Tischlereien oft unwirtschaftlich.

Für die komplexe Freiform-Fertigung geht der aktuelle Trend hin zum Einsatz von Industrierobotern. Insbesondere für sehr große Strukturen stellt das Verfahren eine

alternative zu konventionellen manuellen Prozessschritten dar und steht aktuell im Fokus von Forschung und Entwicklung. [3-9]

Dagegen steht bei kleinen Freiformelementen, wie bspw. Dekorationsobjekten oder auch Komponenten aus dem Möbelbau, der Fertigungsraum der genutzten Schwerlast-Industrieroboter oft in keinem Verhältnis zu tatsächlich notwendigen Abmessungen. Zusätzlich dazu wird eine große Stell- und Bearbeitungsfläche notwendig, welche die Praxistauglichkeit für kleinere Werkstätten mit begrenztem Platzangebot stark verringert. Daher sind Ansätze zu finden, welche die geringen Prozesskräfte bei der Holzbearbeitung ausnutzen und auch kleine Industrieroboter mit entsprechend geringer Traglast befähigen Fräsaufgaben durchzuführen. [7,8] Unter der zusätzlichen Verwendung des Part-to-Tool-Ansatzes können insbesondere die genannten Komponenten somit ohne zeitaufwändige Umspannvorgänge gefertigt werden.

1.1. Anwendungsfall

Wie dargelegt können für die Fertigung von kleinen Holzbauteilen auch kompakte Industrieroboter mit geringer Traglast eingesetzt werden. Daher wird in diesem Beitrag ein Ansatz vorgestellt, in dem das dreidimensionale Freiformfräsen von Vollholzsouvenirs durch einen Roboter erfolgen kann. Hierbei werden die Objekte auf einem

Sockel positioniert, welcher als zusätzliche Option mit einem Schriftzug individualisiert werden kann. Sowohl das Objekt selbst, als auch die Holzsorte können hierbei aus vorgegebenen Optionen gewählt werden.

Die dabei abzubildenden Prozessschritte entsprechen denen der klassischen Holzbearbeitung und umfassen das Schruppen, zwei Schlichtprozesse und das Polieren bzw. Schleifen. Zusätzlich dazu wird noch der Gravurprozess integriert. Zwischen jedem Prozessschritt muss dabei ein Werkzeugwechsel erfolgen, wodurch entweder ein kostenintensives Wechselsystem oder manuelle Vorgänge notwendig werden. Notwendige Zu- bzw. Entladevorgänge bedürfen einer zusätzlichen Aktorik oder manueller Interaktion. Ebenfalls kann es auch bei der Bearbeitung mit Robotern notwendig werden, das Fertigungsteil in seiner Orientierung mindestens um eine Achse zu bewegen, da sonst einige Stellen nicht erreicht werden können. Dies erfordert ebenfalls eine zusätzliche Aktorik oder händische Umspannvorgänge.

Eine weitere Herausforderung stellt die Erstellung des Fräsprogrammes dar. Die wählbaren Individualisierungsoptionen, wie die Holzsorten, das Design oder der Schriftzug auf dem Sockel, machen eine programmatische Neuerstellung mit hohem personellen Aufwand erforderlich.

1.2. Ziele

In diesem Beitrag wird ein Ansatz entwickelt, der die Anforderungen von kleinen Tischlereien bei der automatisierten Fertigung von Holzbauteilen berücksichtigt. Dabei soll der Bedien- und Einrichtungsaufwand durch die Verwendung eines Industrieroboters und dem Part-to-Tool-Ansatz minimiert werden. Ein Werkstückhandling- sowie Spindelsystem soll die Durchführung aller notwendigen Bearbeitungsschritte automatisieren und ohne manuelle Umspannvorgänge ermöglichen. Die Minimierung des Platzbedarfes erfolgt durch die Integration einer Materialzu- bzw. abführung in das Gesamtlayout.

Durch eine zusätzliche Mensch-Maschine-Schnittstelle zur Konfiguration der Frästeile sowie zur Interaktion mit der Fertigungszelle sollen die Mitarbeiter befähigt werden ohne Programmieraufwand Fertigungsaufträge zu erstellen und abarbeiten zu lassen. Durch die durchgehende Automatisierung des Prozesses werden die notwendigen personellen Ressourcen auf ein Minimum reduziert und können für wertschöpfende Tätigkeiten genutzt werden.

2. Hardwarekonzept

Im Vergleich zu Metall unterscheidet sich die Bearbeitung von Holz in zwei signifikanten Punkten. Zum einen ist der Werkstoff inhomogen und anisotrop, was zu einer geringeren erreichbaren Bearbeitungsgenauigkeit führt. [10] Bedingt durch die Inhomogenität können die Prozesskräfte nicht analytisch berechnet, sondern müssen

entweder abgeschätzt oder numerisch für die tatsächliche Anwendung ermittelt werden. [8] Zum anderen ist Holz im Vergleich zu Metall ein weiches Material. Dies führt zu wesentlich geringeren resultierenden Momenten in der Aktorik und ermöglicht die Verwendung von weicheren Mechaniken.

Aufgrund dieser Eigenschaften können hier die notwendigen Bearbeitungsschritte als auch die notwendigen Freiheitsgrade mit einem Roboter mit einer maximalen Traglast von 10kg erfüllt werden. [7]

2.1. Fräskonzept

Derzeitig am meisten genutzt werden Fräsvorgänge, bei denen das Werkzeug über ein fest eingespanntes Werkstück geführt wird. Herkömmliche CNC-Bearbeitungsmaschinen als auch roboterunterstützte Anwendungen arbeiten nach diesem Prinzip. Vorteile sind zum einen eine hohe Steifigkeit und Kompaktheit der Aktorik bei den CNC-Maschinen und der damit verbundenen hohen Genauigkeit und zum anderen die Möglichkeit mit Robotern große und schwere Bauteile zu bearbeiten. [11]

Da jedoch im Bereich der Holzbearbeitung beide Anforderungen nicht in allen Anwendungsfällen gegeben sind, sind verschiedene Bearbeitungszentren zu finden, welche nach dem Part-to-Tool-Ansatz arbeiten. Dabei wird das Werkstück relativ zu einem fest positionierten Werkzeug bewegt. Insbesondere hoch spezialisierte Systeme, wie bspw. zur Bohr- und Kantenbearbeitung sind so ausgelegt. Weiterhin werden auch Schleif- und Polierprozesse nach dem Ansatz erstellt, wie die Intec ABB-IRB2400 Robotic Finishing Cell [10].

Die Vorteile des Part-to-Tool-Ansatzes liegen, wie bereits dargestellt, auch in der Möglichkeit hoch spezialisierte Maschinen für präzise Operationen zu entwickeln und beispielsweise auch die automatisierte Bohrbearbeitung großer Bauteile zu ermöglichen. Weiterhin kann insbesondere unter Verwendung von Industrierobotern eine Anwendung konzipiert werden, in welcher der Roboter das Werkstückhandling im gesamten Prozess übernimmt. Gerade die Aufgaben der Materialzu- bzw. abführung des Prozesses, normalerweise durch zusätzliche Systeme bewerkstelligt, kann somit von der zentralen Aktorik im System ausgeführt werden und somit kostenintensive Erweiterungen obsolet machen.

Im vorliegenden Ansatz wird ein Roboter mit 10 kg Traglast über eine Greiferapplikation befähigt, das Werkstückhandling im gesamten Prozess zu übernehmen. Dies ermöglicht eine vollständige Automatisierung des gesamten Fertigungsprozesses von der Materialzuführung bis zur Ausgabe des Fertigteils.

Bedingt durch implementierte Greiferlösung am Roboter müssen weitere Prozessabschnitte anderweitig abgebildet werden. Dabei kann entweder eine spezielle Spindel mit einem Werkzeug-Wechselsystem kombiniert oder für jeden Schritt eine Spindel mit festem Werkzeug eingesetzt werden. Im vorliegenden Ansatz wird dem

Umstand Rechnung getragen, dass die notwendige Anzahl an unterschiedlichen Bearbeitungswerkzeugen sehr begrenzt ist. Dies macht den Einsatz zusätzlicher Aktorik zum Wechseln sowie die dazugehörigen preisintensiveren Bearbeitungswerkzeuge unwirtschaftlich. Daher wird für jeden Bearbeitungsschritt eine feststehende Hochfrequenz-Spindel verwendet. Durch die vorhandenen Freiheitsgrade des Roboters ist ebenfalls die in Abbildung 1 dargestellte kompakte Bauweise möglich.

Dabei sind die Spindeln für die Bearbeitungsschritte mit den höchsten zu erwartenden Prozesskräften, Schruppen (S1) und Schlichten (S2+S3), möglichst nah an der Roboterbasis platziert. Dies ermöglicht eine kompakte Pose und folglich eine hohe Gesamtsteifigkeit des Systems. Gleichzeitig reduziert diese Anordnung die resultierenden Momente in den einzelnen Gelenken. Die Spindeln für das Schleifen/Polieren (P) als auch für die Gravurarbeiten (G) sind weiter entfernt positioniert, da hier aufgrund der wesentlich geringeren Prozesskräfte eine gestrecktere Pose möglich ist.

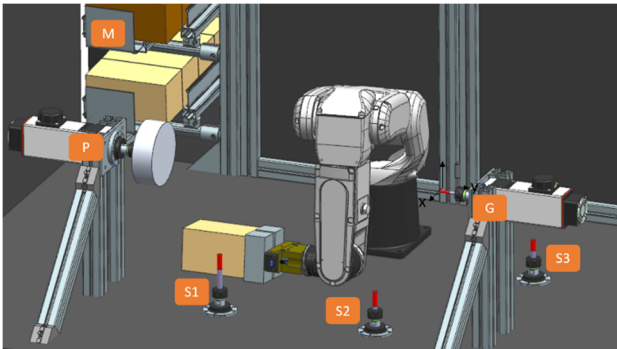


Abbildung 1: Fräskonzept - (S1) Schruppen, (S2) Schlichten, (S3) Schlichten, (G) Gravurfräsen, (P) Schleifen/Polieren, (M) Materialzuführung

2.3. Gesamtlayout

Damit die Materialzuführung bzw. die Ausgabe des Fertigteils ebenfalls vom Roboter übernommen wird, werden beide Komponenten direkt in die Fräszelle integriert.

Durch die möglichen Individualisierungsoptionen sind im Konzept insgesamt drei verschiedene Rohmaterialien zur Verfügung zu stellen. Durch die Verwendung von Schwerkraft-Rollenförderern, Abbildung 1 (M), kann ein fester Übergabepunkt für jede Holzsorte im Vorfeld definiert werden. Dabei sind die Förderer aus Platzgründen übereinander angeordnet und so ausgelegt, dass mindestens 20 Rohteile vorgehalten werden. Der Füllstand der einzelnen Förderer wird dabei elektronisch überwacht und entsprechende Personen benachrichtigt, sobald ein kritischer Wert unterschritten wird.

Die Materialausgabe ist als schiefe Ebene angelegt. Damit kann auch hier zusätzliche preisintensive Aktorik vermieden werden. Um die Kompaktheit der Gesamt-

zelle zu erhöhen ist die Ausgabe direkt unter der Materialzuführung implementiert. Mit dem Separieren der Auffüll- und Entnahmepunkte des Nutzers vom aktiven Teil der Zelle kann der Fräsvorgang auch während des Materialhandlings gefahrlos weitergeführt werden.

Mit der Anbindung an eine zentrale Werkstattabsaugung und einem gefüllten Rohteillager ist die Zelle somit in der Lage bis zu 60 Fertigungsaufträge völlig autonom abzuarbeiten.

3. Programmatisches Erstellen des Fräsprogramms und Nutzerinteraktion

Für die Erstellung von Fräsoptionen werden heutzutage vorrangig CAM-Applikationen eingesetzt. Jedoch bieten diese meist nur geschultem Personal unter hohem Adaptionen- und Programmieraufwand entsprechende Ergebnisse. Zusätzlich dazu sind Part-to-Tool-Anwendungen außerhalb von spezialisierter Software wie SprutCAM Robot oder Robotmaster nicht implementiert. Ebenso würden die Individualisierungsoptionen, wie bspw. der Schriftzug auf dem Sockel des Souvenirs zu einem hohen Adaptionen- bzw. Nachbearbeitungsaufwand in der Software selbst führen.

Um diese Hürden für den aktuellen Ansatz zu umgehen, wird eine einfache Bedienoberfläche implementiert, über welche die möglichen Anpassungen vorgenommen und der adaptierte Fräscode automatisch erzeugt wird.

3.1. Erstellung der Roboterpfade

Den Ausgangspunkt bilden hierbei Fräsprogramme, welche im Vorfeld erzeugt wurden. Dafür wurden zunächst mit einer CAM-Software Tool-to-Part 5-Achs-Operationen erstellt und der entsprechende Automated-Programming-Tool Code (APT-Code) exportiert. Dieser beschreibt maschinenunabhängig die Bewegungen des Fräasers. Die im APT-Code beschriebenen Bahnen setzen sich dabei allein aus linearen und kreisförmigen Bewegungen zusammen. [13-14] Unter Betrachtung der Bewegungszusammenhänge aus Abbildung 2 kann somit ein Postprozessor für den eingesetzten Roboter erstellt werden. Dabei wird bspw. eine lineare Punkt-zu-Punkt-Bewegung durch einfache Koordinatentransformation konvertiert. Kreisbewegungen werden als Drehung des Werkstückes um den Mittelpunkt der entsprechenden Bahn dargestellt.

Die Abbildung der verschiedenen Prozessschritte erfolgt durch die Verwendung von unterschiedlichen Koordinatensystemen (KOS). So erhält jeder Prozessschritt ein eigenes KOS und der im APT-Code abgebildete Werkzeugwechsel wird als Wechsel des KOS übersetzt. Etwaige Bahnen für den Transfer von einer Spindel zur nächsten werden ebenfalls durch den Postprozessor integriert.

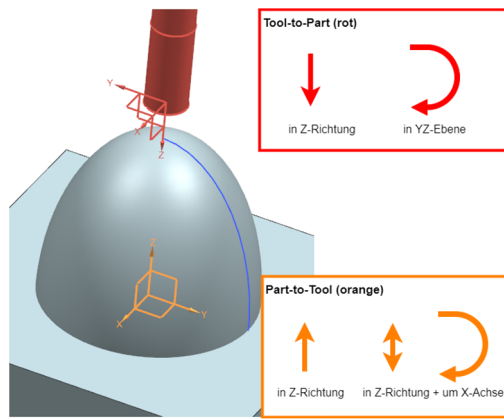


Abbildung 2: Bewegungskombination zum Fräsen der Bahn (blau)

Durch die Verwendung eines Roboters stehen in der aktuellen Anwendung sechs Bewegungsachsen zur Verfügung. Da der Ausgangscode jedoch lediglich mit einer 5-Achs-Maschine erzeugt wurde, ist das Zielsystem unterbestimmt. Dies kann jedoch genutzt werden um die Steifigkeit der jeweiligen Posen zu optimieren und etwaige Kollisionen im Bearbeitungsraum sowie Singularitäten in der Bahnplanung zu vermeiden. [15-17]

3.2. Nutzerinteraktion und Adaption der Fräsbahnen

Über eine einfache Bedienoberfläche erfolgt jegliche Nutzerinteraktion zur Erstellung der Fräsaufträge. Dabei werden durch eine geführte Schritt-für-Schritt Anleitung sämtlich mögliche Parameter abgefragt und in ein Vorschaumodell übersetzt. Mögliche Adaptionen sind bspw. die Wahl der Holzsorte, die Ausrichtung sowie diskrete Änderungen der Position und Größe des Objektes auf dem Sockel. Zusätzlich dazu erfolgt hierüber die Eingabe von Freitext oder Skizzen zur Individualisierung des Sockels. Eine beispielhafte Umsetzung ist dabei in Abbildung 3 dargestellt.



Abbildung 3: Beispielhaftes Designobjekt auf der Bedienoberfläche

Passend zum ausgewählten Designobjekt werden daraufhin die im Vorfeld erzeugten Roboterbahnen abgerufen und an die jeweiligen Parameter angepasst. Die Auswirkungen der Parameteränderungen wurden dabei ebenfalls im Vorfeld analysiert und in Algorithmen zur Adaption der Ausgangsbahnen übersetzt.

Da die Individualisierung des Sockels über Freihand-Eingaben erfolgt, können die resultierenden Bahnen nicht wie bei der Parameteradaption erzeugt werden. Hierbei wird ausgenutzt, dass die Kontur eine reine zweidimensionale Kontur ist und somit die resultierenden Roboterbewegungen nur in einer Ebene umgewandelt werden müssen. Dafür wird die erfolgte Eingabe zunächst in G-Code umgewandelt. Unter Berücksichtigung der oben beschriebenen und in Abbildung 2 dargestellten Bewegungszusammenhänge können die resultierenden Roboterbahnen im Koordinatensystem der Gravurspindel ebenfalls mittels Postprozessor konvertiert werden.

Zusätzlich zur Erstellung von Fräsaufträgen kann der Gesamtzustand der Zelle über die Bedienoberfläche angezeigt werden. Dafür werden die erhobenen Parameter so aufbereitet, dass Fehler auch von nicht geschultem Personal identifiziert und behoben werden können. Dies ermöglicht den Betrieb der Fertigungszelle ohne erhöhten Schulungs- und somit Personalaufwand.

4. Ausblick

Das im vorliegenden Beitrag beschriebene Konzept wird derzeit in einer Tischlerei in Mecklenburg-Vorpommern umgesetzt. Ziel ist es hier, kundenindividuelle Souvenirs automatisiert und mit möglichst geringem Personaleinsatz zu fertigen. Dafür wird die Mensch-Maschine-Schnittstelle so aufbereitet, dass der Kunde sich selbstständig sein Designobjekt erstellen und in Auftrag geben kann. Durch eine zusätzliche abgesicherte Produktentnahme erfordert im gesamten Prozess lediglich das Auffüllen des Rohteillagers Personalressourcen.

Da nach aktuellem Kenntnisstand eine Forschungslücke im Bereich des Einsatzes von kleineren Industrierobotern bei Fräsaufgaben vorhanden ist, können keine belastbaren Angaben zur erreichbaren Genauigkeit oder zu realen Prozesszeiten gemacht werden. Hierfür ist geplant, dass mit der aufgebauten Zelle repräsentative Fräsvorgänge aufgenommen und hinsichtlich der angesprochenen Faktoren bewertet werden. Bei ausreichendem Kenntnisstand sollen die möglichen Fräsobjekte auf Funktionskomponenten im Möbelbau, beispielsweise Handgriffe oder auch Verbindungselemente erweitert werden. Dies würde die Einsatzmöglichkeiten des Systems erheblich erweitern.

Bedingt durch die derzeitige Erstellung der Fräsprogramme im Vorfeld und der einfachen Adaption auf die Individualisierungsoptionen kann der vollautomatisierte Ansatz des Gesamtprozesses noch nicht ausgereizt wer-

den. Hierfür ist es notwendig auch für unbekannte Bauteile eine automatisierte Pfadgenerierung zu implementieren. [18]

Literaturverzeichnis

- [1] Bundesagentur für Arbeit, 2020. Fachkräfteengpassanalyse 2019, Nürnberg (Germany).
- [2] Brunello, G., Wruuck, P. Skill Shortages and Skill Mismatch in Europe: A Review of the Literature, Bonn. IZA, 2019(12346). <http://ftp.iza.org/dp12346.pdf>. Accessed 15 December 2020.
- [3] Brunete, A., Gambao, E., Koskinen, J., Heikkilä, T. et al., 2018. Hard material small-batch industrial machining robot, *Robotics and Computer-Integrated Manufacturing* 54, S. 185-199.
- [4] Çakır, M., Hekimoğlu, B., Deniz, C., 2019. Path Planning for Industrial Robot Milling Applications, *Procedia Computer Science* 158, S. 27-36.
- [5] Gołda, G., Kampa, A., 2014. Modelling of Cutting Force and Robot Load during Machining, *AMR (Advanced Materials Research)* 1036, S. 715-720.
- [6] Song, Y., Chen, Y.H., 1999. Feature-based robot machining for rapid prototyping, *Journal of Engineering Manufacture* 213, S. 451-459.
- [7] Leal-Muñoz, E., Diez, E., Marquez, J., Vizan, A., 2019. Feasibility of machining using low payload robots, *Procedia Manufacturing* 41, S. 594-601.
- [8] Ayari, O., Bouali, A., Méausoone, P.-J., 2020. Cutting forces and accuracy characterization during wood machining with serial robots, *European Journal of Wood and Wood Products* 78, S. 767-775.
- [9] Koch, J. Der eiserne Kollege: Tischlermeister Axel Eigenstetter arbeitet mit einem Fünf-Achs-Industrieroboter von Kuka. <https://www.bm-online.de/praxis-und-kollegentipps/zu-gast-beim-kollegen/der-eiserne-kollege/>. Accessed 9 December 2020.
- [10] Kollmann, F.F.P., Cote, Wilfried, A. Jr., 1968. Principles of Wood Science and Technology: I Solid Wood. Springer Berlin Heidelberg, Berlin, Heidelberg.
- [11] Dryba, S., Meißner, J., Wanner, M.-C., Wurst, O., 2017. Hochpräzises Bearbeiten von Schiffspropellern: High-precision machining of very large ship propellers, *wt Werkstattstechnik online* 107, S. 182-188.
- [12] Rileys Surface World. INTEC - ABB IRB 2400 ROBOTIC FINISHING CELL. <https://www.rileysurface-world.co.uk/live/machines2/25800.pdf>. Accessed 9 December 2020.
- [13] Deutsches Institut für Normung. DIN 66215-1:1974-08, Programmierung numerisch gesteuerter Arbeitsmaschinen; CLDATA, Allgemeiner Aufbau und Satztypen, Berlin. Beuth Verlag GmbH (66215).
- [14] Nof, S.Y., Editor, 2009. Springer handbook of automation: With 149 tables. Springer, Berlin.
- [15] Schneider, U., Diaz Posada, J.R., Verl, A., 2015 - 2015. Automatic pose optimization for robotic processes, in 2015 IEEE International Conference on Robotics and Automation (ICRA), IEEE, S. 2054-2059.
- [16] Xiong, G., Ding, Y., Zhu, L., 2019. Stiffness-based pose optimization of an industrial robot for five-axis milling, *Robotics and Computer-Integrated Manufacturing* 55, S. 19-28.
- [17] Gonul, B., Sapmaz, O.F., Tunc, L.T., 2019. Improved stable conditions in robotic milling by kinematic redundancy, *Procedia CIRP* 82, S. 485-490.
- [18] Nagata, F., Okada, Y., Sakamoto, T., Kusano, T. et al., 2017. Preprocessor with spline interpolation for converting stereolithography into cutter location source data, *IOP Conference Series: Earth and Environmental Science* 69, S. 12115.

Möglichkeiten der Kraftregelung an Leichtbaurobotern

Christian Thormann, Alexander Winkler

Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

In diesem Beitrag werden Möglichkeiten der Kraft-/ Momentregelung an Leichtbaurobotern aufgezeigt. Für die Versuche werden zwei Roboter verwendet, einer ist mit einem internen Kraft-/ Momentsensor ausgestattet, bei dem anderen Roboter werden die Motorströme der Antriebe zur Ermittlung der Kontaktkräfte und -momente benutzt. Bei den meisten kommerziellen Systemen ist es schwierig, Kraftregelalgorithmen zu implementieren. Häufig sind die Programmierfunktionalitäten sehr eingeschränkt. Bei den hier verwendeten Systemen sind Algorithmen zur Kraftregelung vorprogrammiert. Diese sollen getestet und bewertet werden. Dazu soll der Roboter zunächst den Kontakt zu einer Oberfläche herstellen. Danach wird die Kontaktkraft sprungförmig geändert. Zudem soll eine komplexe Aufgabe zur Kraftregelung implementiert werden.

1. Einleitung

In Situationen, bei denen der Roboter physischen Kontakt zu seiner Umwelt hat, ist es sinnvoll, die Roboterbewegung durch die Kontaktkräfte zwischen Roboter und Umwelt zu steuern. Diese können mit einem Kraft-/ Momentsensor (KMS) gemessen werden. Mittels dieser Messwerte kann dann die Roboterbewegung beeinflusst werden. Dieser Vorgang wird als Kraft-/ Momentregelung bezeichnet. Der Sensor ist hierbei üblicherweise am Roboterflansch montiert.

Die Kraft-/ Momentregelung wird seit vielen Jahren bzw. Jahrzehnten erforscht. Eine schöne Übersicht über wichtige Konzepte findet man in [1]. Seit einigen Jahren geht der Trend hin zu sogenannten Leichtbaurobotern. Bei denen ist das Verhältnis zwischen Nutzlast und Eigengewicht näher an 1:1 als 1:10, was bei konventionellen Industrierobotern der Fall ist. Leichtbauroboter wurden ursprünglich für die Raumfahrt entwickelt. Die geringe Eigenmasse führte zu neuen Einsatzgebieten wie z.B. der Mensch-Roboter Interaktion, welche einen wachsenden Trend darstellt [2]. Daher findet man heutzutage zahlreiche Anbieter von Leichtbaurobotern.

Einer dieser Anbieter ist Universal Robots, deren UR-Serie sehr verbreitet ist. Für die Teach-In Programmierung kann der Roboter handgeführt werden. Dazu werden die Motorströme analysiert, ein Kraftsensor ist nicht nötig. Das Handführen entspricht einer sensorlosen Kraftregelung. Damit könnten auch kraftgeregelte Aufgaben, wie z.B. Oberflächenbearbeitung, ohne einen KMS durchgeführt werden. Zudem bietet UR eine erweiterte Baureihe an (URe), welche mit einem KMS im Flansch ausgestattet ist. Dieser ist in der Robotersteuerung integriert.

In diesem Beitrag sollen die Möglichkeiten der Kraft-/ Momentregelungen der Systeme UR und URe miteinander verglichen werden. Die Qualität der Kraftregelung wird anhand der Kraftverläufe bewertet, die beim Kontaktieren der Umgebung durch den Endeffektor auftreten. Des Weiteren werden Sprungantworten bei bestehendem Kontakt ausgewertet. Zur Demonstration der

Kraftregelung wird außerdem eine komplexe und anschauliche Aufgabe implementiert. Der somit entstandene Demonstrator kann z.B. verwendet werden, um Schülerinnen und Schülern die Robotik näher zu bringen.

Dieser Beitrag ist wie folgt gegliedert: Im nächsten Abschnitt wird der Versuchsaufbau beschrieben. Abschnitt 3 erläutert die Implementierung von Kraftregelalgorithmen auf der UR-Robotersteuerung. Dazu werden die notwendigen Funktionen der UR-Script Programmierung [3] vorgestellt. Anschließend werden im Abschnitt 4 die Versuche beschrieben und ausgewertet, um die Qualität der Kraftregelung der beiden Roboter zu vergleichen. Im Abschnitt 5 wird die Implementierung einer komplexen Aufgabe zur Kraft-/ Momentregelung präsentiert. Der Roboter soll hierbei kraftgeregelt ein Bild zeichnen, welches als Datei vorgegeben ist. Im letzten Abschnitt erfolgt eine Zusammenfassung.

2. Versuchsaufbau

Für die Versuche wurden zwei Roboter, UR3 und UR3e, verwendet. Beide sind 6-achsige Gelenkarmroboter mit einer Nutzlast von 3 kg. Sie sind auch im Rahmen der Mensch-Roboter Interaktionen einsetzbar. Dazu verfügen die Roboter über Leistungs- und Kraftbegrenzungen, wodurch sie auch ohne Schutzzaun betrieben werden können. Wie bereits erwähnt, ist die e-Serie mit einem integrierten KMS ausgestattet. Dessen Messbereich beträgt ± 30 N für Kräfte und ± 10 Nm für Momente.

Im ersten Versuch sollen die Roboter daraufhin verglichen werden, wie sie im Modus der Kraftregelung reagieren, wenn sie aus einer Bewegung heraus auf die Kontaktumgebung treffen. Dazu werden die Kontaktkräfte durch einen am Flansch montierter externer KMS aufgezeichnet, da die Messwerte der internen Sensoren von der Robotersteuerung nicht in Echtzeit bereitgestellt werden. Der Sensor KMS40 der Firma Weiss Robotics besitzt einen Messbereich von ± 120 N und ± 3 Nm für die Kräfte bzw. die Momente. Die Datenübertragung zur Robotersteuerung oder an einen PC erfolgt via Ethernet.

In der Abbildung 1 sind zwei Szenarien zum Herstellen des Kontaktes zwischen Roboter und Umgebung dargestellt. Das erste mit einer steifen, das zweite mit einer nachgiebigen Oberfläche. Der externe KMS ist am Flansch des UR3e zu erkennen. Der Aufbau für den UR3 war identisch.

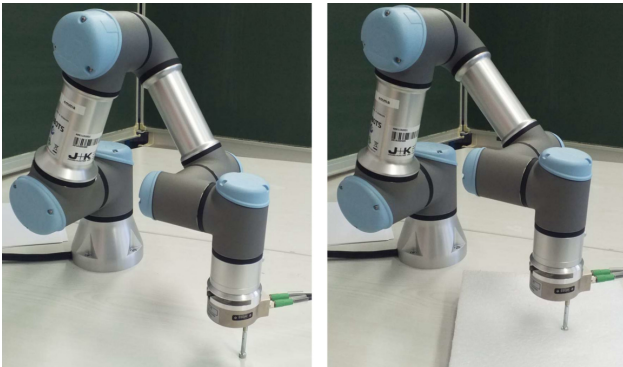


Abbildung 1: Steife und nachgiebige Kontaktumgebung

3. Implementierung der Kraftregelung in die UR-Leichtbauroboter

Die Robotersteuerungen unterstützen Multi-Threading. Die Kraftregelung muss parallel zum Hauptprogramm laufen, welches die Bewegungsbefehle beinhaltet. Die Roboterbewegung kann Achsinterpoliert *movej*, Linearinterpoliert *moveL* oder Kreisinterpoliert *movec* sein. Ein neuer Thread wird mit dem Befehl *thread* deklariert. Es ist notwendig, die *sync*-Anweisung in jedem Thread aufzurufen, damit andere Threads auch ausgeführt werden können. Für die Implementierung der Kraft-/ Momentregelung muss der Befehl *force_mode* (*Feature, Compliance, Force, Type, Limits*) in UR-Script verwendet werden [3]. Die Parameter haben folgende Bedeutung:

- *Feature*: Spezifiziert das Koordinatensystem, in welchem die Regelung arbeiten soll. Hier kann z.B. das Werkzeugkoordinatensystem ausgewählt werden.
- *Compliance*: Aktiviert die kartesischen Freiheitsgrade für die Kraft-/ Moment-Regelung.
- *Force*: Dieser Vektor beinhaltet die Sollwerte der Kräfte und Momente in N bzw. Nm.
- *Type*: Definiert die Transformation. Der Wert 2 bedeutet, dass keine Transformation stattfindet.
- *Limits*: Hier werden die maximale Geschwindigkeiten eingestellt, die der Roboter in der entsprechenden Freiheitsgraden verfahren darf. Für Positionsgeregelte Freiheitsgrade stellen diese Werte maximale Weg-/ Winkelabweichungen dar.

Nach der Deklaration des Threads zur Kraftregelung muss der Sensor tariert werden. Hierfür wird die Funktion *zero_ftsensor()* verwendet. Danach wird eine Wartezeit von mindestens 20 ms empfohlen. Nachdem der Roboter an den Startpunkt bewegt wurde, kann die Kraftregelung mittels der Funktion *run* aktiviert werden. An-

schließend folgt eine Wartezeit von 500 ms. Nach der Aktivierung des Threads werden alle Bewegungsbefehle mit Kraft-/ Momentregelung durchgeführt. Die Regelung arbeitet nur wenn Bewegungsbefehle aufgerufen werden. Ist keine Bewegung gewünscht müssen Bewegungsbefehle mit sehr kleinen Distanzen programmiert werden. Die Befehle *kill* und *end_force_mode* beenden den Regelungs-Thread und stoppen den Sensor. Die Abbildung 2 zeigt beispielhaft ein Programm in UR-Script, in welchem die verschiedenen Funktionen verwendet werden.

```

thread MyForceCtrlThread() :
  while(True) :
    force_mode(get_actual_tcp_pose(),
              [0,0,1,0,0,0],
              [0,0,1.5,0,0,0],2,
              [0.1,0.1,0.1,0.35,0.35,0.35])

    sync()
  end
end

zero_ftsensor()
sleep(0.2)
moveL ...

global MyForceCtrl=run MyForceCtrlThread()
moveL ...
...
kill MyForceCtrl
end_force_mode()
moveL ...

```

Abbildung 2: Beispielprogramm zur Kraft-/ Momentregelung

4. Versuche und Resultate

In diesem Abschnitt werden die Potentiale der beiden Roboter hinsichtlich der Kraft-/ Momentregelung untersucht und ausgewertet. Dazu werden Experimente zur Kontaktierung aus dem freien Raum heraus und Sollwertänderungen bei bestehenden Kontakt zwischen Endeffektor und Umgebung durchgeführt.

4.1. Kraftgeregeltes Kontaktieren

Im ersten Versuch soll das Verhalten beim Kontaktieren verglichen werden. Hierfür wurde das Werkzeug des Roboters in einem geringen Abstand zur Oberfläche platziert. Die Kraftregelung wurde so aktiviert, dass in z-Richtung des Werkzeugs ein Kraftsollwert F_d von 5 N vorgegeben wurde. Die Kontaktkräfte wurden mit dem externen KMS gemessen und aufgezeichnet. Die Abbildung 3 zeigt den Zeitverlauf der Kräfte beim Kontaktieren der harten und der weichen Oberfläche.

Zunächst kann beim Kontaktieren ein großes Überschwingen beobachtet werden. Im Vergleich der Robotersysteme fällt auf, dass das Überschwingen am UR3e größer ist. Der Grund dafür liegt wahrscheinlich daran,

dass der Kraftregelkreis um den Positionsregelkreis herumgeschlossen wurde. Diese Struktur wird auch als positions-basierte Kraftregelung bezeichnet. Diese Regelstruktur besitzt eine schlechte Dynamik und es können große Kraftspitzen auftreten.

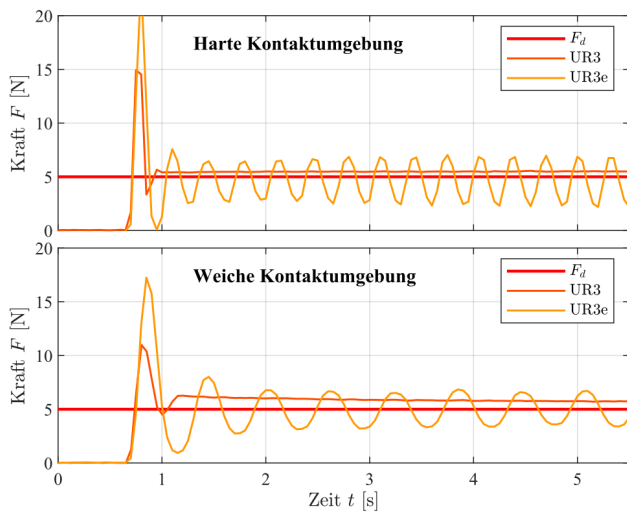


Abbildung 3: Kräfte während des Kontaktierens

Beim UR3 wurde der Kraftregelkreis über die Gelenkmomente bzw. über die Motorenströme realisiert. Daher ist die Kraftregelung am UR3 dynamischer als die am UR3e. Nachdem der UR3e-Roboter die Oberfläche berührt hat, beginnt die Kraft zu schwingen. Die Amplitude beträgt etwa 2 N. Der Grund für die Schwingung liegt wahrscheinlich an der Zeitverzögerung in der positions-basierten Kraftregelung. Je nach Oberfläche und Nachgiebigkeit können unterschiedliche Kraftspitzen während des Kontaktierens beobachtet werden. Eine harte Oberfläche erzeugt eine höhere Kraftspitze. Zudem beeinflusst die Nachgiebigkeit die Periodendauer der Schwingung. Die Schwingungsfrequenz beträgt etwa 3,0 Hz für die harte und 1,6 Hz für die weiche Oberfläche. Der UR3 schwingt nicht, wenn er die Oberfläche berührt. Aufgrund des fehlenden internen Kraft-/ Momentsensors wurde jedoch eine statische Abweichung der Kontaktkraft gemessen.

4.2. Sollwertänderung in der Kraftregelung

Im zweiten Experiment wurden die Sprungantworten der Kraftregelkreise für beide Roboter aufgezeichnet. Die Ergebnisse sind in der Abbildung 4 dargestellt. Zum Beginn des Experimentes berührt das am Roboterflansch befestigte Werkzeug die Oberfläche. Die Sollkraft F_d ist auf 5 N eingestellt. Nach einigen Sekunden wird F_d auf 10 N gesetzt, nach einer weiteren Wartezeit dann auf 20 N. Es wurde festgestellt, dass die Kraftregelung am UR3 recht schlecht funktioniert. Während der Änderung von F_d von 5 N auf 10 N konnte keine Reaktion festgestellt werden. Nach der zweiten Sollwertänderung wird die Kraft am UR3 vergrößert. Eine Abweichung vom Sollwert um ca. 40% kann beobachtet werden. Die Kraftregelung am UR3 kann daher nicht für Aufgaben verwenden

det werden, bei denen eine exakte Kontaktkraft notwendig ist. Die Kraftregelung am UR3e ist im Gegensatz zum UR3 genauer und weist eine bessere Qualität auf. Die Kontaktkraft folgt dem Sollwert. Die im vorherigen Experiment beobachteten Schwingungen sind immer noch vorhanden. Im Fall von $F_d = 20$ N ist die Amplitude bei der steiferen Umgebung etwas geringer.

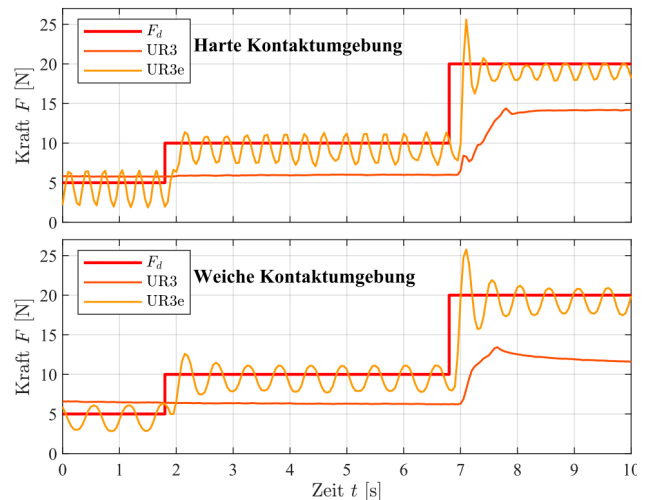


Abbildung 4: Sprungantworten der geschlossenen Regelkreise

5. Verwendung der Kraftregelung am Beispiel einer komplexen Aufgabe

Die Kraftregelung wurde mittels einfacher Experimente bewertet. Nun soll eine komplexere Aufgabe mit dem Leichtbauroboter durchgeführt werden. Beispiele für Aufgaben, welche kraftgeregelt bearbeitet werden können, sind z.B. Montage, Oberflächenbearbeitung oder Mensch-Roboter Interaktion. Ein Beispiel der Oberflächenbearbeitung ist das Zeichnen. Diese Aufgabe kann auch in einem Robotik-Labor implementiert werden.

In unserem Szenario soll der Roboter ein Bild auf ein Blatt Papier, welches auf dem Tisch liegt, zeichnen. Hierfür wurde der UR3e gewählt, da dieser über einen integrierten KMS verfügt, was günstiger für die Kraftregelung schein. Zum Zeichnen wird ein Bleistift verwendet, welcher am Flansch befestigt wurde. Die Bildquelle kann entweder eine Kamera oder eine Bilddatei sein. Zudem wird der Stift gelegentlich kraftgeregelt angespitzt. Die Implementierung besteht aus zwei Teilen.

Zunächst läuft eine Applikation auf einem PC. In dieser werden die Bildverarbeitung und die Bahnberechnung durchgeführt. Die Bewegungsdaten werden an die Robotersteuerung gesendet, welche die Bewegungssteuerung und die Kraftregelung durchführt. Die PC-Applikation wurde in MATLAB erstellt und beinhaltet folgende Funktionalitäten, die mittels einer entsprechenden Bedienoberfläche gesteuert werden können:

- Aufnahme eines Kamerabildes oder laden einer Bilddatei
- Manuelles zuschneiden des Bildes

- Automatische Reduzierung der Auflösung
- Umwandlung in ein Binärbild
- Berechnung der Roboterbahn

Der zweite Programmteil, der auf der Robotersteuerung abgearbeitet wird, empfängt Positionsdaten vom PC und aktiviert die Roboterbewegung in Form einer Linearinterpolation. Parallel hierzu wird die Kraftregelung für die z-Achse des Werkzeugkoordinatensystems aktiviert. Als Sollkraft wurde 1 N gewählt. Diese Einstellung kann auch durch die Vektor \mathbf{F}_d ausgedrückt werden, der die Sollwerte der Kräfte bzw. Momente enthält.

$$\mathbf{F}_d = [- \quad - \quad 1,0 \text{ N} \quad - \quad - \quad -]^T$$

Die Einstellungen können in der *force_mode*-Anweisung von UR-Script geändert werden, die im Abschnitt 3 bereits beschrieben wurde. Die Kraftregelung wird parallel zur Bewegungssteuerung abgearbeitet, welche aus dem Bild, das gezeichnet werden soll, generiert wird. In Abbildung 5 (links) ist zu erkennen, dass das kraftgeregelt gezeichnete Zeichen gut funktioniert. Das Bild wurde durch den Roboter gezeichnet ohne den Stift zu beschädigen.

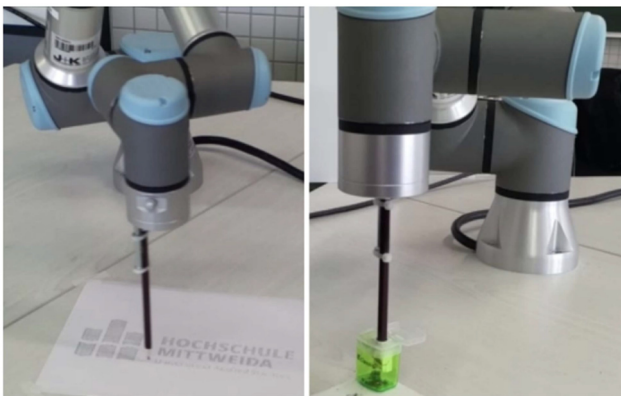


Abbildung 5: Versuchsaufbau für kraftgeregelt gezeichnete Zeichen und Anspitzen des Stiftes

Das rechte Bild in Abbildung 5 zeigt das Anspitzen des Stiftes. Dieses Programm besteht aus zwei Teilen. Zuerst wird der Stift in den Spitzer eingeführt. Hierfür wurde Kraftregelung mittels \mathbf{F}_{d1} wie folgt parametrisiert:

$$\mathbf{F}_{d1} = [0 \quad 0 \quad 1,0 \text{ N} \quad 0 \quad 0 \quad -]^T$$

Im Gegensatz zum Programm des Zeichens wird hier keine Bewegungssteuerung benötigt. Die Kraftregelung wird allerdings nur parallel zu Bewegungsbefehlen abgearbeitet. Daher wurde eine kleine Bewegung von 1 mm programmiert. Typischerweise würde eine solche Bewegung sehr schnell gehen. Die *move*-Anweisung in UR-Script jedoch kann so parametrisiert werden, dass die Dauer der Bewegung eingestellt werden kann. In diesem Fall wurde eine Dauer von 5 s eingestellt, innerhalb der Roboter den Stift in den Spitzer drückt.

Im zweiten Schritt wird das Spitzen durchgeführt. Während der Kraft-/ Momentregelung werden die Freiheitsgrade F_z , M_x und M_y aktiviert, dargestellt in \mathbf{F}_{d2} :

$$\mathbf{F}_{d2} = [- \quad - \quad 5,0 \text{ N} \quad 0 \quad 0 \quad -]^T$$

Parallel zur Kraftregelung wird der Stift um die z-Achse gedreht. Die Drehung wurde mittels Linearinterpolation realisiert. Gedreht wurde 10-mal um 90°. Es wurde festgestellt, dass das kraftgeregelt Anspitzen funktionierte.

6. Zusammenfassung

In diesem Beitrag wurden zwei weit verbreitete Leichtbauroboter der Typen UR3 und der UR3e von Universal Robots bezüglich der Kraftregelung untersucht. Der UR3e besitzt einen integrierten KMS. Die Programmierung mittels UR-Script war für beide Roboter identisch.

Es wurden einige Versuche durchgeführt, in welchem ein Freiheitsgrad, die z-Achse des Werkzeugkoordinatensystems, kraftgeregelt arbeitete. Als erstes kontaktierte der Roboter die Tischfläche. Danach wurde die Kontaktkraft geändert. Das Experiment wurde auf einer weichen und einer harten Kontaktumgebung durchgeführt. Die Kontaktkräfte wurden durch einen externen KMS gemessen. Der UR3e war in der Lage die Kontaktkraft zufriedenstellend zu regeln. Eine genaue Kraftregelung war beim UR3 nicht möglich, da die Regelkreise über die Gelenkmomente bzw. die Motorströme geschlossen wurden. Beim UR3e war ein starkes Schwingen der Kontaktkraft zu beobachten, was nachteilig ist. Beim UR3 konnte keine Schwingung beobachtet werden. Aus diesem Grund kann die Kraftregelung mit dem UR3 für Anwendungen verwendet werden, bei denen ein stabiler Kontakt zwischen dem Endeffektor und der Umgebung erforderlich ist und der tatsächliche Wert der Kontaktkraft nicht so wichtig ist. Aufgrund der verschiedenen Steuerungsstrukturen kann der Schluss gezogen werden, dass der UR3 nachgiebiger ist als der UR3e. Am Ende wurde das Zeichnen einer Figur durch einen Roboter ermöglicht, was eine anspruchsvolle Aufgabe ist. Trotz des relativ ungünstigen Verhaltens der Kraftregelung konnte gezeigt werden, dass es möglich ist die Zeichenaufgabe erfolgreich umzusetzen. Zudem wurde das Anspitzen des Bleistiftes durch den Roboter demonstriert.

Literaturverzeichnis

- [1] G. Zeng and A. Hemami, *An overview of robot force control*, Robotica, vol. 15, no. 5, pp. 473–482, 1997.
- [2] B. Finkemeyer, *Towards safe human-robot collaboration*, in Proc. of the IEEE Int. Conf. on Methods and Models in Automation and Robotics (MMAR), 2017, pp. 883–888.
- [3] Universal Robots A/S, *The URScript Programming Language*, 2018.

Danksagung

Diese Maßnahme wird mitfinanziert mit Steuermitteln auf Grundlage des von den Abgeordneten des Sächsischen Landtags beschlossenen Haushaltes.

Nutzung und Potenziale von Instandhaltungsmanagementsystemen für die Optimierung automatisierter Produktionsprozesse

Christian Brenner, Martin Hiersemann, Rolf Hiersemann
Hiersemann Prozessautomation GmbH, Tuchschererstraße 4a, 09116 Chemnitz

Für Prozessstabilität und Optimierung industrieller automatisierter Produktionsabläufe hat die Hiersemann Prozessautomation das System CEMODAS® - Central Modular Data System - entwickelt und industriell in mehreren Anwendungen im Einsatz. Dieses System umfasst in der Grundversion Module zur Betriebs- (BDE) und Qualitätsdatenerfassung (QDE) inkl. Auswertung dieser Daten.

Durch die Integration von Schnittstellen zu verschiedenen Automationssystemen ermöglicht das System Datenauswertungen, wie NEE und OEE, von gesamten Produktionslinien und Werken.

Weiterhin ist der Instandhaltungsassistent als sog. Helpdesk entwickelt worden, der Warn- und Fehlerinformationen für Instandhaltungszwecke erfasst. Die Informationen werden aktiv an Devices, wie PC, Laptop, Mobile Phone oder Watches, übertragen und beinhalten zudem ein Ticket-System.

Aktuell wird an der Integration von Methoden der künstlichen Intelligenz KI in das System gearbeitet.

1. Einleitung

Die automatisierte Fertigung erfordert zunehmend durchgängig gestaltete Automations- und Informationssysteme, die für eine Optimierung der Produktionsprozesse ermöglichen und die Prozessstabilität sichern. Das System CEMODAS® ermöglicht neben Betriebsdatenerfassung (BDE) und Qualitätsdatenerfassung (QDE) mit Datenauswertung und -visualisierung an Maschinen auch anlagenweite Funktionen wie OEE und NEE sowie ein Predictive Maintenance mittels Instandhaltungsmanagement.

2. Funktionsübersicht

Die Funktionen des Systems sind in folgende Module strukturiert.

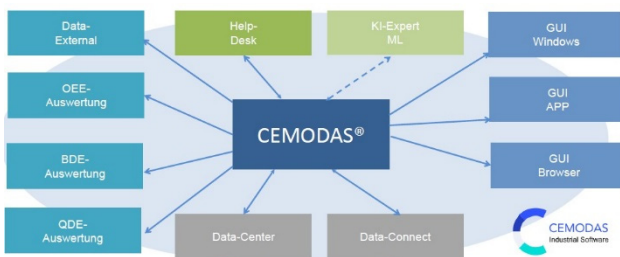


Abbildung 1: Funktionsmodule

Das Modul Data Connect stellt die Verbindung zu den Automationsgeräten der Maschinen (SPS, Roboter, Kamerasationen etc.) her und empfängt die relevanten Daten. Nach einem Datencheck werden die Daten in einer SQL-fähigen Datenbank hinterlegt. Für die Datenerfassung sind Schnittstellen zu verschiedenen Automationssystemen unter Nutzung mehrerer, auswählbarer Datenübertragungsprotokolle entwickelt worden und verfügbar. Auf dieser Basis können erweiterte Auswertefunktionen, wie QDE, BDE und OEE für Produktionslinien und gesamte Werke installiert und betrieben werden.

Mit dem Modul Data external werden Daten auch für sog. Fremdsysteme bereitgestellt.

Das Modul Instandhaltungsassistent ist als sog. Helpdesk verfügbar. Dieser ermöglicht die Visualisierung und damit die Nutzung von Betriebsinformationen, insbesondere von Warn- und Fehlerinformationen, für Instandhaltungszwecke. Die Informationen können am Modul Data Center angefragt werden. Die Datenaufbereitung und notwendige Berechnungen erfolgen vor der Übermittlung innerhalb dieses Moduls. Die Informationen können aktiv an Devices, wie PC, Laptop, Mobile Phone oder Watches, übertragen werden. Weiterhin ist ein Ticket-System implementiert. Mit dessen Nutzung kann lokal einsetzbares Instandhaltungspersonal umgehend kritische Situationen an Maschinen und Anlagen erfassen und beheben und zur Optimierung der Produktionsabläufe beitragen.

Für die breite Anwendbarkeit wurden mehrere Nutzer-/Bediener-Schnittstellen (GUI) entwickelt.

Zur Erweiterung der Funktionalitäten wird derzeit an der Integration von KI-Funktionen gearbeitet, Tools sind in Entwicklung.

3. Grundfunktionen

Zu den Grundfunktionen des entwickelten Systems CEMODAS® gehören Betriebsdatenerfassung (BDE) und Qualitätsdatenerfassung (QDE). Während die Betriebsdaten für das Instandhaltungsmanagement genutzt werden, bilden die Qualitätsdaten die Grundlage für Prozess- und somit Qualitätsstabilität. Die im System abrufbaren Auswertefunktionen betreffen:

Mittelwertbildung

Der arithmetische Mittelwert wird aus einer konkreten Datenmenge, die der Werker vorgeben kann, berechnet und gibt die Tendenz einer Verteilung an. Damit kann der Werker das Wegdriften von Parametern erkennen. Er ist definiert als:

$$\bar{x} = \frac{1}{n} \cdot \sum_{i=1}^n x_i \quad (1)$$

n = Stichprobenumfang

x_i = Wert des i -ten Elements

Standardabweichung

Die empirische Standardabweichung wird auch Stichprobenstreuung genannt. Sie gibt an, wie weit die Stichprobe im Schnitt um das arithmetische Mittel streut.

$$s_g = \sqrt{\frac{1}{n-1} \cdot \sum_{i=1}^n (x_i - \bar{x})^2} \quad (2)$$

Prozessfähigkeit

Die Prozessfähigkeit dient zur statistischen Bewertung eines Prozesses und wird insbesondere beim Einfahren von Produktions- und Prüfanlagen benötigt und angewendet. Angegeben wird damit die Sicherheit des Erreichens vorgegebener Ziele laut Spezifikation. Der Werker hat die Möglichkeit, eine einfache und eine kritische Prozessfähigkeit abzurufen. Damit wird sein Entscheidungsspielraum erheblich erweitert.

Einfache Prozessfähigkeit (C_p)

$$C_p = \frac{OTG - UTG}{6 \cdot s_g} \quad (3)$$

OTG = Obere Toleranzgrenze

UTG = Untere Toleranzgrenze

Kritische Prozessfähigkeit (C_{pk})

$$C_{pk} = \frac{\min(OTG - \bar{x}; \bar{x} - UTG)}{3 \cdot s_g} \quad (4)$$

Insbesondere die Werte Prozessfähigkeit (C_p) und kritische Prozessfähigkeit (C_{pk}) sind für die Nutzung von Prüfsystemen zur Überwachung der Qualitätsparameter der Produktteile von entscheidender Bedeutung. Erst bei Erreichen vorgegebener Werte sind Prüfsysteme für den Produktionseinsatz freigegeben.

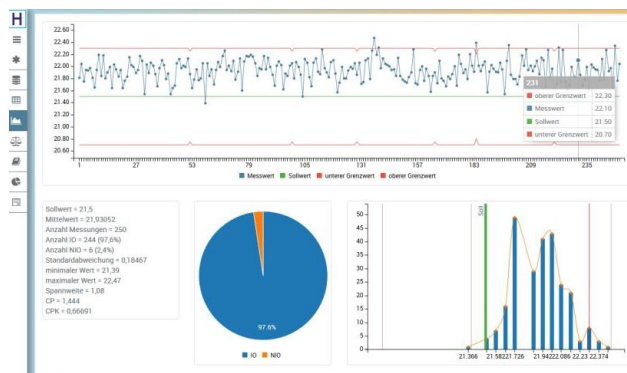


Abbildung 2: Qualitätsdatenauswertung eines Parameters

4. OEE und NEE Berechnungen

Für die Funktionen NEE - Net Equipment Effectiveness (Entspricht der Effektivität der Anlagenlinie unter Berücksichtigung technischer Rahmenbedingungen) und OEE – Overall Equipment Effectiveness (Entspricht der Effektivität der Anlagenlinie unter Berücksichtigung technischer und organisatorischer Rahmenbedingungen, wie Ausfallzeiten (geplant, ungeplant), Stillstand, Wartung/Instandsetzung etc.) werden folgende Stammdaten herangezogen:

Zur Berechnung des Verfügbarkeitsfaktors:

- Schichtplan inklusive Angaben von Pausen
- Wartungsschichten
- Geplante Rüstvorgänge mit hinterlegten Zeiten
- Zusätzliche Pausen (Schulung, Versammlung)

Der Verfügbarkeitsfaktor ist ein Maß für Verluste aufgrund ungeplanter Anlagenstillstände und wird wie folgt ermittelt:

$$f_v = \frac{T_{pp} - \sum T_s}{T_{pp}} = \frac{T_B}{T_{pp}}$$

T_{pp} = geplante Produktionszeit

T_s = Stillstandszeit

T_B = tatsächliche Betriebsdauer

(5)

Für die Berechnung des Leistungsfaktors:

- Soll-Taktzeit an der Maschine für IO und NIO Teile getrennt
- Sollten mehrere Typen auf der Maschine laufen, so muss diese Angabe für alle Typen vorliegen

$$f_l = \frac{T_{pT} \cdot X_g}{T_B}$$

T_{pT} = geplante Taktzeit

X_g = Teile gesamt

T_B = Betriebsdauer

(6)

Der Qualitätsfaktor erfasst Verluste durch fehlerhafte oder nachzuarbeitende Teile nach der Formel:

$$f_q = \frac{X_g - X_{nio}}{X_g}$$

X_g = Teile gesamt

X_{nio} = Teile fehlerhaft

(7)

Die Gesamtanlagenverfügbarkeit OEE wird auf Basis dieser Daten wie folgt berechnet:

$$OEE = f_v \cdot f_l \cdot f_q$$

f_v = Verfügbarkeitsfaktor

f_l = Leistungsfaktor

f_q = Qualitätsfaktor

(8)

Somit gehen auch die Werte aus BDE und QDE in die Berechnungen ein.

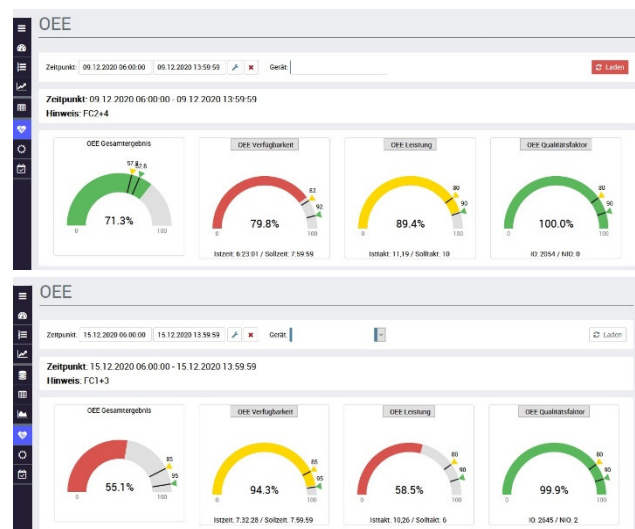


Abbildung 3: OEE-Ermittlungen und Darstellungen einer Anlage

Eine OEE von 72,9 % basiert somit beispielsweise aus einer Beachtung der Faktoren:

- Verfügbarkeitsfaktor = 0,9
- Leistungsfaktor = 0,83086 = 83,09%
- Qualitätsfaktor = 0,96679 = 96,68%
- OEE = 0,9 * 0,83086 * 0,96679 = 0,72294 = 72,29%

Dies ist nur exemplarisch dargelegt. In aktuellen Projekten konnten starke Schwankungen der OEE erfasst werden (siehe Abb. 3).

Die Sichtung der OEE kann somit insbesondere für die vorbereitende Instandhaltung zweckdienlich sein, zeigt diese doch die erschließbaren Potenziale in der Anlagenoptimierung auf.

Weitere Visualisierungen sind als Darstellung der Anlagen-Layouts generierbar. Hierbei können die Kunden die Zuordnung und Anordnung von Darstellungs-Frames zum Layout selbst generieren.

Die Net Equipment Effectiveness Kennzahl ist eine Anlehnung an die OEE. Hierbei wird der Verfügbarkeitsfaktor abgewandelt und die Rüst- und Einrichtzeiten nicht als Verlustzeiten betrachtet. Die NEE ist somit eine Untermenge der OEE. Oft wird in der NEE der Verfügbarkeitsfaktor als Nutzungsgrad bezeichnet, der gegenüber der OEE Berechnungsanteile ausklammert und somit höher ausfällt.

5. Instandhaltungsmanagement

Der Zusammenhang von OEE-Auswertungen und Instandhaltungsmanagement ist in folgender Darstellung einer OEE-Verfügbarkeit zu erkennen.

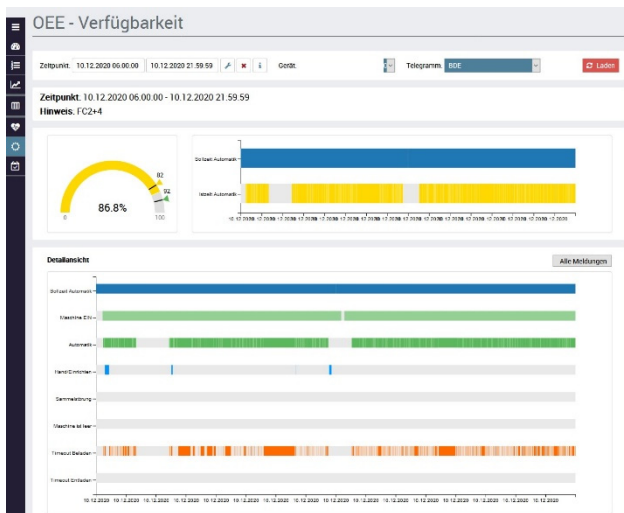


Abbildung 4: OEE-Verfügbarkeit einer Anlage

Zum Heben dieser Potenziale wurde innerhalb des Systems CEMODAS® ein Instandhaltungsmanagementsystem integriert, das als Helpdesk fungiert und mehrere Nutzer integriert.

Das Instandhaltungsmanagement arbeitet mit einem zentralen (virtuellen) Server, der Anforderungen an die Instandhaltung verwaltet. Dieser Server löst Anforderungen aus, die auf allen aktiven, für die Instandhaltung vorgesehenen Netzwerkteilnehmern angezeigt werden, dies in Form von Tickets. Das Ticket kann nun von einen der Instandhalter

aufgegriffen und bearbeitet werden. Dazu erhält dieser die Angaben über den Instandhaltungsgrund. Nach erfolgter Instandhaltung hinterlegt der Instandhalter Informationen zur Instandhaltungsmaßnahme. Diese Informationen werden anschließend im Server analysiert und kategorisiert, z. B. Top-Ten-Störungen. Auf dieser Basis ist anschließend eine vorausschauende Instandhaltung durchführbar.

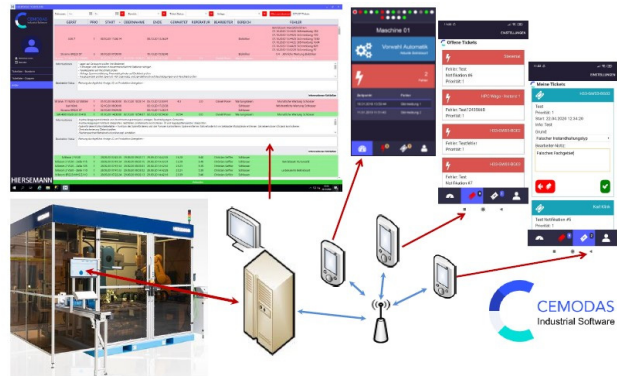


Abbildung 5: Instandhaltungsmanagementsystem Helpdesk-Übersicht

6. Einsatzerfahrungen

Die Einsatzerfahrungen des Systems CEMODAS® sind durchweg positiv. Die Anwendung sowohl der Einzelkomponenten wie QDE und BDE als auch OEE und des Helpdesk dienen zur signifikanten Erhöhung von Produktivität und Prozessstabilität. Anwender erzielen Effekte in Qualitätsverbesserung und Reduzierung von Stillstands- und Wartezeiten, konnte somit Optimierungen realisieren.

Sämtliche Sondermaschinen der Hiersemann Prozessautomation GmbH sind mit dem System CEMODAS® ausgestattet und ermöglichen den Anwendern sowohl die on board Diagnose von Qualität und Betriebszuständen, als auch die rasche Integration in anlagen- und unternehmensweite Auswertesysteme. Auch sind Datenexport in firmenspezifische SCADA-Systeme realisiert.

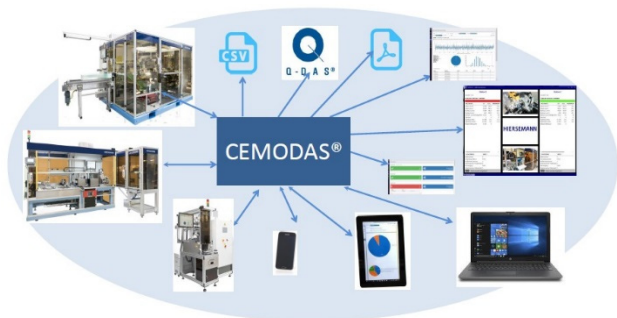


Abbildung 6: Nutzungsübersicht des Systems CEMODAS® mit Anlagen der Firma Hiersemann Prozessautomation

Auf Basis der ersten Installationen der anlagenweiten OEE- und Helpdesk-Funktionen von CEMODAS® bei einem Lieferanten der Automobilindustrie wurden weitere Anlagen mit dem System ausgerüstet. Für alle Neuanlagen sind von diesem Anwender Schnittstellen zum System CEMODAS® vorgeschrieben worden. Dadurch verkürzen sich Zeit und

Aufwendungen des Andockens neuer Anlagen an das System.

Das Ticketsystem wird außerdem im eigenen Haus eingesetzt für Ferndiagnostik und Fernwartung von Anlagen bei weiteren Kunden, die über INTERNET-Verbindung (VPN-gesichert) Instandhaltungsanforderungen generieren.

7. Entwicklungsschwerpunkte

Die Entwicklungsschwerpunkte liegen derzeit auf der Integration von KI-Funktionalitäten sowohl in die anlagenweiten Auswertungen (OEE) als auch in die Helpdesk-Funktionen. Aufbauend auf den Ganglinien von Parametern sollen Vorschauen auf mögliche Ereignisse, wie Qualitätsverlust durch vorgelagerte Prozessschritte oder Ausfallrisiko von Baugruppen, eingebunden werden.

Dazu werden mehrere Methoden des maschinellen Lernens, wie Regressionsverfahren, Clustering, Netztheorie etc., untersucht und getestet. Die Anwendung dieser Methoden befindet sich derzeit im Teststadium auf Demonstratoren.

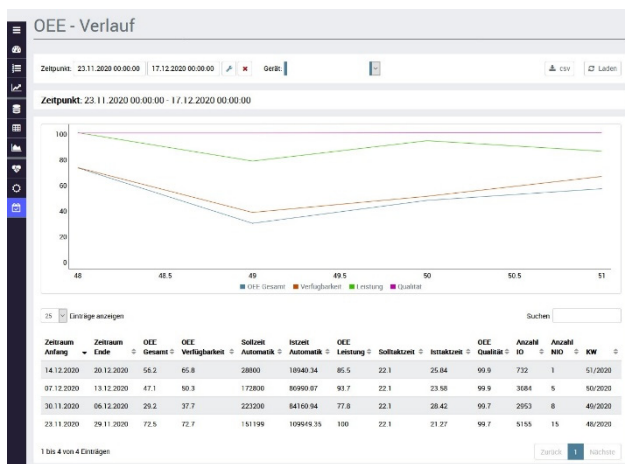


Abbildung 7: Ganglinien des OEE-Verlaufs

Die dargelegten Applikationen basieren auch auf mehreren Forschungs- und Entwicklungsprojekten.

- Modulare funktionsinkludierte Fertigungsstationen (Einzelprojekt *Inklude* ZIM-FK EP180052, gefördert vom BMWi, Projektträger EURONORM [1]).

Im Projekt *Inklude* wurden sämtliche produktionsrelevanten Prozesse in einer Fertigungsstation zusammengefasst, dies unter dem Aspekt der Integration derartiger Stationen in heterogene IIoT-basierte Fertigungsanlagen.

- Harmonisierung der Mensch-Maschine-Interaktion in der Produktion (Verbundvorhaben *HMMI*, Projektträger SAB ANr.100294859, Gefördert von EU-EFRE und SMWA [2]).

Ziel dieses Verbundvorhabens war die Gestaltung von abgestimmten Datenstrukturen in komplexen Fertigungs- bzw. Kommunikationssystemen sowie die Harmonisierung von Visualisierungssystemen

- Harmonisierte Datensicherheit von der Maschine bis in die Cloud (Verbundvorhaben *HDMC*, Projektträger SAB ANr.100376474, gefördert von EU-EFRE und

SMWA [3]).

Ziel dieses Verbundvorhabens ist die Gestaltung sicherer Datenkommunikationen zwischen Anlagen und Datenauswertesystemen. Dies ist insbesondere für die Realisierung von Instandhaltungssystemen mit niedriger Latenz über Firmengrenzen hinweg bedeutsam.

- Produktionsoptimierung zur Ressourcenschonung unterstützt durch künstliche Intelligenz (Verbundvorhaben *KInO*, Projektträger SAB ANr.100361722, gefördert von EU-EFRE und SMWA [4]).

Ziel des Vorhabens ist es, Produktionsanlagen durch die Integration von Methoden der Künstlichen Intelligenz (KI) ressourcenschonender zu gestalten. Dafür werden die an heterogenen Automationsstrukturen abgegriffenen prozessnahen Daten ausgewertet.

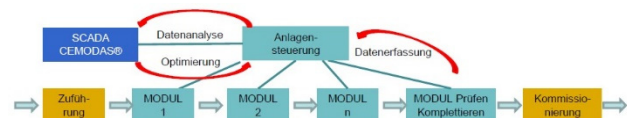


Abbildung 8: Einbindung KI-Funktionen in Produktionsanlagen

Danksagung

Die Fa. Hiersemann Prozessautomation GmbH und die Autoren bedanken sich für die Unterstützung bei Fördermittelegern, Projektträgern und kooperierenden Firmen und Instituten. Die Verantwortung für den Textinhalt liegt bei den Autoren.

Literaturverzeichnis / Quellen

- [1] Hiersemann, R. (2020): Modulare funktionsintegrierte Montage-, Prüf- und Logistikstation für eine effiziente Fabrikintegration (Inklude), Sachbericht, Hiersemann Prozessautomation GmbH, 2020
- [2] Singer, A. (2019): Harmonisierung der Mensch-Maschine-Interaktion in der Produktion (HMMI), Sachbericht, Fraunhofer-Gesellschaft, 2019
- [3] Hoffmann, M. (2019): Harmonisierung der Mensch-Maschine-Interaktion in der Produktion (HDMCI), Projektantrag, Fraunhofer-Gesellschaft, 2019
- [4] Hiersemann, R. (2020): Produktionsoptimierung zur Ressourcenschonung unterstützt durch künstliche Intelligenz (KInO), Zwischenbericht, Hiersemann Prozessautomation GmbH, 2020

Die angegebenen Forschungsprojekte wurden gefördert und unterstützt von.



Einsatz von AutomationML zur automatischen Generierung von Digitalen Zwillingen von Werkzeugmaschinen zur Entwicklung energieeffizienter Produktionssysteme

Nicolai Beisheim, Tobias Ott, Sebastian Amann
Hochschule Albstadt-Sigmaringen, Jakobstraße 1, 72458 Albstadt

Die Entwicklung energieeffizienter Produktionssysteme wie z. B. Werkzeugmaschinen ist ein komplexer Prozess. Viele Fachdomänen müssen während des Konstruktionsprozesses intensiv zusammenarbeiten, um ein optimales Ergebnis zu erzielen. Neben den rein mechanischen Aspekten z. B. dem Leichtbau spielt die Optimierung der Programme der Maschinensteuerung eine immer wichtigere Rolle. Durch Optimierung der Programme hinsichtlich Energieeffizienz kann der Energieverbrauch einer Maschine signifikant gesenkt werden.

Durch den Einsatz von Digitalen Zwillingen von Werkzeugmaschinen können mit Simulationen eine große Anzahl an Parameterstudien für die Optimierung des Energieverbrauchs durchgeführt werden. Da die Generierung solcher akkurater digitaler Zwillinge jedoch sehr aufwändig ist, wurde an einem vernetzten IT-System gearbeitet, welches die Generierung solcher Digitalen Zwillinge von Werkzeugmaschinen automatisiert. Der Artikel beschreibt den Aufbau und die Funktion dieses IT-Systems.

The development of energy-efficient production systems such as production machines is a complex process. Many specialist domains must work together intensively during the design process to achieve an optimum result. In addition to the purely mechanical aspects, e.g. lightweight construction, the optimization of the machine control programs plays an increasingly important role. By optimizing the programs in terms of energy efficiency, the energy consumption of a machine can be significantly reduced.

By using digital twins of production machines, simulations can be used to perform a large number of parameter studies for the optimization of energy consumption. However, since the generation of such accurate digital twins is very time-consuming, work has been done on developing a networked IT system that automates the generation of such digital twins of production machines. The article describes the structure and function of this IT system.

1. Einleitung

Die zunehmende Ausrichtung von Produktionssystemen und Anlagen auf Energieeffizienz ist ein sehr zentraler Bestandteil, um die produzierende Wirtschaft zukunftsfähig zu machen. Zum einen ist die Senkung des Energieverbrauchs aus Umweltschutzgründen und zur Erreichung von globalen Klimazielen von großer Bedeutung, zum anderen hilft die Optimierung des Energieverbrauchs dabei, eine Produktion wirtschaftlicher zu gestalten. Vor allem in Bezug auf die steigenden Energiekosten in Europa ist es für produzierende Unternehmen essenziell, ihre Energiekosten zu senken, um im internationalen Wettbewerb bestehen zu können.

Simulationen als Digitale Zwillinge von technischen Systemen können helfen, Produktionssysteme zu optimieren. An Digitalen Zwillingen kann mit wenig Aufwand eine große Anzahl von Studien durchgeführt werden, um ein ausgewogenes Verhältnis zwischen Produktivität und Energieverbrauch zu ermitteln. Diese Studien können auch parallel zur bereits laufenden Produktion durchgeführt werden, weil die Produktion der realen Anlage nicht unterbrochen werden muss.

Die Modell Generierung von akkuraten Digitalen Zwillingen ist jedoch komplex und es gibt bisher keine einheitliche Software Plattform, welche sowohl die physikalischen Eigenschaften einer Anlage wie auch deren Steuerung abbildet. Da der Energieverbrauch eines solchen Produktionssystems jedoch von verschiedenen Faktoren abhängt, müssen alle an der Entwicklung beteiligten Fachrichtungen eng zusammenarbeiten, um ein optimales Ergebnis zu erreichen. Es wird deswegen eine Plattform benötigt, welche die Generierung von Digitalen Zwillingen weniger aufwändig gestaltet, sowie alle Fachdomänen einer Produktionsmaschine abbildet und ebenfalls fachübergreifend kompatibel zu den bereits heute sich im Einsatz befindlichen Engineering Tools ist.

Der in diesem Artikel vorgestellte Ansatz beinhaltet eine IT-technische Kombination des AutomationML-Schnittstellenformats mit einem automatisierten Prozess der Modell Generierung Digitaler Zwillinge mittels graphenbasierter Entwurfssprachen, um die Anforderungen bei der Entwicklung energieeffizienter Produktionssysteme optimal umzusetzen.

2. Modelle für eine standardisierte Architektur von Industrie 4.0-Projekten

2.1. Referenzarchitekturmodell Industrie 4.0

Industrie 4.0 ist in weiten Bereichen ein sehr abstrakter Begriff, weshalb der Zentralverband Elektrotechnik- und Elektronikindustrie (ZVEI) in Zusammenarbeit mit verschiedenen Industrieunternehmen das Referenzarchitekturmodell Industrie 4.0 (RAMI 4.0) entwickelt.

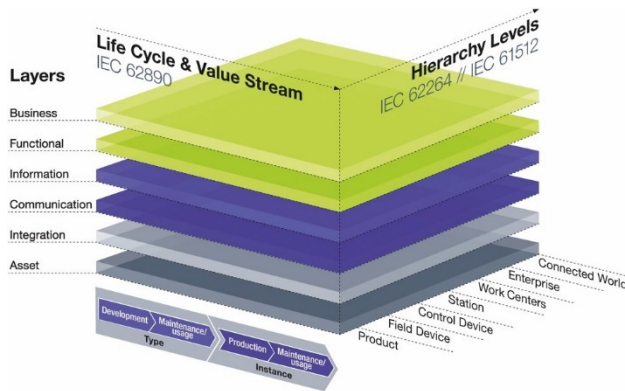


Abbildung 1: Schichtenmodell des RAMI4.0 [1]

1. Schichtenmodell

Den Kern des Referenzmodells bildet ein dreiaxsiges Schichtenmodell, das in Abbildung 1 dargestellt ist. Es bietet die Möglichkeit, beliebige Zustände eines technischen Objekts genannt Asset innerhalb des Produktlebenszyklus darzustellen.

2. Communication capabilities (CP)-Klassifikation

Die CP-Klassifikation soll eine einfache Einordnung von technischen Objekten in das Raster des Referenzarchitekturmodells Industrie 4.0 ermöglichen. Die Matrix der CP-Klassifikation ist in Abbildung 2 dargestellt. Die x-Achse zeigt die Kommunikationsfähigkeit und die y-Achse die Erkennung im System.

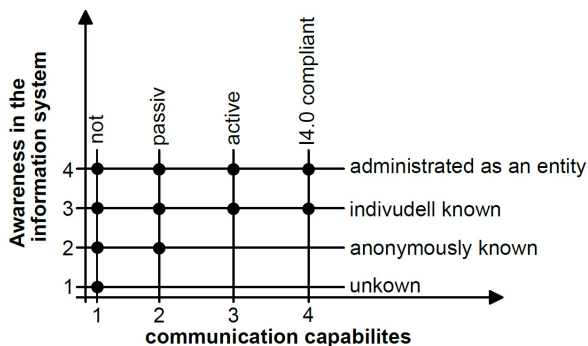


Abbildung 2: CP-Klassifikation des RAMI4.0 [1]

3. Asset-Verwaltungsschale

Um ein technisches Objekt digital abzubilden, wurde im Referenzarchitekturmodell Industrie 4.0 das Konzept der Verwaltungsschale eingeführt. Die Kombination aus

Verwaltungsschale und technischem Objekt wird als Industrie 4.0-Komponente bezeichnet. Nach der bereits vorgestellten Einteilung der CP-Klassifikation entsprechen Industrie 4.0-Komponenten einer CP-Klassifikation von CP43 oder CP44. In diesem Beitrag werden daher nur Elemente dieser Zuordnung betrachtet. Die Verwaltungsschale verwaltet nicht nur die Daten des technischen Objekts, sondern kann auch eigene Funktionen zur Verfügung stellen. Diese werden als digitale Dienste gemäß dem Referenzarchitekturmodell zur Verfügung gestellt. Ein Beispiel für einen solchen Dienst kann die Durchführung einer Diagnose des technischen Objekts durch die entsprechende Verwaltungsschale sein. Auf Basis der erhobenen Daten werden dann z. B. Aussagen über die Restlebensdauer oder den nächsten Serviceeinsatz berechnet.

2.2 AutomationML

Aufgrund der steigenden Komplexität von Industrie 4.0-basierten Produktionssystemen ist es obligatorisch, dass Engineering-Teams verschiedener Abteilungen Informationen effizient austauschen. Ein Format, das mit heterogenen Daten umgehen kann, ist das XML-basierte Datenformat AutomationML (siehe z. B. [3], [4]). Es kann wesentlich mehr Informationen enthalten als z. B. ein typisches CAD-Austauschformat wie STEP oder IGES. Um AutomationML einfach zugänglich zu machen, beinhaltet es mehrere Standards.

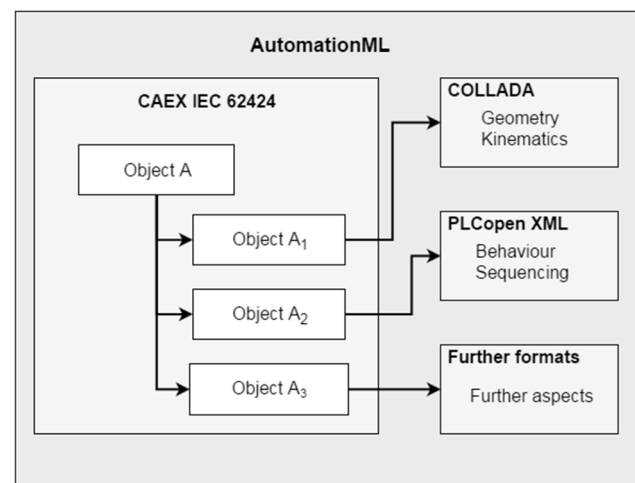


Abbildung 3: AutomationML Überblick

Die offenen Standards, die von AutomationML verwendet werden, sind in Abbildung 3 dargestellt. Die AutomationML-Datei selbst basiert auf dem CAEX-Format (IEC 62424), das nur geringfügig angereichert ist. Da es XML-basiert ist und durch die Möglichkeit, andere Dateien zu referenzieren, ist es leicht erweiterbar. Die vorhandenen Komponenten, die hierarchische Struktur sowie die Verbindung zwischen den Komponenten werden mit dem CAEX-Format beschrieben. Der COLLADA-Standard bietet die Funktionalität für die Darstellung von Geometrie. Er ist in der Lage, Geometrie sowohl als Boundary-Darstellung (typisch für CAD-Software) als auch als triangulierte Netzdarstellung zu speichern. Neben der Geometrie

rie kann COLLADA auch Informationen über die Kinematik und Physik eines Objekts sowie andere geometriebezogene Informationen enthalten. Das PLCopen XML-Format ist ebenfalls in AutomationML enthalten und macht es besonders für die Virtuelle Inbetriebnahme (VIBN) von Produktionssystemen interessant. Da es auf der IEC61131-3 basiert, ist es fähig, Programmiersprachen für SPS-Steuerungen, Embedded Controls und Industri-PCs zu speichern und zu übertragen. Diese Daten können auf Software oder Hardware in den für die virtuelle Inbetriebnahme typischerweise erforderlichen Loop-Systemen ausgewertet werden. Ebenfalls in Abbildung 3 dargestellt ist die Möglichkeit, weitere Formate einzubinden, um die AutomationML um spezielle Funktionalitäten zu erweitern.

3. AutomationML-basierte Asset-Verwaltungsschalen

Eine Verwaltungsschale begleitet ein technisches Objekt über die gesamte Dauer des Produktlebenszyklus. Dabei fallen verschiedenste Daten an. In der Konstruktionsphase sind dies überwiegend Planungsdaten u. a. 3D-CAD-Daten. Sobald das technische Objekt als Instanz genutzt wird, ändert sich auch die Art der zusätzlichen Daten. Es werden z. B. Messdaten sowie Service- und Lebensdauerdaten erzeugt. Um die persistente Erfassung dieser sehr heterogenen Daten zu ermöglichen, ist es notwendig, ein sehr flexibles System bzw. Format für die Anlagenverwaltungsschale zu wählen.

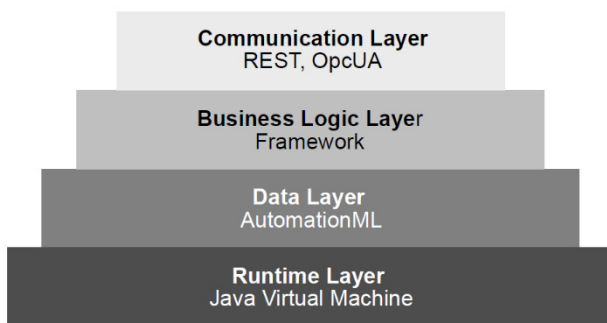


Abbildung 4: Aufbau der Implementierungsschicht

3.1 Aufbau der Implementierungsschicht

Das Referenzarchitekturmodell Industrie 4.0 gibt einen grundlegenden Überblick über die Ziele, die mit dem Modell erreicht werden sollen. Für den Großteil der Komponenten lassen sich jedoch keine Umsetzungsempfehlungen ableiten. Die Autoren treffen daher im Folgenden einige Annahmen, die als Grundlage für die spätere Umsetzung dienen.

- Laufzeitumgebung

Die softwaretechnische Umsetzung einer Verwaltungsschale kann sehr unterschiedlich sein. Zum einen ist es möglich, die Daten und die Laufzeitumgebung der Verwaltungsschalen technischer Systeme zentral in einem datenbankorientierten System abzulegen. Je nach Wahl

der Datenbank können sich jedoch Einschränkungen hinsichtlich der Art und Struktur der Daten ergeben. Eine weitere Möglichkeit ist die dezentrale Einbettung der Verwaltungsschalen z. B. direkt auf dem verwalteten technischen Objekt. Da es für beide Anwendungsszenarien plausible Anwendungsfälle gibt, sollte eine Möglichkeit gewählt werden, die beide Szenarien gleichermaßen ermöglicht.

- Datenablage

Wie bereits in der Laufzeitumgebung erwähnt, können die Daten zentral oder dezentral abgelegt werden. Insbesondere die Wahl des Datenformats, in dem die Daten zur Verfügung gestellt werden, spielt eine zentrale Rolle. Ein proprietäres Datenformat kann zu Integrationsproblemen mit Fremdsystemen führen, insbesondere aufgrund der großen Varianz der verfügbaren Softwareprodukte. Es ist daher ratsam, einen offenen Standard zu wählen, um die Nutzungsmöglichkeiten einer Verwaltungsschale nicht einzuschränken. Das gewählte Datenformat muss in der Lage sein, die bereits erwähnten heterogenen Daten, die während des Produktlebenszyklus anfallen, aufzunehmen.

- Kommunikation

Die Kommunikationsfähigkeit einer Verwaltungsschale ist elementar und sollte daher besondere Beachtung finden. Im Referenzarchitekturmodell wird an dieser Stelle der Begriff "Serviceorientierte Architektur" verwendet. Die Kommunikation auf Basis einer solchen Architektur hat sich in den letzten Jahren in verschiedenen Softwareprojekten bewährt und wird daher empfohlen. Die Autoren sind jedoch der Meinung, dass eine weitere, näher an der Maschine liegende Kommunikationsmöglichkeit die Integration der Verwaltungsschalen auf Maschinenebene erleichtern würde. Daher werden zwei Formen der Kommunikation betrachtet.

Der daraus resultierende schichtenartige Aufbau ist in Abbildung 4 dargestellt. In dieser Abbildung sind die einzelnen Schichten bereits mit Technologien belegt, die die getroffenen Annahmen erfüllen. Eine grundsätzliche Überlegung, die mit der gewählten Programmiersprache adressiert werden muss, ist die Kompatibilität mit unterschiedlichen Ausführungssystemen. Daher wurde ein Ansatz gewählt, der die Ausführung des Codes auf verschiedenen Plattformen wie Windows- oder Linux-Umgebungen ermöglicht. So wurde Java als Programmiersprache gewählt, die es aufgrund der Java Virtual Machine erlaubt, den gleichen Code auf verschiedenen Plattformen auszuführen.

3.2 Java AutomationML Framework

Das vom Industrie Konsortium AutomationML e.V. bereitgestellte Framework ist derzeit nur auf Basis der .NET Framework Programmiersprache C# verfügbar. Ein Einsatz in Java ist damit nicht möglich.

Aus den bereits genannten Gründen wird aber für den beschriebenen Ansatz ein AutomationML-Framework

benötigt, das den effektiven Einsatz von AutomationML unter Java ermöglicht. Da dieses AutomationML-Framework insbesondere für den Einsatz in Verbindung mit den Verwaltungsschalen verwendet werden soll, müssen einige zusätzliche Anforderungen erfüllt werden.

Einfache Integration zusätzlicher Lebenslaufdaten

Die Hauptaufgabe des neu erstellten Frameworks ist die Integration von zusätzlichen Daten, die während des Produktlebenszyklus anfallen. Es soll möglich sein, jede Art von zusätzlichen Daten in die AutomationML-Datei zu integrieren.

Vollständige Serialisierung und Deserialisierung

Um das Framework robuster gegen Störungen zu machen und um den Speicherbedarf zu reduzieren, müssen die Daten als AutomationML-Dateien (*.aml) im korrekten, vom Framework vorgegebenen Format, sowohl gespeichert als auch geladen werden können. Dies übernimmt ein Serialisierungs- und Deserialisierungsmechanismus. So wird Fehlern aufgrund von fehlerhaft aufgebauten Datensätzen vorgebeugt.

Toolkit für mathematische Operationen auf Basis der *FrameAttributeType*-Attribute

Positionen und Drehungen von einzelnen Komponenten können in AutomationML als *FrameAttributeType* gespeichert werden. Dieses *FrameAttributeType*-Attribut enthält die Position und Rotation eines Elements. Die Rotation wird in Euler-Winkeln gehalten, was insbesondere bei komplexen mathematischen Operationen im 3D-Raum problematisch ist. Daher werden zwei neue Klassen für die Rechenoperationen auf Basis der *FrameAttributeType*-Attribute eingeführt. Das *FramePosition*-Element enthält den Positionsanteil des *FrameAttributeType*-Attributs. Das *FrameRotation*-Element enthält den Rotationsanteil des *FrameAttributeType*-Attributs, der in eine Quaternion *FrameRotation* umgewandelt wird. Um die von AutomationML e.V., einem 2006 gegründeten Industriekonsortium, definierte Rotationsfolge nicht zu verletzen, erfolgt die Umrechnung wie in Gleichung 1 dargestellt:

$$q_x * q_y * q_z = q_{res} \quad (1)$$

Die Indizes geben die Drehung um die einzelnen Achsen an. Durch die Umrechnung der Rotationen in Quaternionen werden die erforderlichen Rechenoperationen für räumliche Berechnungen reduziert und das mathematische Problem der "Kardanverriegelung" (siehe auch [8]) wird vermieden. Die Drehung einer Position p_0 um ein gegebenes Quaternion q_n kann dann wie folgt ausgedrückt werden:

$$p_0' = q_n * p_0 \quad (2)$$

Mit diesem System lassen sich komplexe mathematische Operationen auf der Basis des *FrameAttributeType* leicht durchführen.

Integriertes Toolkit zum Erstellen und Ändern von PLCOpenXML-Daten

Um eine Verwaltungsschale und damit auch das verwaltete technische Objekt in die Lage zu versetzen, adaptiv auf veränderte Randbedingungen zu reagieren, kann es notwendig sein, das verwendete SPS-Programm anzupassen. Für die Verwaltung von SPS-Programmen ist der offene Standard PLCOpenXML in den AutomationML-Standard integriert. Um die Modifikation dieser Programme zu vereinfachen, wird ein Toolkit implementiert, das die semantisch und syntaktisch korrekte Modifikation von PLCOpenXML-Daten ermöglicht.

3.3. Asset-Verwaltungs-Shell-Framework

Wie in Abbildung 4 dargestellt, befindet sich das Administrations-Shell-Framework zwischen der Kommunikationsschicht und der Datenschicht und repräsentiert die eigentliche Geschäftslogik. Die Abbildung der Daten auf die Kommunikation ist grundsätzlich in zwei verschiedenen Formen möglich:

1. Aufteilung einer AutomationML-Struktur in einzelne Datenelemente
2. Abbildung der kompletten AutomationML-Struktur als ein einziges Datenelement

Die erste Form eignet sich besonders für Kommunikationsformen, die eine solch granulare Aufteilung erfordern wie z. B. Maschinensteuerungen. In der Regel ist dies für laufzeitvariable Daten notwendig. Die 2. Form ist z. B. für Planungsdaten geeignet, die in eine Software importiert und ggf. um weitere Planungselemente wie neue Geometrien erweitert werden sollen.

Maschine-zu-Maschine-Kommunikation

Ein Aspekt von Industrie 4.0 ist die Verlagerung von Intelligenz durch Einbettung von Steuereinheiten in Baugruppen zu eigenständigen Objekten. Dies erhöht den Bedarf an einer standardisierten Maschine-zu-Maschine (M2M) Kommunikation. Daher muss die Asset Administration Shell in der Lage sein, auf diese standardisierte Weise zu kommunizieren. In der jüngeren Vergangenheit hat sich der OPC UA-Standard als valider Konkurrent für eine zukünftige standardisierte M2M Kommunikation erwiesen. Daher wurde dieser Standard in das Smart Asset Administration Shell Framework implementiert.

Mensch-Maschine-Kommunikation

Da auch in hochautomatisierten Prozessen der Einfluss eines Bedieners notwendig ist, muss die Mensch-Maschine-Kommunikation in einer vergleichbaren Qualität wie die Maschine-Maschine-Kommunikation sein. Um eine Mensch-Maschine-Kommunikation zu ermöglichen, gibt es verschiedene Möglichkeiten. Eine typische Möglichkeit ist heutzutage die Einbettung eines Displays in das technische System wie z. B. als Bedienfeld an einer Werkzeugmaschine. Während dies für Maschinen mit einer einzigen Steuerung die wahrscheinlich beste Option ist, kann sie bei Maschinen, die aus Dutzenden von Steuerungen bestehen, kaum angewendet werden. Daher

wird die Mensch-Maschine-Kommunikation vergleichbar einer serviceorientierten Architektur realisiert, um einem Benutzer die einfache Interaktion mit beliebigen Steuergeräten oder Asset Administration Shells zu ermöglichen.

4. VIRTUELLE INBETRIEBNAHME

Um mit dem Digitalen Zwilling eines Produktionssystems zu arbeiten und die verschiedenen Parameterstudien zur Optimierung des Energieverbrauchs durchzuführen, ist es erforderlich, eine Umgebung für die Virtuelle Inbetriebnahme bereitzustellen. Diese erfordert eine akkurate Simulationsumgebung und Schnittstellen zu weiterer kommerzieller Steuerungssoftware, damit die Programme, welche später auf der realen Maschine laufen sollen, getestet und ggf. modifiziert werden können.

4.1 Unity zur Visualisierung eines Digitalen Zwillings mit AutomationML

Unity ist eine Plattform zur Erstellung von 3D-Inhalten und -Anwendungen. Ursprünglich wurde Unity als GameEngine entwickelt. Aber aufgrund der absolut offenen Software, die Unity bietet, wird die Software nun für alle Arten der Erstellung von 3D-Inhalten verwendet, auch für die virtuelle Inbetriebnahme. Da Unity auf C# basiert, kann es mit einer großen Anzahl von APIs und SDKs für alle Arten von Anwendungsfällen arbeiten. Daher ist Unity ein ausgezeichnetes Programm für die Visualisierung von Objekten des Digitalen Zwillings von Produktionssystemen.

Eine eigens entwickelte Import-Erweiterung für AML-Dateien ermöglicht es dem Benutzer, die generierte AML-Datei zur Laufzeit in die Simulationsumgebung zu laden. Die vom AutomationML-Container bereitgestellten *Collada*-Dateien werden voll automatisiert in der Szene platziert und jedem Bauteil werden seine Position und seine funktionalen Eigenschaften durch die Attribute in der AutomationML-Datei zugewiesen.

LogicFacets ermöglichen darüber hinaus die Animation des Produktionsprozesses ohne Anbindung an eine spezielle Steuerungssoftware. Dies ist sehr hilfreich für eine erste Überprüfung des Systems, des Materialflusses oder für einfache Präsentationsfälle.

Eine akkurate virtuelle Inbetriebnahme-Umgebung benötigt jedoch Schnittstellen zu kommerzieller Steuerungssoftware wie z. B. das Programm TwinCat der Firma Beckhoff, damit die eigentliche Produktionssoftware an dem Digitalen Zwilling getestet werden kann. Dafür gibt es Schnittstellen für industrietaugliche Steuerungssoftware. Die Schnittstelle für TwinCat basiert auf dem TwinCat ADS Framework, mit dem TwinCat mit allen Arten von C#-Applikationen verbunden werden kann.

Wie in Abbildung 5 dargestellt, verbinden sich die Schnittstellen mit den generierten Inputs und Outputs

des Digitalen Zwillings und übernehmen dann die logische Steuerung des Systems. So erhält die TwinCat-Software Inputs von Ereignissen in der Unity-Umgebung z. B. ausgelöste Sensoren wie bei Lichtschranken oder wenn eine ausgeführte Animation beendet ist. Die Logiksteuerung stößt dann weitere Ereignisse an, indem sie die Ausgänge der SPS aktiviert. Auf diese Weise können die generierten SPS-Programme innerhalb der VIBN-Umgebung getestet, verifiziert oder bearbeitet werden.

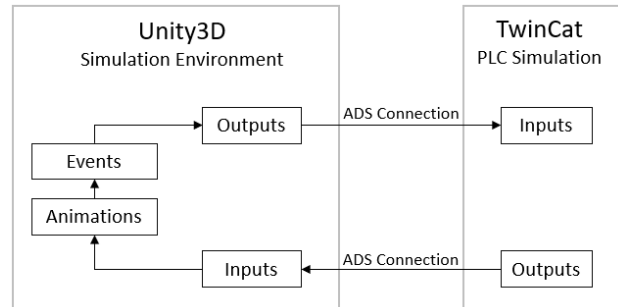


Abbildung 5: Schematische Darstellung der Unity-TwinCat Kommunikation

4.2 ISG-Virtuos für eine physikalisch akkurate virtuelle Inbetriebnahme

Virtuos ist eine Echtzeit-Simulationssoftware für physikalisch exakt abgebildete digitale Zwillinge, die von der Firma ISG mit Sitz in Stuttgart angeboten wird. Virtuos stellt eine sehr große Bibliothek mathematischer Modelle zur Verfügung, die in der Lage sind, ein genaues Verhalten aller Arten von verwendeten Aktoren eines Produktionssystems (z.B. hydraulische/pneumatische Zylinder, Servomotoren, Getriebe usw.) darzustellen. Hier kann durch Beschleunigungen, Geschwindigkeiten etc. auf den Energieverbrauch der Maschine geschlossen werden. So kann durch eine Variation der Parameter (Beschleunigungen, Geschwindigkeiten etc.) ein optimales Verhältnis zwischen Energieverbrauch und Produktivität erreicht werden.

Um dieses mathematische Modell innerhalb der Unity VIBN-Umgebung nutzen zu können, wurde ein eigenes C++ UDP-Interface entwickelt. Das Interface basiert auf dem Publisher-Subscriber-Prinzip. Für jede Aktion, die in Virtuos berechnet werden muss, gibt es ein spezielles Attribut in der AutomationML-Datei.

Wie in Abbildung 5 dargestellt, werden die Berechnungen der einzelnen Vorgänge z. B. Hydraulikzylinder vorwärts durch eine Schreibenforderung aus der Unity-Umgebung ausgelöst. Die Berechnung des Fortschritts wird in dem VIBN-Programm Virtuos berechnet und die Unity-Umgebung abonniert diesen Vorgang, wodurch die Simulationsumgebung die aktuelle Position der Bewegung über die UDP-Schnittstelle aktualisieren kann. Die Updates werden in jedem Frame angezeigt, was bedeutet, dass alle Bewegungen mit einer Framerate von bis zu 30fps aktualisiert werden, so dass eine sehr flüssige Visualisierung erreicht werden kann. Auf diese

Weise kann die Simulation des Produktionsprozesses sehr detailliert und genau erfolgen.

Virtuos unterstützt nicht die Verwendung von AutomationML-Dateien. Es nutzt das Dateiformat *.ecf für den Import von Virtuos Modellinhalten in einer Projektdatei, was im Anwendungsfall der automatisch erstellten Simulationsmodelle einige Herausforderungen mit sich bringt.

Die Lösung ist die separate Erstellung einer passenden *.ecf-Datei, die als komplettes Projekt in Virtuos geladen werden kann. Diese Datei wird zusätzlich zur AutomationML-Datei von der graphenbasierten Entwurfssprache automatisch erstellt.

4.3 Weitere Schnittstellen für realistische Simulationen

Mit den beschriebenen Erweiterungen in Unity gibt es damit eine Lösung für die Visualisierung, für genaue physikalische Berechnungen sowie für die Anbindung an eine SPS-Steuerungssoftware. Um die VIBN-Umgebung noch vollständiger zu machen, wird eine Schnittstelle für eine kommerzielle Robotersteuerungssoftware benötigt.

Die Software KukaOffice Light von KUKA ist ein Programm für die Roboterprogrammierung. Es wurde eine UDP-Schnittstelle entwickelt, die eine Kommunikation zwischen der Simulationsumgebung und der Kuka-Steuerungssoftware ermöglicht. Die Simulation ermöglicht eine Optimierung der Pfade der Roboter. So kann beispielsweise die Verfahrgeschwindigkeit eines Handlingsroboters an die Taktzeit der Produktionsmaschine angepasst werden, um keine Energie für unnötig große Beschleunigungen und hohe Verfahrgeschwindigkeiten aufzuwenden.

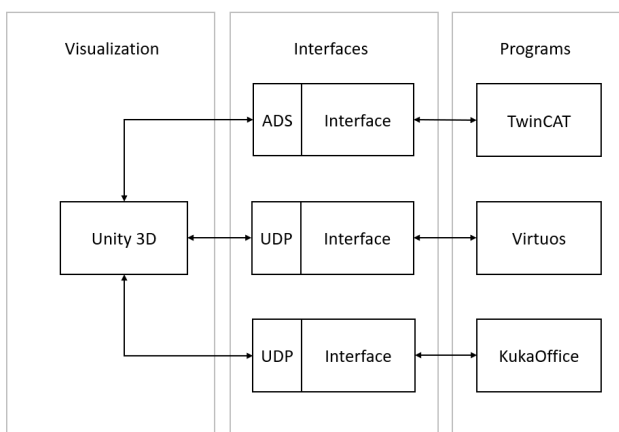


Abbildung 6: Gesamtüberblick über die VIBN-Umgebung

Wie in Abbildung 6 zu sehen ist, ermöglicht die offene Architektur von Unity die Erstellung einer kompletten VIBN-Umgebung mit Schnittstellen zu allen Arten von kommerzieller Software, die in der Lage ist, viele Anforderungen an die Simulation als Digitale Zwillinge von Produktionssystemen zu erfüllen und mit dem Automa-

tionML-Daten-Container zu arbeiten. Weitere Schnittstellen zu anderen Arten von Software für kommerzielle Anwendungen können noch hinzugefügt werden.

5. FAZIT UND WEITERE FORSCHUNG

Digitale Zwillinge verfügen über ein sehr großes Potential, zur Optimierung von Produktionssystemen und -anlagen sowohl im Hinblick auf die Produktivität sowie auch im Hinblick auf Energieeinsparungen. Die Akzeptanz von Industrie 4.0-Komponenten und dem Referenzarchitekturmodell Industrie 4.0 wird stark davon abhängen, ob die Hersteller der Systeme eine gemeinsame Datenbasis finden. Die Kombination aus AutomationML und dem Referenzarchitekturmodell Industrie 4.0 könnte eine solche datentechnische Basis darstellen und damit zur verbesserten Interoperabilität dieser Systeme beitragen. Solche Modelle können durch eine Produktionspipeline auf Basis graphenbasierter Entwurfssprachen, wie von Kiesel et al. [5] und Beisheim et al. [6], [7] beschrieben, bereits heute automatisch generiert werden, was eine höhere Anzahl von Simulationen zur Funktionsvalidierung und Parameteroptimierung ermöglicht.

Die neu geschaffene VIBN-Umgebung mit den Erweiterungen in Unity liefert ein recht genaues Modell inkl. des Verhaltens von Produktionssystemen. Es ist jedoch notwendig, den Abbildungsgrad in weiteren Projekten zu validieren. Eine weitere Aufgabe ist die kontinuierliche Verbesserung der erstellten Software und Schnittstellen, da an noch einige manuelle Arbeitsschritte z. B. die Verknüpfung der SPS mit der VIBN-Software notwendig sind sowie um das gesamte System hochzufahren und alle Software zu verbinden. Speziell bezogen auf den Energieverbrauch ist eine Erweiterung um Funktionen geplant, die standardmäßig während jeder Simulation Daten hinsichtlich Energieverbrauch sammeln und auswerten, um zukünftige Produktionsmaschinen energieeffizienter zu gestalten.

Danksagung

Das Forschungsprojekt 'Digitaler Produktlebenszyklus (DIP)' wurde durch einen Zuschuss des Europäischen Fonds für regionale Entwicklung und des Ministeriums für Wissenschaft, Forschung und Kunst Baden-Württemberg, Deutschland, unterstützt. Weitere Informationen: <http://dip.reutlingen-university.de>.

Literaturverzeichnis

- [1] DIN SPEC 91345, Referenzarchitekturmodell Industrie 4.0 (RAMI4.0), Beuth Verlag, 2016.
- [2] AutomationML e.V., Standardized data exchange in the engineering process of production systems, Accessed: 15.01.2020. [Online]. Available: <https://www.automationml.org/o.red/uploads/dateien/1544706233-automationml.pdf>

- [3] R. Draht, A. Lüder, J. Peschke and L. Hundt, AutomationML - the glue for seamless automation engineering, IEEE International Conference on Emerging Technologies and Factory Automation, Hamburg, September 15-18, 2008, IEEE Xplore, 2008
- [4] A. Lüder, L. Hundt and A. Keibel, Description of manufacturing processes using AutomationML, IEEE 15th Conference on Emerging Technologies & Factory Automation (ETFA 2010), Bilbao, September 13-16, 2010, IEEE Xplore, 2010
- [5] M. Kiesel, P. Klimant, N. Beisheim, S. Rudolph and M. Putz, Using Graph-based Design Languages to Enhance the Creation of Virtual Commissioning Models, Procedia CIRP, Vol. 60, 2017, pp. 279-283.
- [6] N. Beisheim, M. Kiesel and S. Rudolph, Digital manufacturing and virtual commissioning of Intelligent Factories and Industry 4.0 systems using graph-based design languages, Proceedings of the 25th ISPE Inc. International Conference on Transdisciplinary Engineering, Modena, July 3-6, 2018. Vol. 7, IOS Press, 2018, pp. 93-102
- [7] Beisheim, N.; Kiesel, M.; Linde, M.; Ott, T., Using AutomationML and graph-based design languages for automatic generation of digital twins of cyber-physical systems, Proceedings of the 27th ISPE Inc. International Conference on Transdisciplinary Engineering, Warschau/online, July 1-3, 2020., IOS Press Ebooks; ISBN 978-1-64368-110-8

Sprachsteuerung für einen industriellen Leichtbauroboter

Thomas Wachsmuth, Christian Thormann, Alexander Winkler

Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

In diesem Beitrag wird die Sprachsteuerung eines Leichtbauroboters vorgestellt. Im Alltag findet man viele Beispiele zur Sprachsteuerung. Darum sollte dieser Ansatz auch in der industriellen Automatisierung etabliert werden, wie z.B. in der Robotik. Leichtbauroboter eignen sich aufgrund ihrer intuitiven Programmierung sehr gut für solche Aufgaben. Zudem enthalten sie üblicherweise Sicherheitsfunktionen in Form von Geschwindigkeits- oder Kraftbeschränkungen was sie auch in Kombination mit der Spracherkennung sicherer macht. Für die Implementierung der Roboter-Sprachsteuerung musste eine geeignete Spracherkennungssoftware ausgewählt werden. Hierfür wurden verschiedene Systeme verglichen. Bei der Sprachsteuerung des Roboters wurde zwischen dem Teach-In mit Sprachbefehlen und der Bewegungssteuerung mittels Sprache unterschieden. Beide Konzepte wurden an einem Demonstrator umgesetzt und die Algorithmen in Versuchen verifiziert. Anschließend wurden das Teach-In mit Sprachbefehlen mit dem per Handbediengerät in Bezug auf die Arbeitszeit verglichen.

1. Einführung

Im Allgemeinen besteht die Programmierung industrieller Roboter aus einer Kombination von Teach-In und textueller Programmierung. Teach-In bedeutet, dass der Bediener den Manipulator zu allen Positionen bewegt, die für die Aufgabe notwendig sind. Dabei kann der Manipulator z.B. über die Tastatur oder per Joystick gesteuert. Nach dem Teach-In muss die Programmierung textbasiert vervollständigt werden. Hierfür gibt es viele verschiedenen Roboterprogrammiersprachen. Ein Überblick kann in [1] gefunden werden.

Seit einigen Jahren geht der Trend hin zu sogenannten Leichtbaurobotern. Das bedeutet, dass das Verhältnis von Traglast zu Eigenmasse nahezu 1:1 beträgt, im Gegensatz zu konventionellen Industrierobotern mit nahezu 1:10. Leichtbauroboter wurden ursprünglich für die Raumfahrt entwickelt. Allerdings führte die geringere Masse auch zu neuen Anwendungsgebieten, wie z.B. der Mensch-Roboter Interaktion, welche zunehmend an Bedeutung gewinnt [2]. Aus diesem Grund können gegenwärtig viele Leichtbauroboter von unterschiedlichen Herstellern am Markt gefunden werden.

Häufig können diese Leichtbauroboter zum komfortablen und intuitiven Teach-In vom Bediener handgeführt werden. Für diesen Zweck werden oftmals die Motorströme analysiert oder es kommen Gelenkmomentsensoren zum Einsatz.

Eine weit verbreitete Gruppe von Leichtbaurobotern ist die UR-Serie von Universal Robots. Diese verfügt über Energie- und Kraftbegrenzungsfunktionen, welche den Betrieb des Roboters auch ohne trennenden Schutzeinrichtungen erlaubt. Wie bereits erwähnt, kann das Teach-In auch per Handführen erfolgen. Danach muss man das Programm jedoch noch vervollständigen, wobei die entsprechenden Anweisungen auf dem Touchscreen des Handbediengerätes ausgewählt werden [3]. Nach unserem Verständnis passt diese Vorgehensweise jedoch nicht zum Teach-In per Handführen. Außerdem

sollte der Roboter auch im Automatikmodus nicht ausschließlich mittels Kontrollpanel bedient werden können. Bewegt sich der Roboter z.B. im kraftgeregelten Modus [4] und der Bediener greift in den Prozess ein, so wird der Roboter die Kraft weiter erhöhen, wodurch auch der Bediener mehr Kraft aufwendet. Dies ist der Beginn einer Interaktion, welche der Roboter gewinnen wird. In einem solchen Fall könnte ein instinktiver „Stopp“-Ruf des Bedieners den Roboter unter Kontrolle bringen.

Aus den genannten Gründen scheint es notwendig zu sein, alternative Programmier- und Bedienkonzepte zu betrachten. Dazu existiert eine Vielzahl von Arbeiten. Dabei wird z.T. zwischen Experten und Amateuren unterschieden. Im Fall der Experten, würde es hilfreich sein, eine genormte Programmiersprache zu haben. Jedoch ist der Versuch einer Normung, mittels der standardisierten Sprache IRL [5], in der Vergangenheit gescheitert. Für den Fall das auch Nicht-Experten in die Roboterprogrammierung und Bedienung einbezogen werden sollen, spielt die Intuitivität eine wichtige Rolle. Praktische Ansätze können den Roboter durch Berührung, Gesten, Sprache oder 3D-Tracker steuern.

Dieser Beitrag ist im Weiteren wie folgt aufgebaut. Das folgende Kapitel gibt einen Überblick über den aktuellen Stand der Technik in Bezug auf Spracherkennungssoftware. In Abschnitt 3 stellen wir alternative Ansätze zum Teach-In und zum Bedienen eines UR3e-Leichtbauroboters über Sprache vor. Das neue Teach-In-Konzept wird in Abschnitt 4 untersucht. Zu diesem Zweck wird die benötigte Zeit für das Einlernen mit der Dauer herkömmlicher Programmierkonzepte verglichen. Eine Zusammenfassung der Ergebnisse wird in Abschnitt 5 gegeben.

2. Systeme zur Spracherkennung

Eine Art der intuitiven Programmierung und Steuerung ist per Sprache. Sprachsteuerung hat sich schon im Alltag etabliert. Ebenso auf dem Feld der Roboter, wo man bereits diverse Ansätze finden kann. Allerdings sind die

meisten davon für mobile Roboter vorgesehen. Im Gegensatz zu anderen Konzepten, arbeitet die Sprachsteuerung ohne spezielle Hardware. Nur ein Mikrofon wird benötigt, welches in den meisten modernen Kommunikationsgeräten integriert ist, wie im Tablet PC, Laptop oder Handy. Die Sprachsteuerung arbeitet auch unter ungünstigen Lichtverhältnissen und in der Dunkelheit, was für spezielle Befehle, wie den Not-Halt, erforderlich ist. Ein weiterer Vorteil liegt darin, dass der Bediener währenddessen auch etwas Anderes mit seinen Händen machen kann. So dürfte es möglich sein den Roboterarm per Hand zu führen und die Teach-In-Positionen per Sprachbefehl zu speichern.

Zur Realisierung einer Robotersteuerung per Sprache wurden die aktuell relevanten Sprachanalysesysteme recherchiert und in Tabelle 1 aufgeführt.

| Sprachanalyse-Software | Eigenschaften | | | | |
|--------------------------------|---------------|---------|-----------------------|---------------------|--------------------------|
| | Online | Offline | kostenfreier Download | kostenfreie Nutzung | Entwickler-Schnittstelle |
| Apps (Microsoft, Android, MAC) | X | X | X | X | |
| Amazon Web Service Lex | X | | | X | X |
| CMU Spinx | | X | X | X | X |
| Google Cloud Speech API | X | | | X | X |
| Houndify API | X | | | | X |
| IBM Watson | X | | | | |
| Julius (bis 2014) | | X | X | X | X |
| Nuance | | X | | | |
| Mozilla CommonVoice | | X | X | X | |
| Mozilla DeepSpeech | | X | X | X | X |
| MS Cortana | X | | | X* | X* |
| MS Speech Assistant | | X* | X* | X* | X* |
| Picovoice | | X | X | | X |
| Porcupine (HotWordDetection) | | X | X | | X |
| Snips Flow | X | X | X | | |
| Snowboy (HotWordDetection) | X | X | X | | X |
| wit.ai | X | | X | X | X |

Gewünschte Eigenschaft
 X* Abhängig von Windows-Lizenz

Tabelle 1: Vergleich aktueller Spracherkennungssoftware

Dabei kann grundsätzlich zwischen Onlinesprachanalyse, deren Verarbeitung in der Cloud stattfindet, und der Offlinesprachanalyse unterschieden werden. Es gibt auch hybride Varianten, welche zum Anlernen den Sprachalgorithmus und die Rechenkapazitäten der Cloud nutzen, ihre eigentliche Arbeit jedoch offline ausführen können. Die bedeutendsten drei Sprachanalyse-systeme sind die von Google, Microsoft und IBM. Sie bieten, durch starke Rechner-server, sehr schnelle Analysezeiten mit geringen Fehlerraten. Dieses Cloudcomputing setzt allerdings eine Onlineverbindung der Maschine voraus. Weiterhin ist die Nutzung dieser künstlichen Intelligenz kostenpflichtig und wird je nach Anzahl der transkribierten Wörter oder der eingewählten Zeit abgerechnet.

Im Zuge der vierten industriellen Revolution hält das Internet of Things derzeit Einzug in die Produktion und die Onlinekommunikation der Maschinen nimmt zu. Es gibt jedoch einige Aspekte, wie Betriebskosten, Produktionsgeheimnisse oder eine sensible IT-Sicherheitsinfrastruktur, die für einen cloudbasierten Maschinenbetrieb als kritisch anzusehen sind. Hier bietet eine Offline-Variante Vorteile.

Die Offlinesprachanalyse ist jedoch weniger komfortabel als die cloudbasierte Variante, da ein lokaler Algorithmus mit beschränktem Datensatz erst die Stimme des Bedieners erlernen muss. Die intensive Nutzung der Offline-Spracherkennung verringert diesen Nachteil, da sich der Algorithmus kontinuierlich verbessert, die Produktionsdaten im Unternehmen verbleiben und kostenlos sind. Wenn allerdings die Offlinesprachsteuerung für dieses Projekt geeignet ist, dann sollte auch die Online-sprachsteuerung funktionieren. Deswegen wird im vorliegenden Szenario eine Offline-Variante verwendet.

Fast alle Offline-Sprachanalyse-systeme sind kostenfrei benutzbar, aber nur etwa 60 % der betrachteten Systeme sind auch kostenlos zu erwerben und nur etwa 1/3 bieten eine Entwicklerschnittstelle, um sie in eigenen Anwendungen zu implementieren. Aufgrund der weiten Verbreitung des Betriebssystems Microsoft in Unternehmen und den zugehörigen Entwicklungspaketen wurde der Microsoft Speech Assistent zum Aufbau eines Demonstrators ausgewählt.

3. Funktionsmuster einer Sprachsteuerung für Industrielle Leichtbau-Roboter

Wie bereits erwähnt, verwenden wir den Microsoft Sprachassistenten für die Sprachanalyse. Dieser ist in jedem aktuellen Windows Betriebssystem vorhanden und vielfältig verwendbar. Die Spracheingabe erfolgt über ein einfaches USB-Mikrofon. Der Sprachassistent des Betriebssystems stellt die offline analysierten Wörter den verschiedenen Anwendungen auf dem PC zur weiteren Verarbeitung zur Verfügung. Je nach Sprachqualität muss der Offline-Algorithmus vom Benutzer trainiert werden. Wenn das Betriebssystem neue Sprachdaten empfängt, sendet es diese Daten in unserem Fall an MATLAB. Dies werdet die Daten aus und sendet entsprechende Befehlsfolgen an die Steuerung des Roboters. Die Kommunikation zwischen PC und Robotersteuerung wurde mittels Ethernet hergestellt.

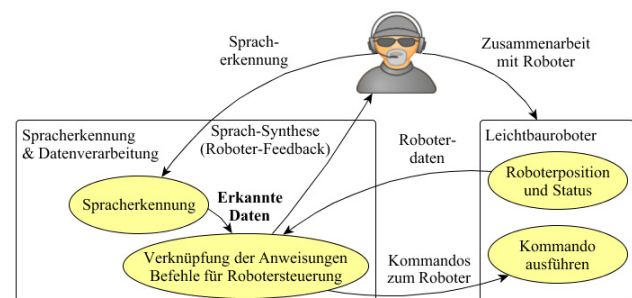


Abbildung 1: Konzept der Sprachsteuerung

Die Kommunikationsstruktur ist in der Abbildung 1 zu sehen. Für die Interaktion mit dem Roboter wird zwischen sprachgesteuerten Teach-In und sprachgesteuerten Bewegungen des Roboters unterschieden.

3.1. Teach-In mittels Sprachbefehl

Während des Teach-In, wird der Roboter per Hand zur Teach-In-Position geführt. Wenn diese erreicht ist, sagt der Programmierer den Namen dieser Position. Dieser wird durch die Sprachanalyse übersetzt und als Variablenname gespeichert. Diese Variable kann nun um Eigenschaften, wie Bewegungsart, Geschwindigkeit oder Beschleunigung erweitert werden. Gleichzeitig werden die Positionsdaten des Roboters ausgelesen und im genannten Variablennamen gespeichert. Dabei entsteht eine Datenstruktur, die für beliebig viele Teach-In-Punkte erweitert werden kann. Die so angelernten Positionen können dem Roboter offline vor Programmstart übermittelt werden oder die Bahn auch online kontinuierlich beeinflussen. Die Abbildung 2 zeigt den Programmablauf.

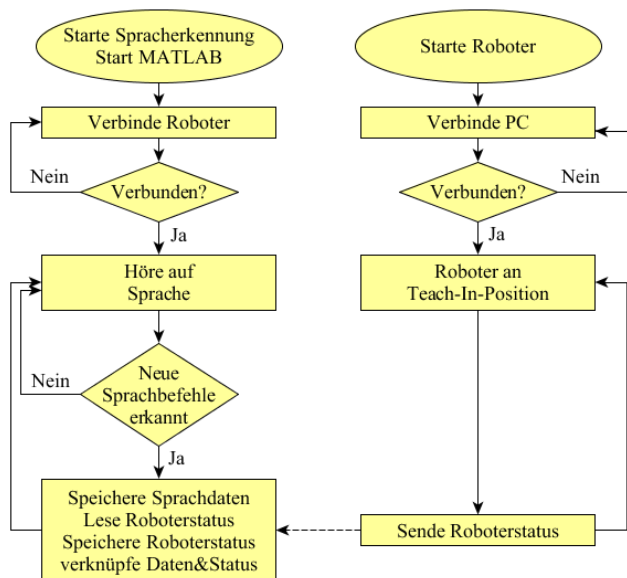


Abbildung 2: Ablaufdiagramm Teach-In mittels Sprache

3.2. Steuerung des Roboters mittels Sprache

Im Weiteren soll der Roboter während des Betriebs durch sprachliche Interaktion in bestimmte Steuerungszustände versetzt werden. Es soll z.B. möglich sein, ihn im laufenden Betrieb durch die Wörter „Stopp“ oder „Halt“ in einen Sicherheitszustand zu versetzen. Außerdem soll die Bewegung des Roboters durch gesprochene Befehle gesteuert werden. Dazu analysiert das Betriebssystem fortwährend die aufgezeichneten Umgebungsgeräusche. Sobald das Signalwort identifiziert wird, sendet der PC einen entsprechenden Befehl an die Robotersteuerung. Dadurch wird die Zusammenarbeit zwischen Mensch und Roboter wesentlich erleichtert und die Sicherheit insbesondere bei der kollaborativen Mensch-Roboter Interaktion erhöht. Die Abbildung 3 zeigt das Funktionsprinzip der Sprachsteuerung.

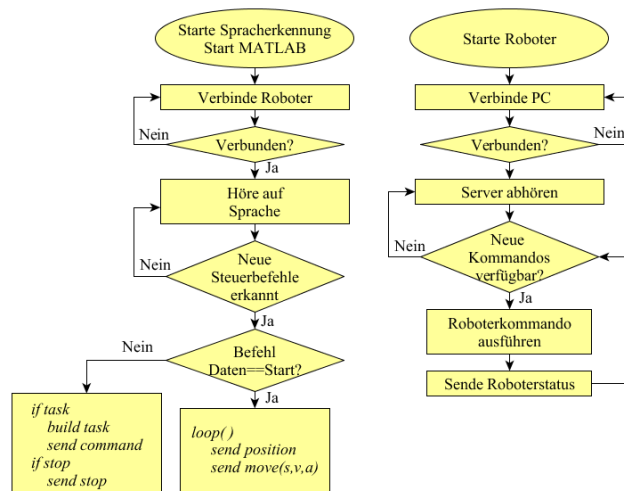


Abbildung 3: Ablaufdiagramm der Sprachsteuerung

4. Vergleich zwischen neuem und konventionellem Teach-In

Universal Robots bewirbt seine Roboter mit einem einfachen und schnellen Programmierverfahren. Tatsächlich erleichtert der vorgefertigte Befehlssatz die Programmierung, sodass diese sehr schnell erfolgen kann. Unsere Absicht war es, die vorgefertigten *URScript*-Befehle mit dem Teach-In per Sprache zu vergleichen. Der aufgebaute Demonstrator ist in Abbildung 4 zu sehen. Dieser besteht aus dem Roboter UR3e mit einem elektrischen Greifer. Dieser hält ein Werkzeug in Form eines Stiftes. Die Bahnpunkte sind auf einem kleinen Tisch markiert.

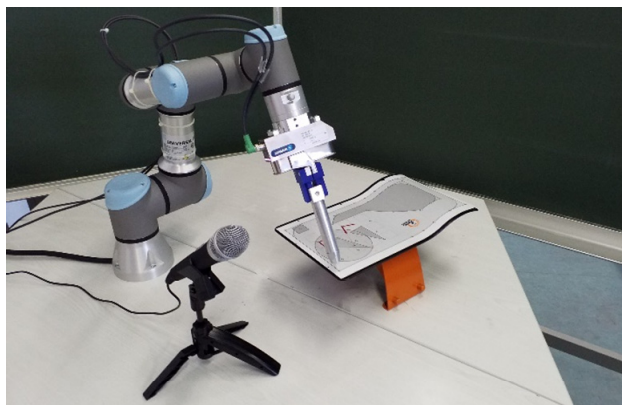


Abbildung 4: Versuchsaufbau zum Vergleichen der Teach-In-Konzepte

Die Bewegung des Roboters zu den Punkten hängt von den Fähigkeiten des Bedieners ab und wird im Vergleich nicht betrachtet. Es wird nur die Programmiermethode berücksichtigt. Um beide Teach-In-Verfahren zu vergleichen, wurden je 10 Punkte aufgenommen, zuerst über herkömmliches Teach-In und dann über Sprach-Teach-In. Des Weiteren wurde das Teach-In mit einem anderen Roboter vom Typ KUKA KR6 Agilus durchgeführt. Beim konventionelle Teach-In wurden bei beiden Robotern jeweils die vorgefertigten Funktionen mit Standardeinstellungen verwendet, um die Teach-In-Zeit zu minimieren. Dies beinhaltet auch das automatische Beschriften und

Inkrementieren der Variablen. Die Abbildungen 5 und 6 zeigen die Bedienoberflächen der Roboter. Das Teach-In per Sprache erzeugt dieselben Bewegungseinstellungen, jedoch mit individuellen Beschriftungen.

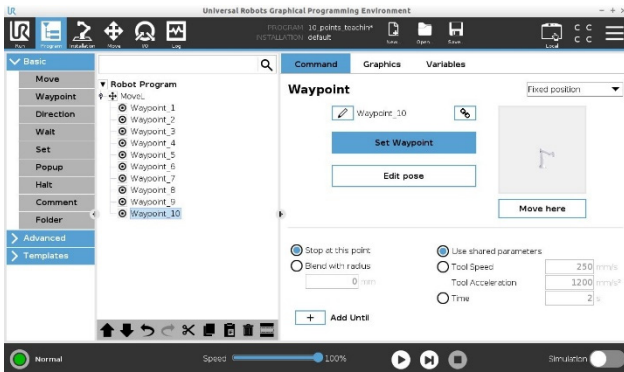


Abbildung 5: Bedienoberfläche des UR3e-Panels

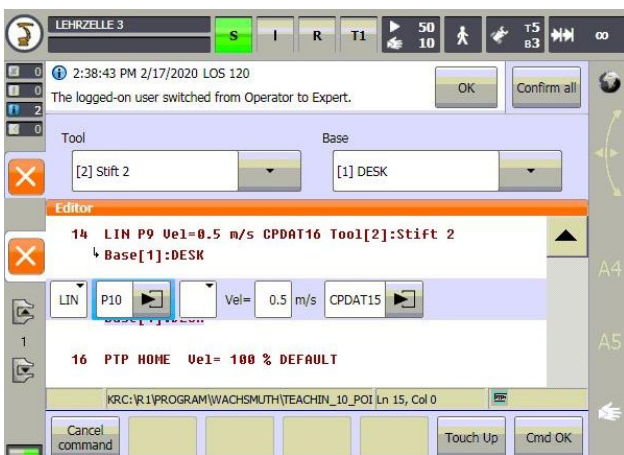


Abbildung 6: Ausschnitt der Oberfläche des KUKA-Bedienfeldes

| Position | Universal Robots UR3e | | KUKA AGILUS 6 | |
|----------|-----------------------|----------------------|----------------------------|----------------------------|
| | Sprach-Teach-In total | Teach-In Bedienpanel | Teach-In Bedienpanel total | Teach-In Bedienpanel total |
| 1 | 9,07 | 5,57 | 10,35 | 15,60 |
| 2 | 8,57 | 5,07 | 8,25 | 13,67 |
| 3 | 6,94 | 3,44 | 8,65 | 12,47 |
| 4 | 7,86 | 4,36 | 5,83 | 26,70 |
| 5 | 6,94 | 3,44 | 7,88 | 12,56 |
| 6 | 5,80 | 2,30 | 6,22 | 9,70 |
| 7 | 6,02 | 2,52 | 8,17 | 10,27 |
| 8 | 6,40 | 2,90 | 5,56 | 10,80 |
| 9 | 5,82 | 2,32 | 7,53 | 9,36 |
| 10 | 8,50 | 5,00 | 7,84 | 11,38 |

Tabelle 2: Vergleich der verschiedenen Teach-In-Konzepte

Die Tabelle 2 zeigt die angelernten Positionen und die dazugehörigen Teach-In-Zeiten. Wie zu erkennen ist, benötigt das Einlernen beim KUKA die längste Zeit. Möglicherweise hängt dies von der Empfindlichkeit und Latenz des Bedienfeldes ab. Das konventionelle Tech-In des UR3e ist dahingegen fast doppelt so schnell. Die Messung des UR3e-Sprach-Teach-In zeigt nahezu die gleichen Werte wie das herkömmliche UR3e-Teach-In. Die Datenverarbeitung unseres Systems beinhaltet jedoch eine Latenz von 3,5 Sekunden ohne Spracherkennung. Die an die Verarbeitungslatenz angepasste Teach-In-Zeit ist in der Spalte „total“ zu sehen und zeigt, dass

das Teach-In per Sprache genau betrachtet doppelt so schnell ist wie das herkömmliche Einlernen. Im Vergleich zum verwendeten konventionellen Industrieroboter wurde sogar nur ein Viertel der Zeit benötigt. Ein weiterer Vorteil würden sich bei der Verwendung von aussagekräftigeren und somit komplexeren Positionsnamen ergeben, indem z.B. Schreibfehler vermieden werden.

5. Fazit

In diesem Beitrag wurde ein Konzept zur intuitiven Bedienung eines Leichtbauroboters per Sprache vorgestellt. Es besteht aus Sprach-Teach-In und sprachbasierter Bewegungssteuerung. Für die Spracherkennung wurde der Microsoft Sprachassistent gewählt. Die Offline-Sprachsteuerung erfordert etwas Zeitaufwand, um den Algorithmus zu trainieren. Während der Tests wurde festgestellt, dass dabei einige Probleme auftreten, z.B. Kollisionen mit anderen Programmen wie Mozilla Firefox. Hinzu kommt eine hohe Auslastung des Betriebssystems. Daher sollten zukünftig alternative Offline-Algorithmen in Betracht gezogen werden.

Die Untersuchung der Sprachsteuerung für industrielle Leichtbauroboter hat einige Vorteile für die Zusammenarbeit zwischen Mensch und Roboter gezeigt. Wie erwartet war das Sprach-Teach-In schneller als die klassische Methode. Darüber hinaus können Sicherheit und Produktivität durch diese Schnittstelle erhöht werden. Es ist möglich, den Roboter mittels eines Ausrufs in einer riskanten Situation anzuhalten. Ebenso kann der Roboter intuitiver gesteuert und bewegt werden. Es ist auch denkbar, den Roboter so in gefährlichen Umgebungen fernzusteuern.

Literaturverzeichnis

- [1] S. Yang, X. Mao, B. Ge, and S. Yang, *The roadmap and challenges of robot programming languages*, in Proc. of IEEE Int. Conf. on Systems, Man, and Cybernetics, vol. 19, 2015, pp. 328–333.
- [2] B. Finkemeyer, *Towards safe human-robot collaboration*, in Proc. of IEEE Int. Conf. on Methods and Models in Automation and Robotics (MMAR), 2017, pp. 883–888.
- [3] Universal Robots A/S, *PolyScope Manual*, 2018.
- [4] Universal Robots A/S, *The URScript Programming Language*, 2018.
- [5] Deutsches Institut für Normung, *DIN 66312, Industrial Robotic Language (IRL)*, 1996.

Danksagung

Diese Maßnahme wird mitfinanziert mit Steuermitteln auf Grundlage des von den Abgeordneten des Sächsischen Landtags beschlossenen Haushaltes.

Einsatzszenarien für Electronic Shelf Label in der Produktion

Leif Goldhahn, Sebastian Roch

Hochschule Mittweida, Fakultät Ingenieurwissenschaften,

InnArbeit - Zentrum für innovative Arbeitsplanung und Arbeitswissenschaft,

Technikumplatz 17, 09648 Mittweida

Auf dem Weg zur weiteren Digitalisierung der Produktionsprozesse hilft die gezielte Adaption von Digitalisierungstechniken aus anderen Branchen und dem Consumer-Bereich. Im Fertigungsbereich stehen die Unternehmen vor der Herausforderung, bisher auf Papier geführte Daten und Dokumente in elektronischer Weise darzustellen. Hierfür bieten Electronic Shelf Label (ESL) eine sinnvolle Ergänzung zu klassischen Displays. Diese elektronischen Etiketten, auch als „ePaper“ bekannt, eignen sich insbesondere für die dynamische Anzeige von Daten mit niedriger Änderungsfrequenz. Auch bei hellem Tageslicht sind sie ohne Stromverbrauch sehr gut lesbar. Sie benötigen keine eigene Rechenleistung und erfordern Strom nur während der funkgestützten Übertragung neuer Datensätze.

Im Beitrag werden industrielle Anwendungsszenarien für ESL zur Unterstützung der Logistik und der manuellen Montage gezeigt. Ausgehend von Applikationen zur Beschriftung von Lagersystemen und Behältern in der Materialbereitstellung, wird der Einsatz als digitale Auftragsbegleitkarte behandelt. Des Weiteren werden Anwendungen in der Montage sowie der Arbeitsplatz- und Maschinenbeschriftung beschrieben. Der Einsatz der ESL ist sowohl in der auftragsbezogenen Kleinserienfertigung als auch bei variantenreicher Großserienfertigung nützlich.

1. Motivation

Die Entwicklung und Verbreitung von elektronischen Etiketten begann in den frühen 90er Jahren durch die Firmen Pricer und SES, welche erste Electronic Shelf Label mit LC-Displays mit Segmentanzeige auf den Markt brachten. Als erstes großes Unternehmen setzte die Handelskette Metro diese Displays 1995 in seinen deutschen Märkten ein. [1], [2]

Durch ihre Fähigkeit von einem Server dynamisch gesteuerte Preisänderungen anzeigen zu können, konnten ESL ihre Vorzüge gegenüber herkömmlichem Papier im Einzelhandel schnell unter Beweis stellen. Durch die Verwendung neuer Technologie wie vollgrafikfähige Dot-Matrix E-Paper-Displays, Funk-Kommunikation, Geolocation und Geofencing bieten ESL heutzutage vielfältige Möglichkeiten zum Einsatz [3]. So sind ESL mittlerweile auch im Fertigungsbereich angekommen und bieten den Unternehmen neue Möglichkeiten zur Digitalisierung [4].

Durch die zunehmende Verbreitung und Massenfertigung von Electronic Shelf Labels sind die Kosten in den letzten 5 Jahren laut Schätzung der Anbieter um ca. 30% gesunken [5]. Dadurch werden ESL als Mittel zur Digitalisierung nun auch für KMU eine interessante und lohnende Technologie. Dabei ist es für KMU wichtig, diese neuen Technologien in bestehende Softwareumgebungen zu integrieren, um eine weitgehend automatische Datengenerierung zu ermöglichen und die Mitarbeitenden bei der Pflege neuer Software-Systeme zu entlasten.

Innerhalb des Forschungsprojektes Framework Enriched Data Assembly - „FrEDA“ [6] wurde hierzu ein

ESL-System untersucht, welches über eine wartungsarme und kostengünstige Software verfügt. Zudem bietet das ESL-System über eine Schnittstelle die Möglichkeit zur softwaregestützten Anbindung an ein ERP-System an. [7]

Für eine möglichst breite Anwendung von ESL im Produktionsprozess wurden verschiedene Anwendungsszenarien herausgearbeitet und an Beispielen getestet, welche im Beitrag vorgestellt werden.

2. Grundlagen

2.1. Technologie (Hardware)

Die eingesetzten Electronic Shelf Label (ESL) der Firma TroniTAG GmbH verwenden vollgrafikfähige Dot-Matrix E-Paper-Displays, welchen das Prinzip der Elektrophorese zugrunde liegt. Im Display selbst sind kleine Farbkugeln enthalten, die sich durch Anlegen einer positiven oder negativen Spannung nach oben oder unten verlagern und damit sichtbar oder verdeckt werden [8].

In den Displays sind in der Regel drei verschiedene Farben von Kugeln vorhanden (schwarz, weiß, rot). Die e-Paper-Displays sind somit in der Lage dreifarbig und vollgrafisch Darstellungen anzuzeigen.

E-Paper-Displays bieten gegenüber ihrer Vorgänger-Technologie (LCD mit Segmentanzeige) den Vorteil, dass sie nach erfolgtem Bildaufbau kein Strom mehr benötigen, da die Farbkugeln ihre Position stromlos beibehalten. Da echte Farbpartikel eingesetzt werden, wird zudem von den ESL keine Strahlung/Licht emittiert, wodurch sie angenehm für die Augen sind. Zusätzlich

sind sie voll Grafikfähig, d.h. die Darstellung von QR-Codes, Bildern ist möglich.

Nachteilig ist, dass sie bei schwachem Umgebungslicht schlechter ablesbar sind als herkömmliche Display, da sie kein eigenes Licht emittieren.

Zur Übertragung der Daten verfügen die ESL zudem über ein Funkmodul, welches in Kombination mit einem Gateway die Kommunikation zum Server herstellt. Je nach ESL-System können hier verschiedene Funk-Frequenzen und Übertragungsprotokolle zum Einsatz kommen. Im Falle der verwendeten ESL der Firma TroniTAG GmbH handelt es sich um den Standard IEEE 802.15.4 bei einer Frequenz von 2,4 GHz.

2.2. Software

Die grundlegende Funktionsweise aktueller ESL-Systeme ist in Abbildung 2 dargestellt. Über ein Programm oder ein Webinterface kann von einem Endgerät auf den ESL-Server zugegriffen und das System gesteuert werden. Der Server kommuniziert über ein Gateway mit den ESL und kann somit Daten auf die Label via Funkverbindung übertragen und deren Zustand (z. B. Batteriestand) abfragen. [9], [10]

Im Falle des verwendeten ESL-Systems wird auf den Endgeräten ein Webinterface verwendet. Eine Besonderheit des Systems ist die Integration des Servers im Gateway. Dadurch können Unternehmen zusätzlich entstehende Kosten durch einen Server und dessen Wartung vermeiden. Im Webinterface werden die ESL mit den anzuzeigenden Datensätzen verknüpft. Das „Label-Design“ wird dabei über eine zusätzliche Software (Template-Editor) generiert. Die Datensätze der ESL werden entweder über die vorhandene API direkt aus der ERP-Datenbank generiert, oder in ein Austauschformat (hier: .csv) gespeichert und auf den Server / Gateway hochgeladen. [7]

2.3. Anwendungen in der Praxis

ESL werden heute schon vielfältig eingesetzt. Im Einzelhandel erlaubt die elektronische Preisauszeichnung eine dynamische Preisgestaltung, wodurch auf Preise der Konkurrenz schnell reagiert werden kann.

Ein weiteres Beispiel für den ESL-Einsatz ist die Raumbeschilderung im Bereich der Hotellerie und Gastronomie. Diese kann dabei nicht nur die Raumnummer anzeigen, sondern z. B. auch die Dauer der Belegung. [3]

Bisherige industrielle Einsätze beschränken sich auf Großunternehmen. Dabei wurden ESL sowohl als Lagerbeschriftung, als auch für die Auftragsbegleitung verwendet. [4]

3. Anwendungsszenarien

3.1. Versuchsstand

Zum Test der im Projekt FrEDA herausgearbeiteten Szenarien für den Einsatz elektronischer Etiketten (ESL) sowie deren Steuerung aus einer mobilen App heraus wurde im Institut InnArbeit der Hochschule Mittweida die Laborausstattung (Montagebereich) um mehrere ESL erweitert. Diese eignet sich durch seine Vielseitigkeit, mehrere Einsatzszenarien für ESL in der Logistik und Montage zu entwickeln und zu testen. Abbildung 1 stellt den Versuchsaufbau dar. Er besteht grundlegend aus einem Montagetisch mit Gelenkarm, einem Beistellwagen und einem separat aufstellbaren Handregal (für mehrere Arbeitsplätze gemeinsam nutzbar).

Am Montagearbeitsplatz werden verschiedene Variationen von Verteilersteckdosen gefertigt. Im Handregal werden die dafür benötigten Bauteile aller Steckdosenvarianten in verschiedenen Behältern aufbewahrt.

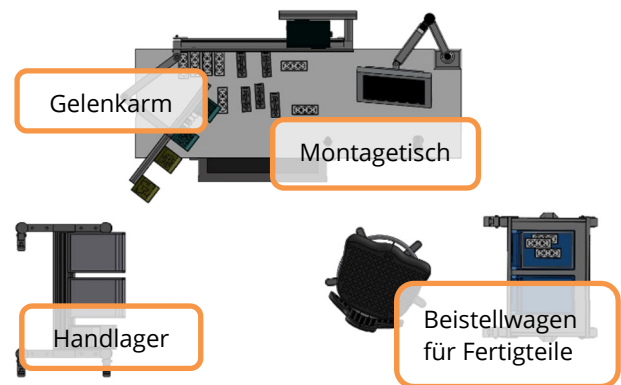


Abbildung 1: Aufbau der Testumgebung

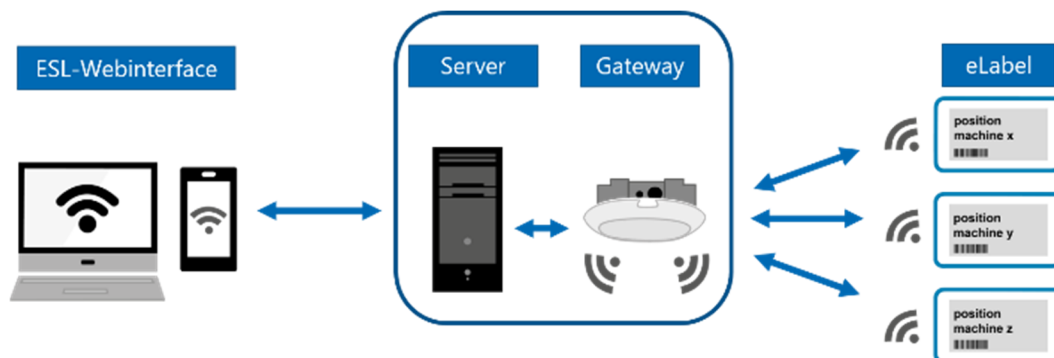


Abbildung 2: Funktionsweise ESL-System

Die Bauteile für die jeweils zu montierende Variante werden aus dem Handregal entnommen und in entsprechenden Behältern des Gelenkarms platziert. Nach erfolgter Montage werden die Fertigteile in den vorgesehenen Behältern auf dem Beistellwagen gelegt.

Dabei erhielten das Handregal, der Gelenkarm, die Behälter für Fertigteile, als auch der Montagetisch selbst ESL verschiedener Größen. ESL mit 2,9 Zoll Größe werden dabei für die Beschriftung der Behälter des Handregals und des Gelenkarms verwendet. Diese dienen der Unterstützung der Materialbereitstellung und der Montage. ESL mit 4,2 Zoll Größe werden zur Beschriftung des Arbeitsplatzes genutzt und ESL mit 5,8 Zoll Größe realisieren digitale Auftragsbegleitkarten an den Kleinladungsträgern (KLT).

3.2. Anwendungsszenario Behälterbeschriftung für die Materialbereitstellung

Für die Materialbereitstellung wurden elektronischen Etiketten am Handregal angebracht, so dass sich an jedem Behälter mit den darin befindlichen Einzelteilen jeweils ein ESL befindet. Auf den ESL werden Artikelnummer, Artikelname, Regalplatz und die aktuell verfügbare Menge angezeigt (Abbildung 3). Durch das Einscannen eines darauf befindlichen QR-Codes erfolgt die Weiterleitung zur zugehörigen Baugruppe.

Als neuartiges Pick-by-Vision-System befindet sich eine Anzeige (weißer Kreis) auf dem Label. Wird ein Kommissionierauftrag für eine Baugruppe ausgelöst, färbt sich dieser Kreis an allen zugehörigen Bauteilbehältern rot. Durch Quittieren der abgeschlossenen Materialbereitstellung wird der dargestellte Kreis wieder weiß und die angezeigte Menge an Bauteilen um die im Auftrag benötigte Menge reduziert. Die auftragsbezogene Kommissionierung lässt sich damit wirkungsvoll unterstützen (Aufwandsreduzierung, klare Bauteil Auswahl).



Abbildung 3: ESL für Materialbereitstellung im Ausgangszustand (li.) und ESL mit aktivem Pick-by-Vision für Auftrag (re.)

3.3. Anwendungsszenario Lagerbeschriftung

Im Lagerbereich können Lagerbereiche, Regale, Lagerplätze, Transporteinheiten und / oder Behälter durch ESL bezeichnet werden. Wichtige Informationen sind hier: Lagername, Regalbezeichnung (Reihe, Spalte), Sicherheitshinweise, Lagerfachbezeichnung und -inhalt, Behälter- und Ladeeinheitenbeschriftung, Artikelstatus oder Kommissionierhinweise. So wird es Mitarbeitenden erleichtert, Waren an den richtigen Plätzen zu lagern und schnell wieder aufzufinden. Der Pick-by-Vision-Ansatz lässt sich auch realisieren.

3.4. Anwendungsszenario Auftragsbegleitkarte

Des Weiteren wurden ESL als digitale und dynamische Auftragsbegleitkarte (ABK) zur Anwendung gebracht. Aufgrund fehlender Normen zur auftragsbezogenen digitalen Etikettierung von KLTs [11] [12] wurde ein eigenes Label-Design entworfen.

Das ABK-Label-Design enthält Informationen zum Artikelname, Artikelnummer, Auftragsnummer, Barcode, QR-Code für relevante Aktionen / Hinweise zum aktuellen Bearbeitungsschritt (z. B. Weiterleitung zu mobiler App), Kunde, Fertigungsmenge, Arbeitsgangnummer, Arbeitsgangname, Soll-Termin und Arbeitsplatz.

Für die Versuchsumgebung wurde am Beispiel der Baugruppen „2er Steckdose“ und „3er Steckdose“ ein Datensatz aufgebaut, welcher vier Arbeitsgänge enthält. Diese sind die Materialbereitstellung, Montage, Prüfung und Versand.

Der Ablauf des Einsatzszenarios „Auftragsbegleitkarte“ ist in Abbildung 4 dargestellt. Nach Eingang des Auftrags erfolgt das „Matching“ des zum Auftrag zugehörigen Datensatzes mit einem verfügbaren Label. Dies ist notwendig um den geforderten Auftrag auf einem ESL abzubilden. Jeder Datensatz enthält zudem ein „Steuerungsfeld“ („Promotion“), wodurch vordefinierte Anzeigen angesteuert und angezeigt werden können. Im Falle der ABK wurden unter anderem Markierungen des aktuellen Arbeitsgangs sowie Anzeigen zu Sperrung und Auftragsabschluss hinterlegt.



Abbildung 4: Ablauf des Einsatzszenarios „Auftragsbegleitkarte“

Mittels „Promotion“-Feld werden die Arbeitsgänge nacheinander angesprochen. Dadurch wird dieser Arbeitsgang im ABK-ESL hervorgehoben. Im Arbeitsgang Materialbereitstellung werden zusätzlich die ESL des Handregals angesteuert. Die ESL an den zum Auftrag gehörenden Bauteilbehältern werden mittels Pick-by-Vision-Anzeige markiert (siehe Abschnitt 3.2.).

Beim Arbeitsgang „Prüfung“ kann zusätzlich ein Sperrvermerk auf dem ESL eingeblendet werden. Nach Abschluss des Versands erfolgt die visuelle Anzeige des Auftragsabschlusses auf dem ESL.

3.5. Anwendungsszenario Montageunterstützung

Für die Unterstützung der Montage wurden die am Greifarm befindlichen Behältern mit ESL ausgerüstet. Darauf wird eine Kurzbeschreibung des Montageschritts, für den das Bauteil benötigt wird, dargestellt. Des Weiteren zeigt es die Arbeitsgangnummer, die benötigte Menge des Bauteils für diesen Arbeitsgang sowie einen QR-Code an, welcher auf eine Webseite mit detaillierten Informationen zum Arbeitsgang weiterleitet. Hierfür wird eine spezielle FrEDA-App zur Montage genutzt [6]. Montageaufgaben lassen sich somit unterstützen und Fehler (z. B. Verbau falscher Teile) vermeiden.

3.6. Anwendungsszenario Arbeitsplatz-/Maschinenbeschriftung

ESL bieten Mitarbeitenden und Vorgesetzten einen schnellen Überblick über den Zustand von Maschinen und Arbeitsplätzen. Dazu wurde ein Label-Design entworfen, welches Informationen zum Namen des Arbeitsplatzes / der Maschine, des Wartungsstatus (letzte Wartung und nächste Wartung), der Mängel, der Verantwortlichen Person sowie wichtige Arbeitsanweisungen und QR-Codes für Bedienungsanleitung, Sicherheitshinweise und Wartungsdokumente abbildet.

Im Falle einer Wartung kann zudem der Inhalt des ESL umgeschaltet werden. Dies blendet deutlich sichtbar einen Warnhinweis ein. So kann z. B. die Außerbetriebnahme des Arbeitsplatzes / der Maschine einfach und gut sichtbar vermittelt werden, was dem Schutz des Menschen förderlich ist.

4. Zusammenfassung/Ausblick

Die dargestellten Einsatzszenarien für ESL in der Montage und Logistik konnten am Versuchsstand getestet und validiert werden. Die Digitalisierung und zentrale Steuerung von Dokumenten, welche bisher nur in Papierform verfügbar waren, kann Unternehmen bei Änderungen helfen, Daten aktuell zu halten sowie korrekte und zügige Kommissionierung und Montage fördern.

Die gewonnenen Ergebnisse können nun in Unternehmen weiter getestet und ausgearbeitet werden. Hierzu gibt es Bestrebungen seitens des KMU CIMPCS GmbH das ESL-System bei Kunden zu testen und die Datengenerierung für die ESL im eigenen ERP-System „ProdatSQL“ zu implementieren [7].

Danksagung

Die Autoren danken dem Europäischen Fonds für regionale Entwicklung und dem Freistaat Sachsen für die finanzielle Unterstützung des Forschungsprojektes „Entwicklung eines neuartigen Framework zur mobilen Anreicherung, Auswertung und synergetischen Applikation von Daten für Montageverfahren in Fertigungsbetrieben“ (FrEDA). Das Forschungsprojekt wird vom Zentrum für innovative Arbeitsplanung und Arbeitswissenschaft

(InnArbeit) der Hochschule Mittweida und der CIMPCS GmbH Dresden bearbeitet.



Europäische Union

Europa fördert Sachsen.



Europäischer Fonds für regionale Entwicklung



Diese Maßnahme wird mitfinanziert durch Steuermittel auf Grundlage des von den Abgeordneten des Sächsischen Landtags beschlossenen Haushaltes.

Literaturverzeichnis

- [1] M. Chalberg, 6 Benefits of Electronic Shelf Labels and How ESLs Work, <https://koronapos.com/blog/benefits-of-electronic-shelf-labels/>, (2019).
- [2] Pricer AB, Die Geschichte von Pricer, <https://www.pricer.com/de/uber/geschichte-von-pricer/>, (2020).
- [3] M. Almeida, Elektronische Preisauszeichnung erklärt, <https://www.woutex.de/elektronische-preis-auszeichnung-erklart/>, (2020)
- [4] SES imagotag, Fujitsu – Business Case / Industrial <https://www.ses-imagotag.com/de/fujitsu/>, (2016).
- [5] B. Reiferscheid, Pricing: erst elektronisch, dann dynamisch, <https://www.stores-shops.de/technology/pricing-erst-elektronisch-dann-dynamisch/>, (2020).
- [6] L. Goldhahn, D. Bock, R. Eckardt, M. Friedemann, Webbasiertes Framework und Apps für die Montage – Entwicklung eines Frameworks und Beispiel-Apps für Montageverfahren in Fertigungsbetrieben, Zeitschrift für wirtschaftlichen Fabrikbetrieb (ZWF), Band 115 – Ausgabe 5, pp 321-326, (2020)
- [7] L. Goldhahn, D. Bock, Einsatz des webbasierten Framework der Montage (FrEDA) in Verbindung mit digitalen Beschriftungssystemen, 26. Interdisziplinäre Wissenschaftliche Konferenz an der Hochschule Mittweida (IWKM), (2021)
- [8] BECK GmbH & Co. Elektronik Bauelemente KG, E-Paper Display (EPD) - Informationen zur E-Paper Technologie, <https://www.beck-elektronik.de/produkte/displays/e-paper-display-epd/>, (2020).
- [9] Z. Chunhui, M. Pan, H. Liwen, L. Kezhong, W. Yuanqiao, An Electronic Shelf Label System Based on WSN, Advanced Materials Research, ISSN: 1662-8985, Vols. 765-767, pp 1718-1721 (2013)
- [10] T. Suh, H. Kim, J. Ko, V. Badrinarayanan, S. Bahk, Electronic Shelf Label: Prototype Development and Validation Using a Design Science Approach, Journal of Information Technology Management Volume XXIX, 23-38, (2018).
- [11] Verband der Automobilindustrie (VDA), VDA 4500 – Kleinladungsträger (KLT)-Systeme, (2018).
- [12] Verband der Automobilindustrie (VDA), VDA 4994 – Global Transport Label, (2020)

Einsatz des webbasierten Frameworks der Montage (FrEDA) in Verbindung mit digitalen Beschriftungssystemen

Leif Goldhahn, Dorit Bock

Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

Die Industrie fordert verstärkt die mobile Einsetzbarkeit von ERP-Systemen, verbunden mit einer hohen Daten- und Informationssicherheit [1]. Zur räumlich und zeitlich flexiblen Erfassung und Nutzbarmachung digitaler Daten mittels mobiler Endgeräte erfolgte im Verbundforschungsprojekt Framework Enriched Data Assembly – FrEDA die webbasierte Entwicklung eines praxistauglichen und industrieorientierten Frameworks. Dieses Framework ist verantwortlich für den Datenaustausch zwischen Mini- oder Kleinanwendungen (Apps) und datenbankbasierter Software (ERP-System) und findet Anwendungen in Montageprozessen und der dazugehörigen Logistik.

Der Beitrag zeigt Möglichkeiten, wie an manuellen Montagearbeitsplätzen die Daten aus dem ERP-System über das Framework direkt mit den digitalen Etiketten (Electronic Shelf Label - ESL) verknüpft und visualisiert werden können. Ein Beispiel ist die Applikation der Auftragsbegleitkarte aus dem ERP-System, um damit die Prozessschritte auf den Electronic Shelf Labels am Montagearbeitsplatz automatisiert anzuzeigen.

1. Motivation

Die fortschreitende Digitalisierung und die damit verbundene Verbreitung und Nutzung mobiler Geräte in Unternehmen nimmt immer weiter zu. Damit steigt auch die Datenmenge im Unternehmen auf verschiedensten Speichermedien, auf welche meist nur ein sehr beschränkter Personenkreis Zugriff hat bzw. weiß das die Daten dort zu finden sind. Eine Software, welche von den vielen Mitarbeitern täglich im Unternehmen genutzt wird, ist das üblicherweise vorhandene ERP-System.

Eine Verknüpfung der Daten/ Informationen von verschiedenen Systemen (z. B. aus dem zentralen ERP-System und aus dezentral angewendeten Apps) hat den Vorteil, dass damit Arbeitszeit gespart werden kann (Entfall doppelter Datenhaltung im ERP-System und auf mobilen Geräten). Daten für und von Tätigkeiten an Montagearbeitsplätzen stehen in Echtzeit zur Verfügung und helfen zur Vereinfachung und Qualitätssteigerung.

Im Forschungsvorhaben Framework Enriched Data Assembly - „FrEDA“ wurde dafür ein praxistaugliches und industrieorientiertes Framework entwickelt, welches mobile Mini- oder Kleinanwendungen (Apps) mit datenbankbasierter Software (ERP-System) verbindet. Dieses Framework verfügt über Technologien, die Bereiche wie Zugriffskontrolle, sicheren Datenaustausch und Datenhoheit abdecken. Außerdem konnte eine Eliminierung von Defiziten, Ungenauigkeiten und fehlender Transparenz sowie eine Reduzierung großer Datenmengen im mobilen Bereich umgesetzt werden.

2. Grundlagen

2.1. Framework FrEDA

Ein Framework ist ein modernes Softwaregerüst, das dem Programmierer einen Rahmen und grundlegende

Bibliotheken für die Entwicklung seiner Anwendungsprogramme zur Verfügung stellt und damit die Architektur der Software- und ihrer Schnittstellen bestimmt. Es wird vorwiegend in der objektorientierten Programmierung eingesetzt. Die Basisbausteine des Frameworks existieren in Form von abstrakten und konkreten Klassen und unterstützen das Erstellen von Applikationen. Es spezifiziert den Datenfluss und die Schnittstellen zwischen den Klassen. [2]

Als methodische Strukturierungsmöglichkeit des Frameworks wurde für FrEDA ein Ansatz in Anlehnung an das REFA-Arbeitssystem gewählt. Abbildung 1 stellt den grundlegenden Aufbau des Frameworks dar. Die Elemente des Arbeitssystems wurden in die wesentlichen Elemente des Frameworks transformiert. Der Mensch wurde im Modell durch Funktionen und die Betriebsmittel durch Programmstrukturen substituiert. Die Aufgabe des Systems ist einerseits die Erfassung der Daten per App bzw. die Bereitstellung der Daten aus dem ERP-System (z. B. Artikelnummer, -bezeichnung) und andererseits die Speicherung der mittels der Apps generierten Daten (z. B. Bilder, Videos).

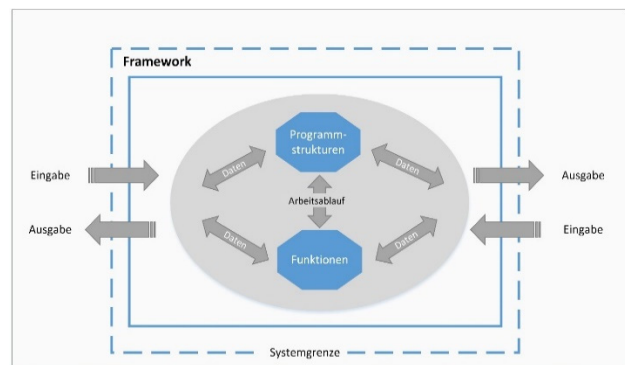


Abbildung 1: Ansatz Framework (i. A. a. REFA-Arbeitssystem [3])

Das Framework wurde in der Projektlaufzeit weiter spezifiziert und modular aufgebaut, um ein auf den jeweiligen Einsatzfall (Industrieanforderungen) anpassbares und erweiterbares System zu schaffen. Dabei sind Klassen bspw. für das Administrations-User-Interface (Admin-UI) und die Middleware-Datenbank (Abbildung 2) implementiert, welche unabhängig von mehreren Apps

gemeinsam genutzt werden können. Eine Anbindung an die Unternehmens-ERP-Datenbank als Datenquelle und zur automatischen Erkennung der Daten-Schemata realisiert eine GraphQL-standardisierte Schnittstellensprache. Dies ermöglicht auch die Freigabe von Schreib- und Leserechten sowie der Funktionalität des Frameworks für die Apps.

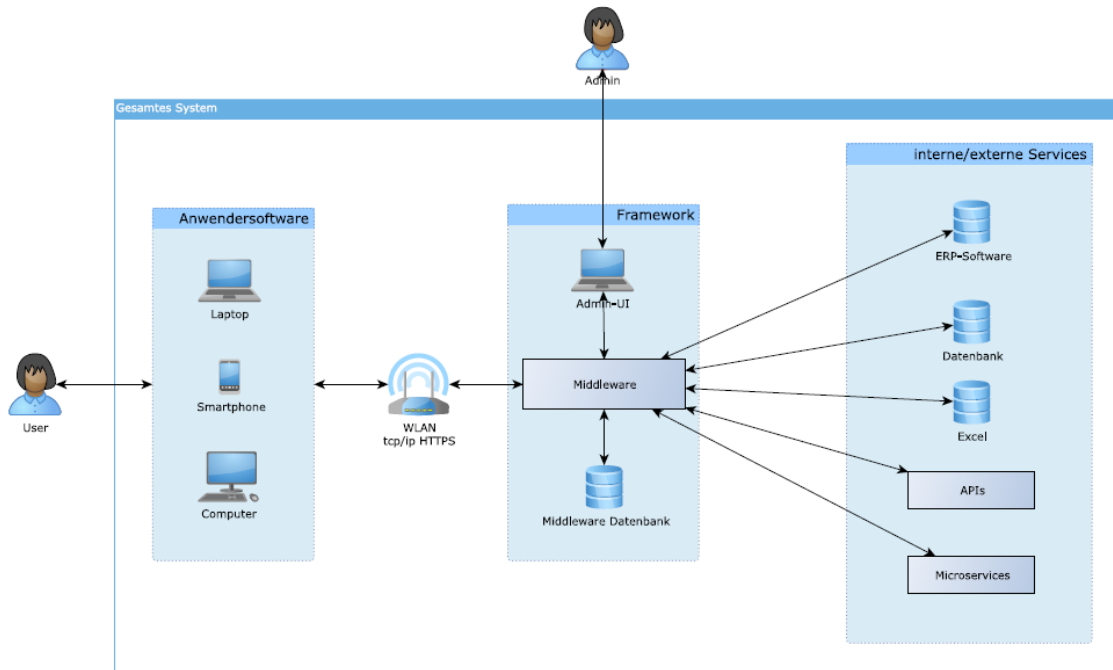


Abbildung 2: FrEDA-Framework

2.2. Apps

Die Entwicklung von Apps erfolgt meist, um eine einfache Kernfunktion, bspw. das Auslesen des QR-Codes, zu erfüllen. Apps sind im Allgemeinen klein und hochspezialisiert, informationsorientiert, mobil sowie mit Touch-Oberfläche, einfach bedienbar und plattformspezifisch. [4]

Bei der App-Entwicklung spielen wirtschaftliche und technische Ziele eine wesentliche Rolle. Dazu zählen Aspekte wie die Optimierung von Arbeitsabläufen sowie die Verbesserung der Qualität und der Zusammenarbeit. Ebenso müssen in diesen Zusammenhang Punkte wie die Verkürzung der Produktionszeiten und eine Produktionsüberwachung als mögliche Anwendungsfelder genannt werden. [5]

Die acatech – Deutsche Akademie der Technikwissenschaften unterstützt sowohl die Entwicklung bedarfsge-rechter Informationen in Apps als auch die Entwicklung aufgabenspezifischer Apps für den Fertigungsbereich [6].

Im Forschungsprojekt FrEDA wurde der Fokus auf Web Apps gelegt, da diese in der Regel geringere Entwicklungskosten haben als bspw. native Apps. Dies ist auch günstiger als die Entwicklung nativer Apps für ein einziges Betriebssystem. Außerdem sind diese Web Apps

plattformunabhängig, da sie über die Browser zur Anwendung kommen. Ein ebenfalls wichtiger Aspekt ist, dass die Web Apps auch ohne ständigen Internetzugang funktionieren, da HTML5 auch die Offline-Speicherung von Daten ermöglicht. Außerdem können Web Apps in Sekundenschnelle veröffentlicht und aktualisiert werden, da diese keinen Zulassungsprozess durchlaufen müssen. [7]

2.3. Elektronische Etiketten

Durch die zunehmende Digitalisierung, auch begünstigt durch Industrie 4.0, finden digitale Etiketten (Electronic Shelf Label – ESL) ihre Anwendung in Bereichen, in welchen sich die Kennzeichnungen, Beschriftungen und Beschilderungen häufig ändern. Hier entsteht Potenzial für die Kleinserien- oder kundenindividuelle Fertigung sowie die Logistik. Die Änderung der Beschriftung bei Auftragsänderungen erfolgt per Knopfdruck, teilweise auch automatisiert durch Verknüpfung mit einem Warenwirtschaftssystem. Auf den mehrfarbigen Displays (übliche Farben sind rot, schwarz, weiß) können wichtige Informationen (z. B. Produktdaten, Auftragsdaten, QR-Codes oder Barcodes) direkt hervorgehoben werden. Des Weiteren verfügen die Displays über einen ausgezeichneten Kontrast, so dass die dargestellten digitalen Inhalte aus verschiedenen Blickwinkeln und auch bei direkter Sonneneinstrahlung gut sichtbar sind. [8]

Damit besitzen die elektronischen, digitalen Etiketten grundsätzlich das Potential Papier als Beschilderungsinstrument zu ersetzen. Darüber hinaus ermöglichen sie schnelle Updates während der laufenden Produktion, was zu einer Automatisierung von Prozessschritten begünstigt.

3. Testspezifikation/ -anforderungen und Aufbau Versuchsstand

Im Forschungsprojekt wurde ein an der Hochschule Mittweida vorhandener Einzelmontagearbeitsplatz ausgewählt, welcher u. a. mit einem Montagetisch, Gelenkarm, Handregallager, Beistelltisch und Computer ausgestattet ist (Abbildung 3). Am Montagearbeitsplatz werden verschiedene Varianten von Verteilersteckdosen (2er, 3er und 4er Steckdose) gefertigt. Im Handregal (Abbildung 3, linke Seite) sind die dafür benötigten Bauteile aller Steckdosenvarianten in verschiedenen Behältern aufbewahrt. Die Bauteile für die jeweils zu montierende Variante werden aus dem Handregal entnommen und in entsprechenden Behältern des Gelenkarms platziert. Nach erfolgter Montage werden die Fertigteile in den vorgesehenen Behältern auf dem Beistellwagen (rechte Seite der Abb. 3) gelegt. Während der Projektlaufzeit wurde der Arbeitsplatz mit QR-Codes und ESL erweitert. Während die QR-Codes bspw. zur weiteren Informationsgewinnung für die verschiedenen Bauteile genutzt werden, sind die kleineren ESL (Größe: 2,9 Zoll) als Beschriftungsfeld der Behälter am Beistelltisch, am Gelenkarm und im Handregallager vorgesehen. Die größeren ESL (Größe: 5,8 Zoll) dienen als Auftragsbegleitkarte an Kleinladungsträgern (KLT) am Montagearbeitsplatz. Die ESL werden über einen zum System der Firma TroniTag GmbH gehörigen Server verwaltet und aktualisiert. Zu diesem musste im Projekt eine Schnittstelle realisiert werden.



Abbildung 3: Versuchsstand Hochschule Mittweida

4. Anwendung App eLabel

Bei der Entwicklung der App eLabel handelte es sich um eine gemeinschaftliche Arbeit der Projektpartner. Die CIMPCS GmbH erweiterte ihr ERP-System, um eine XML-

Exportfunktion der Auftragsbegleitkarten (ABK). Die Hochschule Mittweida implementierte die Schnittstellen-Programmierung zum eLabel-System der Firma TroniTAG GmbH. Anschließend realisierte die Hochschule Mittweida die App eLabel. Die Schritte die zur Übertragung der Daten notwendig sind, zeigt Abbildung 4.

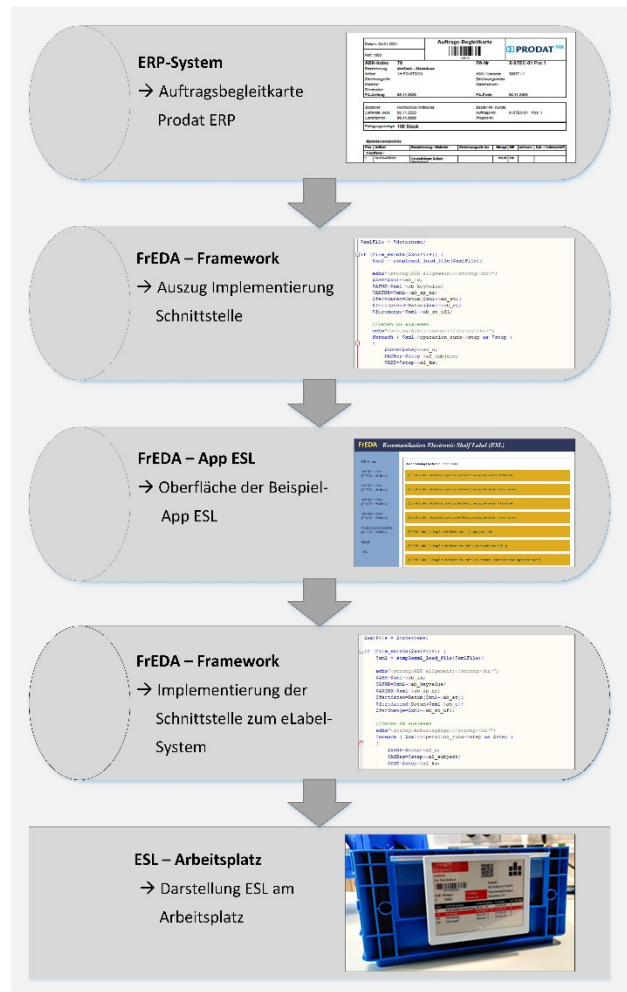


Abbildung 4: Schnittstelle zum eLabel-System zur Datenaufbereitung und -übertragung

Ziel war es auf einem mobilen Endgerät durch Aufruf einer ABK im ERP-System über das Framework und einer App, die entsprechenden Informationen auf das eLabel automatisiert zu übertragen. Dazu ist es notwendig im ERP-System als Vorbereitung den Menüpunkt „ABK-XML-Export“ für die entsprechende ABK einmalig durchzuführen. Ist die XML-Datei vorhanden, kann dann mit der App eLabel die entsprechende XML-Datei zur ABK ausgewählt werden. Mittels der App eLabel werden nun für die entsprechenden Etiketten am Montagearbeitsplatz die benötigten Daten ausgelesen und in eine CSV-Datei übertragen, welche die Daten zu den verschiedenen elektronischen Etiketten übermittelt.

Die App eLabel wurde auch anhand verschiedener Test-szenarien während und nach der Implementierung geprüft [9]. Während es in einem ersten Test um die Vollständigkeit der für die elektronischen Etiketten relevanten Daten ging, wurde in späteren Tests Wert auf die

Usability, eine intuitive Bedienung und auf eine möglichst zeitnahe Aktualisierung der eLabel gelegt. Hier sind aber die für die elektronischen Etiketten vorgegebenen Systemgrenzen zwecks Aktualisierung zu berücksichtigen, welche bspw. eine Aktualisierung in Echtzeit verhindern.

5. Validierung App eLabel

Für die Validierung des Frameworks und der Apps wurden verschiedene Einsatzszenarien im Montage und Logistikbereich herausgearbeitet [9, 10] und dafür die App eLabel getestet. So wurden die Arbeitspläne der zu fertigenden Bauteile (z. B. 2er/ 3er Steckdose, PTO-Getriebe) im ERP-System Prodat ERP hinterlegt. Damit konnten im ERP-System Aufträge zur Fertigung der Bauteile angelegt und die dazugehörigen Auftragsbegleitkarten erstellt sowie die XML-Dateien zur ABK generiert werden. Durch die App eLabel erfolgte nun die Datenverarbeitung und der Datentransfer zu den digitalen Etiketten am Arbeitsplatz. Nach dem Datenaustausch mittels des Frameworks konnte überprüft werden, ob die korrekten Daten von der Auftragsbegleitkarte aus dem ERP-System auf den entsprechenden Etiketten am Montagearbeitsplatz angezeigt werden. Die Lösung erwies sich als durchgehend korrekt und erreichte akzeptable Anzeigeverzögerungen. Diese Laufzeit ist abhängig von der Anzahl gleichzeitig zu ändernder Label und der Datengrößen.

6. Zusammenfassung

Für die Beispiel-App eLabel konnte nachgewiesen werden, dass die mobile Ansteuerung der eLabel aus dem ERP-System funktioniert und damit eine Übertragung auf Unternehmen möglich ist.

Dieser Beitrag sollte anhand eines praktischen Anwendungsfalls zeigen, wie vielfältig das FrEDA-Framework genutzt werden kann. Durch das Framework ist jetzt eine Middleware vorhanden, mit welcher die Nutzung von kleinen mobil anwendbaren Apps in Verbindung mit einem ERP-System möglich ist. So haben Unternehmen die Möglichkeit durch verschiedene kleine Apps mehr mobil zu arbeiten, bei Bedarf auch unabhängig vom ERP-System.

Danksagung

Die Autoren danken dem Europäischen Fonds für regionale Entwicklung und dem Freistaat Sachsen für die finanzielle Unterstützung des Forschungsprojektes „Entwicklung eines neuartigen Framework zur mobilen Anreicherung, Auswertung und synergetischen Applikation von Daten für Montageverfahren in Fertigungsbetrieben“ (FrEDA). Das Forschungsprojekt wird vom Zentrum

für innovative Arbeitsplanung und Arbeitswissenschaft (InnArbeit) der Hochschule Mittweida und der CIMPCS GmbH Dresden bearbeitet.



Europäische Union

Europa fördert Sachsen.



Europäischer Fonds für regionale Entwicklung



Diese Maßnahme wird mitfinanziert durch Steuermittel auf Grundlage des von den Abgeordneten des Sächsischen Landtags beschlossenen Haushaltes.

Literaturverzeichnis

- [1] Trovarit AG: ERP in der Praxis - Anwenderzufriedenheit, Nutzen & Perspektiven 2020/2021, Aachen, 2020
- [2] Framework, <https://www.itwissen.info/Framework-framework.html>, verfügbar am 11.01.2021
- [3] Autorenteam REFA – Verband für Betriebsorganisation und Unternehmensentwicklung e. V.: REFA Kompakt-Grundausbildung 2.0; 2. überarb. Aufl., 2013
- [4] Forum IT@Automation: App-Entwicklung in der Industrie: Grundlagen und Entscheidungshilfe, Industriekreis Smart Devices, VDMA, 2014
- [5] Tomic, M.: Apps für KMU: Praktisches Hintergrundwissen für Unternehmer; Springer Gabler, Springer Fachmedien Wiesbaden, 2015
- [6] Schuh, G.; Anderl, R.; Gausemeier, J.; ten Hompel, M.; Wahlster, W. (Hrsg.): Industrie 4.0 Maturity Index – Die digitale Transformation von Unternehmen gestalten, acatech – Deutsche Akademie der Technikwissenschaften, 2017, S. 27, 41, 60
- [7] Franke, F.; Ippen, J.: Apps mit HTML5 und CSS3. Galileo Press, Bonn, 2012, S. 22ff.
- [8] e-shelf-labels, Elektronische Etiketten in Produktion und Logistik https://www.e-shelf-labels.de/hardware/elektronische-etiketten.html?gclid=EAlaIQob-ChMI9Kb_wc_A7gIVUzCh3RfQr3EAAYBCAAE-glfKfD_BwE, verfügbar am 29.01.2021
- [9] Goldhahn, L., Roch, S.: Einsatzszenarien für Electronic Shelf Label in der Produktion, 26. Interdisziplinäre Wissenschaftliche Konferenz Mittweida, Mittweida, 2021
- [10] Goldhahn, L.; Bock, D.; Eckardt, R.; Friedemann, M.: Webbasiertes Framework und Apps für die Montage, ZWF (ISSN 0947-0085 Jahrg. 115 (2020) 5, doi.org/10.3139/104.112291, S. 321 - S. 326

Entwicklungsbegleitende Planung neuer Montageprozesse in der Zulieferindustrie

Michael Willnich

Pierburg Pump Technology, Sonnenstraße 29, 04746 Hartha

Darstellung über das Zusammenspiel von Entwicklung und Prozess im Projekt. Was muss beachtet werden, um Produkte von Projektstart an herstellbar und montagegerecht zu gestalten, um Kosten und Prozesszeiten zu reduzieren sowie Probleme in der späteren Montage zu vermeiden? Welche Voraussetzungen werden hierfür benötigt und welche Hilfsmittel helfen bei der Entscheidungsfindung?

1 Einleitung

Das magische Dreieck - Zeit, Kosten und Qualität - ist jedem im Projektmanagement bekannt und die Balance zwischen den drei Faktoren zu halten ist nicht immer leicht. In diesem Artikel geht es um das Zusammenspiel von Entwicklung und Prozess während der Produktentstehung von der Anfrage bis hin zur Serienreife. Was führt beim Produktentstehungsprozess zu einem montagegerechten und kostenoptimierten Produkt unter Einhaltung der jeweiligen Anforderungen? Hier ein Beispiel aus der Praxis.

2 Der Produktentstehungsprozess

Der Produktentstehungsprozess umfasst alle Tätigkeiten zur Umsetzung eines Produktes beginnend von der Kundenanfrage bis zum Erreichen der Serienreife. Dabei geht es nicht nur darum ein Produkt gemäß den geforderten Kundenspezifikationen zu erstellen, sondern auch ein montagegerechtes und kostenoptimiertes zu erhalten. Um den Vorgang zu steuern, wird der Prozess in Phasen eingeteilt. Im Hause Pierburg erfolgt die Einteilung beispielsweise in sieben Sektionen innerhalb vier Phasen (siehe Abb. 1). Jede dieser Phasen schließt mit festgelegten Meilensteinen ab. Diese werden zu Projektstart, in Abhängigkeit der Kundentermine, im Projektterminplan festgeschrieben. Die Meilensteine helfen somit, das Gesamtziel nicht aus den Augen zu verlieren.

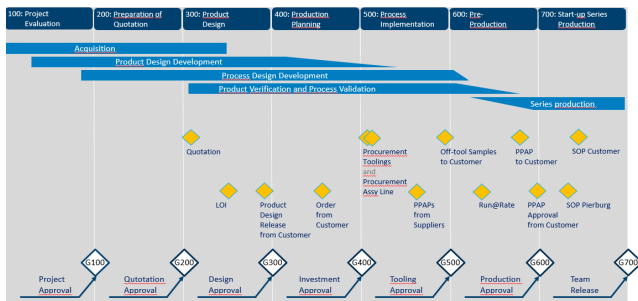


Abbildung 1: Produktentstehungsprozess [1]

Wir befinden uns seit Jahren in einem Wandel, in der die Schnelllebigkeit eingezogen ist. Die Elektronikbranche, gerade im Bereich der Smartphones, macht es vor und lebt von Innovationszyklen unter 1 Jahr. Doch auch in der Zulieferindustrie für die Automobilbranche steigt die Produktvielfalt mit immer kürzer werdenden Entwick-

lungszeiten. Wer hier zu langsam ist, läuft Gefahr abgehängt zu werden. Im Regelfall sind Entwicklungszeiten aber bereits derart optimiert, dass man diese, in Bezug auf die Entwicklungsqualität, kaum weiter verkürzen kann. Daher ist eine Parallelisierung der Teilprozesse notwendig sowie ein bereichsübergreifendes „Kompetenz-Team“, damit alle Aspekte während der Produktentstehung betrachtet werden. Das Team besteht dabei in der Regel aus Entwicklung, Produktion, Qualität, Logistik, Controlling, Vertrieb und Einkauf. Dadurch umfasst der Produktentstehungsprozess auch angrenzende Bereiche und Prozesse. Doch dieser Vorgang ist nicht neu in der Automobilindustrie. So hat bereits eine Studie Mitte der 80iger Jahre gezeigt, dass japanische Hersteller deutlich schneller (3 Jahre) mit der Einführung eines neuen Modells sind als europäische (9 Jahre) [2].

2.1 Simultaneous Engineering

Doch wie schafft man es, die Projektdauer zu verkürzen? Während traditionell die Produktentwicklung und Produktionsplanung getrennt und nacheinander erfolgt, benötigt man einen Prozess, in dem sich die einzelnen Phasen überschneiden. Man spricht bei dieser Vorgehensweise von Simultaneous Engineering (auch als Concurrent Engineering benannt). Hierbei arbeitet das Kernteam gleichzeitig und parallel an seinen spezifischen Aufgaben. Dies bringt den Vorteil, dass Fehler schneller erkannt und die Projektdurchlaufzeit deutlich verkürzt wird. Abbildung 2 zeigt schematisch den Zeitgewinn, welcher durch Simultaneous Engineering (blau) gegenüber einem traditionellen Projektlauf (grau) erzielt werden kann.

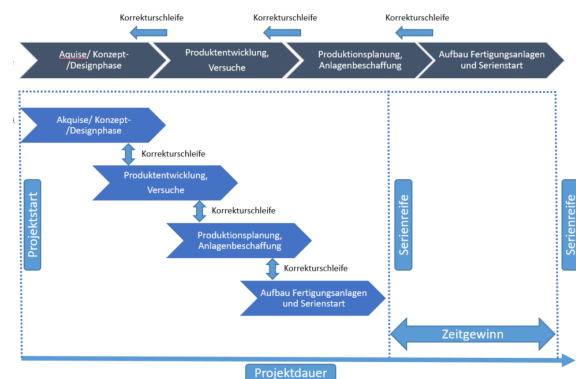


Abbildung 2: Vergleich Traditioneller Projektlauf vs. Simultaneous Engineering.

Jedoch wird dadurch nicht nur die Durchlaufzeit verkürzt, sondern es können auch Kosten durch die frühe Vermeidung von Fehlern eingespart werden. Hier kann man eine Faustformel aus dem Qualitätswesen, die Zehnerregel, anwenden, welche besagt: je später man den Fehler entdeckt, desto höher werden die Aufwände und Kosten zur Behebung des Fehlers (siehe Abb. 3). Zum anderen erzielt man eine Kosteneinsparung durch Vermeiden des Weiterführens unnötiger Konzepte. So kann ein früher Designentwurf der Entwicklung am Anfang der Projektphase sowohl kostenseitig als auch auf technische Machbarkeit durch die Produktion geprüft und entsprechend bewertet werden.

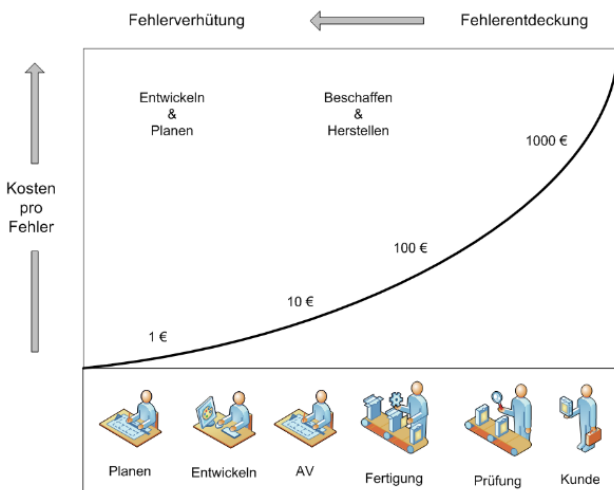


Abbildung 3: Zehnerregel [3]

Die genannten Vorteile erkaufte sich das Simultaneous Engineering allerdings mit hohem Aufwand in den ersten Phasen des Produktentstehungsprozesses. Da bereits parallel zur Produktentwicklung die Industrialisierung geplant wird, muss fachübergreifend genügend Manpower zur Verfügung stehen, um die verschiedenen Konzepte zu bewerten und zu finalisieren. Des Weiteren steigt die Menge an unvollständigen Informationen aufgrund parallel ablaufender Teilprozesse, welche zum Zeitpunkt noch nicht abgeschlossen sind. Es drohen Schnittstellenverluste.

Daher bedarf es Leitsätze für das Simultaneous Engineering. Diese sind Parallelisieren, Standardisieren und Integration [4].

Parallelisieren:

Parallelisieren bedeutet das Beginnen von Prozessen vor Abschluss der Vorgängerprozesse, wodurch Zeitverkürzungen bzw. Zeitoptimierungen entstehen. Die Grundlage dafür sind Informationen, die bereits nach kurzer Zeit der jeweiligen Phase zur Verfügung stehen. Dadurch kann ein zeitlich vorgezogener Nachfolgeprozess bereits starten.

Standardisieren:

Die Standardisierung von Prozessen hilft besonders bei der Parallelisierung komplexerer Themen. Ziel ist das

Vermeiden von Wiederholungen, unnötigen Arbeiten sowie das Einbringen von lessons learned. Die Standardisierung bezieht sich sowohl auf technische als auch auf prozessuale und organisatorische Aspekte.

Integration:

Hierbei ist gemeint, dass im Produktentstehungsprozess verschiedene Bereiche in interdisziplinären Teams an einem ganzheitlichen Ziel arbeiten, um Schnittstellenverluste zu vermeiden.

2.2 Das Team

Am Beispiel von Pierburg besteht dieses Kernteam, dem sogenannten SE-Team (siehe Abb. 4), aus Projektmanager - Entwicklung, Projektmanager - Produktion, Vertrieb, Einkauf, Controlling und übergeordnet aus dem Programmmanager, welcher die Gesamtprojektverantwortung innehat. Somit nutzt das Team die Ansätze und Denkweisen verschiedener Fachabteilungen. Jedes dieser Teammitglieder erarbeitet wichtige Lösungen und ist essentiell zur Umsetzung des Gesamtprojekts. Des Weiteren ist jedes Teammitglied dafür verantwortlich die wesentlichen Informationen zu filtern und in die entsprechenden Abteilungen zu verteilen. Nur wenn ein klarer Informationsfluss stattfindet, kann das Auftreten kosten- und zeitaufwendiger Korrekturschleifen minimiert werden.

Auf Grund der großen Menge an Informationen, die durch die einzelnen Bereiche eingebracht werden und um einen Kontrollverlust im Projekt zu vermeiden, trifft sich das Team regelmäßig. Hier werden Lösungsansätze in Bezug auf Kosten, Qualität und Termin besprochen. Ebenfalls helfen die im Produktentstehungsprozess verankerten Meilensteine beim Erzielen von Teilerfolgen.

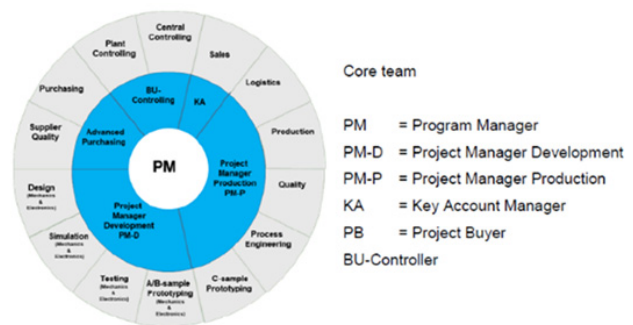


Abbildung 4: SE-Team [5]

Hierbei zeigt sich ein weiterer Nachteil. Fehlt ein Teammitglied, fehlen möglicherweise wichtige Informationen. Eine schnelle Entscheidungsfindung kann somit ins Stocken geraten. Daher ist es ungemein wichtig alle Informationen in den entsprechenden Teammeetings zur Verfügung zu haben. Des Weiteren ist eine agile Arbeitsweise erforderlich. Das Team muss flexibel sein, sich auf unvorhersehbare Probleme und auf plötzliche Planänderungen einstellen sowie darauf entsprechend handeln.

2.3 Die Hilfsmittel

Doch welche Hilfsmittel werden benötigt, um Produkte von Projektstart an herstellbar und montagegerecht zu gestalten? An erster Stelle soll hier nochmals das Projektmanagement auf Basis einer interdisziplinären Zusammenarbeit genannt werden.

Wie bereits beschrieben, lebt das Simultaneous Engineering von einer hohen Kommunikation. Dabei kann im Kernteam aber nicht alles besprochen und geklärt werden. Damit ein neues Produkt bereits in der Designphase montagegerecht gestaltet wird, sind engste Abstimmungen zwischen Entwicklung und Produktion unabdingbar. Hier wird ganz gezielt über die Themen „Produktanforderung, Herstellbarkeit, Kosten und Umsetzung“ diskutiert. Um dabei innerhalb des magischen Dreieck zu bleiben, helfen Verfahrensweisen wie Design for Manufacturing (DFM) und Design for Assembly (DFA). Bei Ersterem wird, wie der Name verrät, bereits in der Designphase der Fokus auf die Produktion gelegt. Bei der DFA liegt der Schwerpunkt auf der Automatisierbarkeit und der Umsetzung von Kundenwünschen (Vielfalt). Beide sorgen dafür, das Produkt später schneller, kostengünstiger und in hoher Qualität zu fertigen. Die frühe Einbindung eines Anlagenherstellers bringt ebenfalls wichtige Informationen über die Herstellbarkeit. Sie helfen ein Produkt kostengünstig und montagegerecht zu gestalten. Denkt man nun wieder zurück an die Zehnerregel, stellt man fest, wie wichtig diese Abstimmungsphasen in einem Projekt sind. Denn mit fortschreitender Projektdauer erhöhen sich die Aufwände für etwaige Änderungen.

Ein weiteres Hilfsmittel stellt die, von Anfang an durchgehende, Dokumentation und Datenspeicherung dar. Mit Hilfe von ersten Zeichnungen in der Designphase können bereits zeitig Flächen und notwendige Toleranzen festgelegt werden, die später den Zusammenbau vereinfachen bzw. die Voraussetzung einer automatisierten Montage bilden. Die Speicherung der Daten in einem ERP System helfen zudem die Dokumente für alle Bereiche einfacher zugänglich zu machen.

Ein Modell sagt mehr als tausend Worte. Hier in abgewandelter Form ist das Sprichwort sehr zutreffend. Neben der engen Abstimmung zwischen den einzelnen Bereichen über das Produkt, hilft ein 3D-Modell für das Finden erster Lösungen. Mehr als ein 3D-Modell kann durch Rapid Prototyping erzielt werden. Vor allem im Bereich des 3D-Drucks konnten hier in den letzten Jahren wahnsinnige Fortschritte erzielt werden, die bei der Auslegung von Prozessen nicht mehr wegzudenken sind. Denn mit Hilfe eines Prototyps können Diskussionen und Problemstellungen erheblich schneller geklärt werden. Die Genauigkeiten, die mittlerweile erzielt werden, sind so hoch, dass man damit bereits Vorrichtungen, Aufnahmen oder entsprechende Verpackungen auslegen und finalisieren kann, wenn noch keine Teile aus dem Serienwerkzeug zur Verfügung stehen. Dies spart

nicht nur Zeit, sondern auch Kosten. Gerade bei der Integration einer neuen Variante auf eine bestehende Fertigungsanlage können dadurch alle notwendigen Informationen in kürzester Zeit ermittelt werden. Abbildung 5 zeigt einige Beispiele für 3D-Drucke.

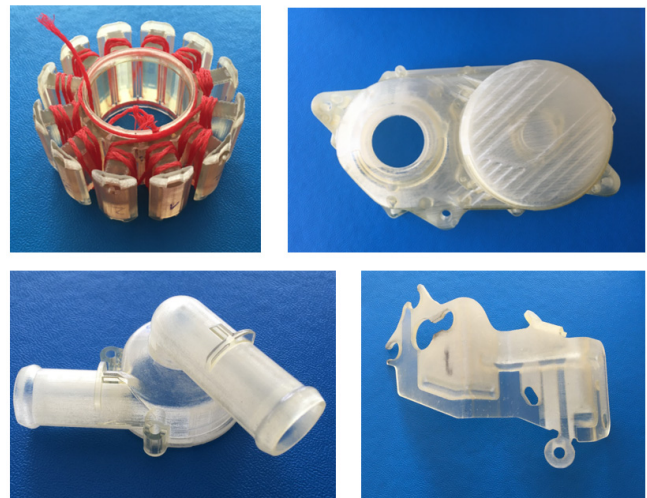


Abbildung 5: Beispiele für 3D-Druck (obere Reihe: Stator, Motorgehäuse; untere Reihe: Pumpengehäuse, Halter)

Das Ganze funktioniert natürlich auch anders herum, indem man Vorrichtungen beispielsweise mittels Rapid Prototyping erstellt, um einen ersten Eindruck von einem Montageprozess zu bekommen.

3 Zusammenfassung

Die entwicklungsbegleitende Planung neuer Montageprozesse ist notwendig, um die Projektdurchlaufzeit zu verkürzen, Aufwand und Kosten niedrig zu halten sowie die Produktqualität zu steigern. Das Simultaneous Engineering bietet durch Parallelisierung, Standardisierung und Integration hierfür die Möglichkeit. Notwendig ist ein Projektmanagement mit einem interdisziplinären Team. Die Vorteile erfordern allerdings eine hohe Qualifikation der Mitglieder im SE-Team, ein agiles Handeln, einen enormen Kommunikationsbedarf innerhalb der Fachabteilungen und den Prozessen sowie das gemeinsame Arbeiten an dem Gesamtziel. Eine kleine Gruppe an Personen zeichnet sich verantwortlich über Erfolg und Misserfolg in der Produktentstehung. Durch die Verfahren DFM und DFA kann gezielt ein montagegerechtes Produkt innerhalb des magischen Dreiecks entworfen werden. Darüber hinaus helfen von der Entwicklung zur Verfügung gestellte technische Hilfsmittel, zum Beispiel Zeichnungen, 3D-Modelle, Stücklisten und Prototypen, bereits in der Designphase Probleme frühzeitig zu erkennen und Produkte gezielt zu gestalten.

Literaturverzeichnis

- [1] Pierburg Dokument (2017): Produktentstehungsprozess, Verfasser: Volker Lankes
- [2] W. Bochtler, (1995): Simultaneous Engineering - Erfahrungen aus der Industrie für die Industrie, Seite 1, Springer Verlag, Heidelberg

- [3] R. Schnurr (Hrsg.): Fehlerkosten 10er Regel Zehnerregel (Rule of ten), <https://www.sixsigmablackbelt.de/fehlerkosten-10er-regel-zehnerregel-rule-of-ten>, (06.02.2021)
- [4] A. Stanke, S. Berndes (1997): Forschungs- und Entwicklungsmanagement, Seite 15 – 17, Vieweg+Teubner Verlag
- [5] Pierburg Dokument (2020): Mechatronics SE Team – Core Functions, Verfasser: Volker Lankes

Standard Assembly Time Setting in an Early Stage of Product Development

Izabela Kutschenreiter-Praszkiewicz

University of Bielsko-Biala / Willowa 2, 43-300 Bielsko-Biala, Poland

Standard assembly time is an important piece of data in product development that is used to compare different product variants or manufacturing variants. In the presented approach, standard time is created with the use of a decision tree regarding standard manual and machine-manual operations, taking into consideration product characteristics and typical tools, equipment and layout. The analysed features include, among others: information determined during product development, such as product structure, parts characteristics (e.g. weight, size), connection type, as well as the information determined during assembly planning: tools (e.g. hand screw driver, power screw driver, pliers), equipment (e.g. press, heater), workstation layout (e.g. distance, way of feeding). The object-attribute-value (OAV) framework was applied for the assembly characteristic. An example of the decision tree application to predict standard assembly time was presented for a mechanical subassembly. The case study was dedicated to standard time prediction for a bearing assembly. The presented approach is particularly important for the enterprises which offer customized products.

1. Introduction

Standard time is the time required by an average skilled operator, working at a normal pace, to perform a specified task using a prescribed method [1]. Standard time is widely used in industrial engineering for workforce planning, line balancing, production system simulation, cost accounting etc. Standard time can be determined with the use of many techniques, including: time study, predetermined motion time system, standard data system or work sampling. Time standards are the basic data in ERP (Enterprise Resource Planning) systems. There is a gap in time standard setting method which can be effectively used for new product planning in ERP systems. The article presents decision tree application and rules induction focused on time standard setting for new product assembly in a production process. Time standard is especially important in assembly planning for manual and machine-manual operations. Authors [2][3] have focused their research on establishing an assembly information model to integrate product information from CAD model with assembly manufacturing information. The main difficulties are related to assembly complexity and plenty of data.

2. Time standard setting - the proposed approach

The proposed approach of time standard prediction for assembly tasks of a new product is based on the following steps:

- Decomposing the assembly process into subassemblies
- Calculating time standards for typical subassemblies,
- Assigning time standards to classes,

- Developing a training set with attributes and values characterising subassemblies,
- Building a decision tree,
- Formulating decision rules for assembly time standard prediction.

2.1. Decomposition of the assembly process into subassemblies

In the first step of the proposed approach focused on assembly process decomposition, it is possible to use the graph theory. According to literature review, two widely used methods for assembly problem decomposition are the graphs of precedences for assembly and disassembly, and the graph of connections between components [4] (liason graph), which is particularly useful.

2.2. Calculation of time standards for typical subassemblies

The next step in the proposed approach is time standard calculation. The following methods can be used to determine time standards [5]:

- Estimation
- Historical records
- Work measurement technique (time study, predetermined motion time systems, standard data system, work sampling)

The most useful methods of time standards calculation for typical subassemblies are time study or predetermined motion time system.

2.3. Assignment of time standards to classes

A catalogue of typical assembly operation can be used in the proposed approach. Standard time for classes can be created basing on the equation (1).[6]

$$w_j = \frac{e^{\sqrt{12T_m(t_n)_j}}}{196} \quad (1)$$

Where:

w_j - the width of the j class

e - accuracy at the confidence level of 95%

T_m - settlement period

$(t_n)_j$ -time standard for j class

Each typical assembly task should be assigned to the appropriate class.

2.4. Development of a training set

In the proposed approach, the assembly process is represented by the object-attribute-value OAV scheme, in which an object is associated with a set of attributes and each attribute is described by appropriate values. The OAV scheme gives a concise data structure for organising the features of a selected process [7]. The main attributes affecting assembly standard time are presented in Figure 1. [2]

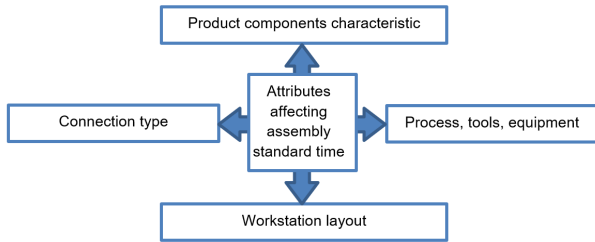


Figure 1. Attributes influencing assembly standard time

Based on the OAV framework, a training set can be created (Tab. 1).

Table 1. A training set

| Attributes charactering assembly process (values) | | | | Decisions (classes) |
|---|------------------------|-----|------------------------|---------------------|
| $A_1(V_1, \dots, V_n)$ | $A_2(V_1, \dots, V_m)$ | ... | $A_k(V_1, \dots, V_2)$ | C_1, \dots, C_j |
| | | | | |

The values in a training set can be divided into intervals according to the following method:

The number of intervals i_p in a training set for a given attribute can be calculated according to the formula (2) [6].

$$i_p = \frac{\log t_{\max} - \log t_{\min}}{\log q} \quad (2)$$

Where:

t_{\max} – max of time standard

t_{\min} – min of time standard

q – quotient calculated according to formula (3) [6]

$$q = \frac{1 + e_t}{1 - e_t} \quad (3)$$

Where:

e_t – precision

Based on formulas (4) and (5), the intervals are fixed.

$$t_i = t_{\min} \cdot q_s^i \quad \text{dla } i=1,2,\dots,i_p \quad (4)$$

$$X_i = \frac{t_i - b}{a} \quad \text{dla } i=1,2,\dots,i_p \quad (5)$$

Where:

q_s - quotient for a total number of intervals

a, b – coefficients in relation between x (variable – attribute being analysed) and t (time fixed according to the regression analysis).

2.5. Building a decision tree

Classical decision trees belong to popular classification models [8]. A decision tree is a graph which can be used as a model of a categorical variable (attributes). A decision tree aims at predicting a categorical (numerical or linguistic) output variable from a set of numerical or linguistic input variables [9]. Decision trees are useful in solving classification and prediction problems [10], [11]. The structure of a decision tree involves a root node, internal nodes, leaf nodes and edges which joint nodes, also called branches (fig. 1) [12].

The well-known calculations of the decision trees is ID3 [13][14] [8] which use the information gain which based on the probability theory (Shannon entropy).

ID3 algorithm steps includes:

- Calculating the entropy of every attribute using the data set according to (equation (6)),
- Splitting the set into subsets using the attribute (formula (7)) for which entropy is minimal (or, equivalently, information gain is maximal),
- Making a decision tree node containing that attribute,
- Recursing on subsets using the remaining attributes.

Entropy I is a measure of the amount of uncertainty in a data set (i.e. entropy characterizes a data set).

$$I = \sum_{i=1}^n (-p_i \log_2 p_i) \quad (6)$$

Where:

p – proportion of the number of elements in a class to the number of elements in a set (probability that element from i class occurs)

$$I(C/A_k) = \sum_{j=1}^{M_k} p(a_k, j) \cdot \left[- \sum_{i=1}^N (p(c_i/a_{k,j}) \cdot \log_2 p(c_i/a_{k,j})) \right] \quad (7)$$

Where:

M_k number of values taken by attribute A_k

N number of classes

k number of attributes

$p(a_k, j)$ probability that a_k takes value j

$p(c_i/a_{k,j})$ probability that class c_i occurs, when $a_k=j$

Information gain is calculated according to formula (8).

$$\Delta I(A_k) = I - I(C/A_k) \quad (8)$$

Decision tree induction is closely related to rule induction; each path from the root of a decision tree to one of its leaves can be transformed into a rule [15], which is

one of the most popular approaches to knowledge representation.

2.6. Formulating decision rules

Rule-based systems are built around rules, which consist of an *if* part and a *then* part [16].

Rules, sometimes called IF-THEN rules, can take various forms e.g.:

- IF condition THEN action
- IF premise THEN conclusion

3. An example of time standard setting with the use of the proposed approach

An example relates to a bearing assembly in toothed gear subassembly. An example of the liaison graph built for a subassembly (Fig. 2) is presented in Figure 3.

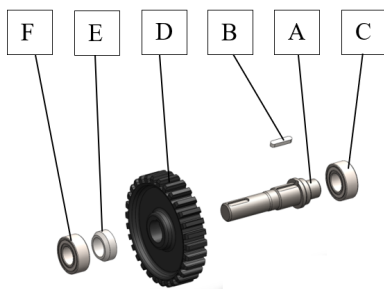


Figure 2. Subassembly

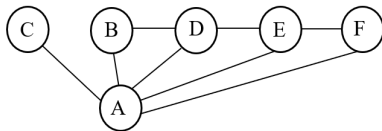


Figure 3. A Liaison graph

The connection A-C in the liaison graph was analysed. Bearing internal diameter BID, bearings heating concurrently HC and hitting H were the attributes taken into consideration. Standard time ST was calculated according to the predetermined motion time system and was divided into classes (tab. 2).

Table 2. ST classes definition

| Class | Lower bound of the class [h] | Upper bound of the class [h] | Average value of the class[h] |
|-------|------------------------------|------------------------------|-------------------------------|
| A | 0,00 | 0,15 | 0,07 |
| B | 0,15 | 0,32 | 0,23 |
| C | 0,32 | 0,49 | 0,41 |
| D | 0,49 | 0,67 | 0,58 |

The analysed attributes and their values were presented in table 3.

Table 3. The analysed data set

| BID | HC | H | ST[h] | ST[s] | Class |
|-----|----|----|----------|-------|-------|
| 60 | 0 | No | 0,106944 | 385 | A |

| | | | | | |
|-----|---|-----|----------|-----|---|
| 80 | 6 | Yes | 0,153889 | 554 | B |
| 50 | 2 | Yes | 0,140278 | 505 | A |
| 50 | 6 | Yes | 0,157222 | 566 | B |
| 200 | 6 | Yes | 0,203889 | 734 | B |
| 210 | 2 | Yes | 0,190278 | 685 | B |

Values of attributes BID and HC were divided into proper categories.

The BID analysis – intervals calculation (Fig. 4)

$$\begin{aligned}
 l_p &= 1,591421 \\
 q_s &= 1,380758 \\
 t_1 &= 532 & x_1 &= 79,81705 \\
 t_2 &= 734 & x_2 &= 224,4356
 \end{aligned}$$

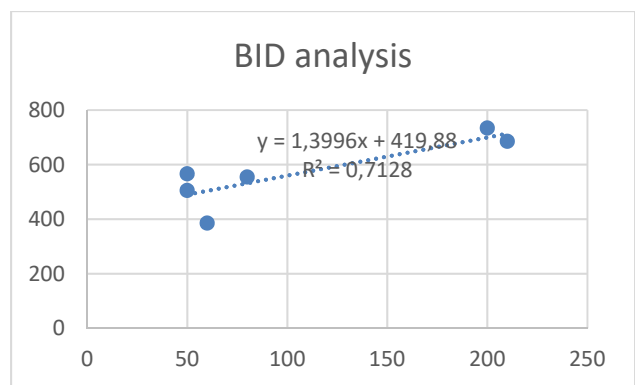


Figure 4. The BID analysis

The HC analysis – intervals calculation (Fig. 5)

$$\begin{aligned}
 l_p &= 1,591421 \\
 q_s &= 1,380758 \\
 t_1 &= 532 & x_1 &= 2,167139 \\
 t_2 &= 734 & x_2 &= 9,848951
 \end{aligned}$$

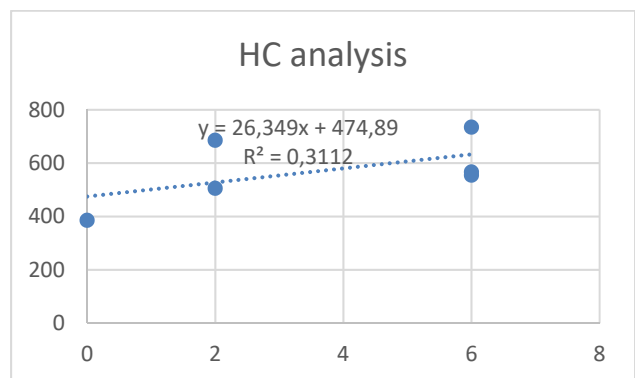


Figure 5. The HC analysis

And finally the training set is presented in table 4.

Table 4. The training set

| BID | HC | H | ST |
|-----|----|-----|----|
| <80 | ≤2 | No | A |
| ≥80 | >2 | Yes | B |

| | | | |
|-----|----|-----|---|
| <80 | ≤2 | Yes | A |
| <80 | >2 | Yes | B |
| ≥80 | >2 | Yes | B |
| ≥80 | ≤2 | Yes | B |

Entropy calculation according to ID3 algorithm:

$$I = 0,918296$$

$$I(\text{BID}) = 0,540852$$

$$I(\text{HC}) = 0,540852$$

$$I(\text{H}) = -0,33333$$

The decision tree with the use of ID3 algorithm is presented in figure 6.

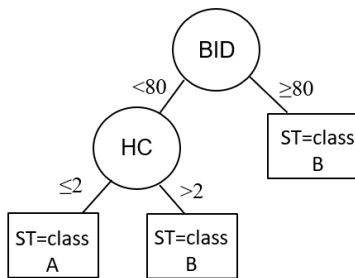


Figure 6. Decision tree

The rules induced based on the decision tree are:

IF $BID < 80$ and $HC \leq 2$ THEN $ST = \text{class A}$

IF $(BID < 80$ and $HC > 2)$ or $BID \geq 80$ THEN $ST = \text{class B}$

4. Conclusions

The presented approach is focused on time standard setting for assembly. The proposed rule based approach is useful in expert systems which can be joined with ERP systems. The proposed approach included the following steps: decomposing the assembly process into subassemblies, calculating time standards for typical subassemblies, assigning time standards to classes, developing a training set with attributes and values characterising subassemblies, building a decision tree, formulating decision rules for assembly time standard prediction which can be applied for assembly as well as for disassembly tasks. In the presented approach, it is especially important to identify the attributes which can be used in the modelled assembly process, and to divide the values into intervals. Too wide intervals can have a negative influence on the preciseness of the results. Too narrow intervals can cause difficulty in developing a decision tree which is easy to use. A decision tree can be updated when new cases in the training set are added.

Literature

- [1] Zandin K.: Maynard's Industrial Engineering Handbook, Mc Graw-Hill, 2001
- [2] Kutschenreiter I.: Neural network application for time standards setting in assembly and disassembly. Journal of Machine Engineering, 20 (3):106-116, 2020

- [3] Chen Z., Bao J., Zheng X., Liu T.: Assembly Information Model Based on Knowledge graph. J.Shanghai Jiao Tong Univ. (Sci.), 25(5): 578-588, 2020
- [4] Seabra Lopes L., Camarinha-Matos L.M.: Planning, Training and Supervision of Learning in Flexible Assembly Systems, in L. Camarinha-Matos et al. (eds.), Balanced Automation Systems, © Springer Science+Business Media Dordrecht, 1995
- [5] Groover M.: Work Systems, Pearson, 2016
- [6] Wołk R., Strzelecki T.: Badanie metod i normowanie pracy. Wydawnictwo Politechniki Warszawskiej, Warszawa, 1993
- [7] Young, H. T., Tsai, D. H. An integrated expert operation planning system with a feature-based design model. International Journal of Advanced Manufacturing Technology, 9:305-310, 1994
- [8] Wang X., Nauck D.D., Spott M., Kruse R.: Intelligent data analysis with fuzzy decision trees. Soft Computing, 11:439-457, 2007
- [9] Voisine N., Boullé M., Hue C.: A: Bayes Evaluation Criterion for Decision Trees. Published in: F. Guillet et al. (Eds.): Advances in Knowledge Discovery and Management, SCI 292: 21-38. Springer-Verlag Berlin Heidelberg, 2010
- [10] Micheal N.: Artificial intelligence: a guide to intelligence systems. Addison Wesley, Great Britain, 2002
- [11] Zhou BH, Xi LF, Cao YS.: A beam-search-based algorithm for the tool switching problem on a flexible machine. The International Journal of Advanced Manufacturing Technology, 25(9-10):876-882, 2005
- [12] Kuo Y., Lin K.-P.: Using neural network and decision tree for machine reliability prediction. The International Journal of Advanced Manufacturing Technology, 50:1243-1251, 2010
- [13] Chandra I, Sivakumar N, Gokulnath CB., Parthasarathy P.: IoT based fall detection and ambient assisted system for the elderly. Cluster Computing volume 22:2517-2525, 2019
- [14] Varadharajan R., Priyan M., Parthasarathy Panchatcharam K., Vivekanandan S., Gunasekaran M.: A new approach for prediction of lung carcinoma using back propagation neural network with decision tree classifiers. Journal of Ambient Intelligence and Humanized Computing. Springer-Verlag GmbH Germany, 2018
- [15] Rokach L., Maimon O.: Decision Trees. Data Mining and Knowledge Discovery Handbook. Springer US, 2005
- [16] Chrystolouris G.: The Design of Manufacturing Systems. In: Manufacturing Systems: Theory and Practice. Mechanical Engineering Series. Springer, New York, 2006

Gestaltung von Anleitungen für die Montage variantenreicher Erzeugnisse

Sven Hinrichsen

Technische Hochschule Ostwestfalen-Lippe, Labor für Industrial Engineering,
Campusallee 12, 32657 Lemgo

Die manuelle Montage variantenreicher Erzeugnisse verlangt von einem Montagebeschäftigten, dass er sich mit jedem Auftragswechsel auf die veränderten Anforderungen einstellt. Je mehr unterschiedliche Erzeugnisse von einem Beschäftigten zu montieren sind, desto mehr Informationen sind vom Beschäftigten aufzunehmen und zu verarbeiten. Mit steigender Komplexität der Montageaufgabe helfen Montageanleitungen dabei, dass die Tätigkeitsausführung effektiv und effizient erfolgen kann und gleichzeitig eine Überbeanspruchung des Beschäftigten vermieden wird. In diesem Beitrag werden daher wichtige Hinweise zur Gestaltung von Montageanleitungen vorgestellt.

1. Einordnung und Bedeutung

Die manuelle Montage variantenreicher Erzeugnisse verlangt von einem Beschäftigten, dass er sich mit jedem Auftragswechsel auf die veränderten Anforderungen einstellt. Je mehr unterschiedliche Erzeugnisse bzw. Erzeugnisvarianten von einem Beschäftigten zu montieren sind, desto mehr Informationen sind vom Beschäftigten aufzunehmen und zu verarbeiten. Die Komplexität der Aufgabenbearbeitung nimmt zu. Weitere Komplexitätstreiber sind eine hohe Anzahl verschiedener, zu montierender Komponenten und eine hohe Dynamik an Änderungen einzelner Erzeugnisse während ihres Lebenszyklus. Darüber hinaus wird die Komplexität der Montageaufgabe, die im Kern aus Handhabungs- und Füge-tätigkeiten besteht, durch die Art und Anzahl der zusätzlich auszuführenden Tätigkeiten beeinflusst (z. B. Aufspielen einer Firmware, Prüfen einzelner Erzeugnisfunktionen, Kennzeichen und Verpacken des Erzeugnisses) [1].

Mit zunehmender Komplexität der Montageaufgabe steigen die kognitiven Anforderungen an den Montagebeschäftigten. Das Entwickeln von Routinen wird deutlich erschwert. Jeder Montageauftrag beinhaltet spezifische Anforderungen und verlangt nach einem eigenen „kognitiven Mindset“ des Beschäftigten [2]. Um die mit dieser Komplexität verbundenen Probleme – wie Orientierungs- und Suchzeiten sowie Fehlhandlungen mit ihren negativen Auswirkungen auf die Arbeitsproduktivität [3] und Beanspruchung der Beschäftigten [4] – zu vermeiden, lassen sich zwei Strategien des Komplexitätsmanagements in der Montage unterscheiden [2].

Die erste Strategie zielt auf eine Reduktion von Komplexität ab, indem etwa die Arbeitsteilung erhöht und die Standardisierung von Arbeitsprozessen forciert wird. In der Folge entwickeln sich durch wiederholende Vorgänge bei den Beschäftigten Routinen heraus, die die Arbeitsproduktivität positiv beeinflussen können [5]. Mit zunehmender Vielfalt der zu montierenden Erzeugnisse und kleiner werdenden Losgrößen sind diesem Ansatz

der Arbeitsteilung aber Grenzen gesetzt, so dass ein weiterer Ansatz darin besteht, Komplexität bereits in den frühen Phasen der Produktentstehung zu reduzieren bzw. zu vermeiden. So lässt sich durch eine modulare Gestaltung der Produktstruktur die Vielfalt der Bauteile und Baugruppen in der Montage deutlich reduzieren, ohne dass es zu Leistungseinschränkungen gegenüber den Kunden kommt [6, 7]. Dieser Ansatz wird auch als Plattform- oder Modulbaukastenstrategie bezeichnet und führt auch in Kombination mit der Anwendung von Prinzipien der montagegerechten Produktgestaltung (z. B. Minimierung der Teileanzahl, Verwendung von Funktionselementen der Selbstzentrierung) zu einer Verringerung der Komplexität und der mit ihr verbundenen Kosten [8].

Mit dem Trend zur Digitalisierung und der Etablierung einer industriell ausgerichteten Individualmontage ist eine strategische Neuausrichtung verbunden. Das Management von Komplexität zielt weniger darauf ab, Komplexität zu reduzieren. Vielmehr geht es bei dieser zweiten Strategie darum, Komplexität explizit zuzulassen und für die Beschäftigten beherrschbar zu machen [2]. Dazu werden Ursachen von Komplexität identifiziert und Maßnahmen zu ihrer Bewältigung abgeleitet [9]. Im Zentrum stehen dabei informatorische Montageassistenzsysteme, die den Beschäftigten im Montageprozess mittels Montageanleitungen unterstützen und so auch die zunehmende Variantenvielfalt sicher beherrschbar machen [10, 2].

2. Begriff und Merkmale von Montageanleitungen

Montageanleitungen, die auch als Montageinstruktionen bezeichnet werden, zielen darauf ab, Informationen zu Montageauftrag und -ablauf so auszuwählen und darzustellen, dass die sensorische Aufnahme und Verarbeitung dieser Informationen zu einer effektiven (Fehlervermeidung) und effizienten (geringe Ausführungszeit) Auftragsbearbeitung führen und gleichzeitig die mittels

der Anleitung gewährte Unterstützung von den Beschäftigten in hohem Maße als nützlich empfunden wird (hohe Gebrauchstauglichkeit).

Grundsätze der Gestaltung von Montageanleitungen können sich in Anlehnung an DIN EN 82079-1:2013-06 [11] auf Inhalt, Informationsbereitstellung, Medium und Technikeinsatz sowie die Organisation beziehen. Daher werden diese vier Kategorien zunächst erläutert. Der Inhalt von Montageanleitungen ist zwar in hohem Maße abhängig vom Montageauftrag und dem zu montierenden Erzeugnis, allerdings lassen sich generelle Anforderungen an den Inhalt einer Montageanleitung aufstellen. Auch wenn alle wichtigen Inhalte zu einem Montageauftrag in einer Instruktion dargestellt werden, kann diese Instruktion trotzdem vom Montagebeschäftigten als wenig nützlich empfunden werden, weil Informationen nicht so aufbereitet sind, dass dieser sie zügig finden, erfassen und richtig interpretieren kann. Daher sind Regeln zur anforderungsgerechten Informationsbereitstellung zu beachten. Informationen können dabei als interpretierbarer Code oder als interpretierbare Daten oder Signale (z. B. Begriffe, Anzeigen) aufgefasst werden [12]. Daten, Signale oder Codes sind zwar objektiv wahrnehmbar, aber nur verwertbar, wenn sie vom Beschäftigten interpretiert werden können. Die Interpretation durch den Menschen kann daher als Decodierung bezeichnet werden [12]. Mit der Decodierung entstehen Informationen. Voraussetzung für eine Decodierung ist, dass der Mensch über sogenannte Signalinventare im Gedächtnis verfügt [12], also in der Lage ist, tätigkeitsleitende Signale oder Daten richtig zu deuten. Je abstrakter und umfangreicher die übermittelten Informationen bzw. Daten sind, desto länger dauert der Decodierungsprozess. Daher besagt das Kompatibilitätsprinzip [13], Daten bzw. Informationen so bereitzustellen, dass der Beschäftigte die zur Aufgabenausführung benötigten Informationen in möglichst kurzer Zeit identifizieren und richtig interpretieren kann, um dann die passende Handlung fehlerfrei auszuführen. Dabei ist auch ein geeignetes Medium mit der entsprechenden Technik auszuwählen. Durch dieses Medium wird festgelegt, wie die Kommunikation der bereitgestellten Informationen an den Empfänger, den Montagebeschäftigten, erfolgt und wie dieser über das Medium interagiert. Um inhaltlich angemessene und ergonomisch aufbereitete Montageanleitungen über ein geeignetes Medium den Beschäftigten bereitzustellen, bedarf es einer anforderungsgerecht gestalteten Organisation, der vierten Kategorie von Gestaltungsgrundsätzen. Nachfolgend werden zu diesen vier miteinander zusammenhängenden Kategorien wichtige Prinzipien der Gestaltung von Montageinstruktionen dargestellt.

3. Inhalte

Wichtige inhaltliche Anforderungen an Montageinstruktionen beziehen sich auf die Relevanz, Aktualität, Übersichtlichkeit, Akzentuierung und Vollständigkeit. Unter Relevanz wird verstanden, dass – nach dem Grundsatz

„weniger ist mehr“ – nur Informationen ausgegeben werden, die für die Durchführung der Montageaufgabe erforderlich sind. Aktualität bedeutet, dass eine Montageinstruktion stets den gleichen Revisionsstand aufweisen muss wie die ihr zugrundeliegenden Produktdaten (z. B. Stückliste, CAD-Zeichnung) [14].

Übersichtlichkeit bezieht sich an dieser Stelle nicht auf die Art der Informationsbereitstellung (s. Abschn. 4), sondern auf die Ausgabe von Informationen, die dazu dienen, dem Montagebeschäftigten zu Beginn eine Übersicht zum Montageauftrag zu geben (z. B. Bild des Erzeugnisses, Stückliste, Losgröße, Liste der Montagevorgänge) und während der Auftragsbearbeitung den Bearbeitungsstatus auszugeben (z. B. Schritt 5 von 30 Montageschritten). Akzentuierung beinhaltet die Ausgabe von gesonderten Informationen, um bei einzelnen Montageschritten zum Beispiel auf Fehlermöglichkeiten, qualitätskritische Vorgänge oder Verletzungsgefahren hinzuweisen. Diese Zusatzinformationen sind gesondert – zum Beispiel mit Hilfe von Bildzeichen (s. Abschnitt 4) – kenntlich zu machen.

Unter Vollständigkeit wird in Anlehnung an Kothes [15] verstanden, dass eine Montageanleitung sämtliche Informationen enthalten muss, die den Montagebeschäftigten in die Lage versetzen, einen Montageauftrag korrekt auszuführen. Bei dieser Anforderung ergibt sich die Herausforderung, dass Beschäftigte über unterschiedliche Qualifikationen und Vorerfahrungen verfügen, aus denen interindividuell unterschiedliche Informationsbedarfe resultieren [15]. Erschwerend kommt hinzu, dass im Zeitverlauf die Informationsbedarfe einer Person durch Prozesse des Lernens und Vergessens variieren können [14]. Der Nutzen einer Montageinstruktion, die Kompetenzentwicklung von Beschäftigten und die Komplexität der Montageaufgabe stehen dabei in enger Wechselwirkung zueinander [1]. So führt die Zunahme an Erfahrung des Beschäftigten mit der Montageaufgabe zu einer Verringerung der wahrgenommenen Aufgabenkomplexität. In der Folge verliert eine detaillierte Montageinstruktion aus Sicht des Beschäftigten an Nutzen und kann sogar als störend empfunden werden und keine Akzeptanz mehr erfahren [1]. Informationen sollten daher möglichst dynamisch auf die Bedarfe einzelner Nutzer oder Nutzergruppen angepasst werden. Wer über umfangreiche Erfahrungen mit der Montage eines bestimmten Produktes verfügt, benötigt im Regelfall deutlich weniger Informationen als ein Beschäftigter, der diesen Produkttyp erstmalig montiert. Eine pragmatische Umsetzung dieser Anforderung kann zum Beispiel erfolgen, indem eine Montageanleitung eine Kurzform (Übersicht mit allen relevanten Informationen für einen erfahrenen Beschäftigten) und eine Langform (z. B. Schritt-für-Schritt-Anleitung mit allen wichtigen Informationen für einen wenig erfahrenen Beschäftigten) enthält [14].

4. Informationsbereitstellung

Anforderungen an die Informationsbereitstellung beinhalten insbesondere die Art der Informationsausgabe, die Strukturierung von Informationen, die (Mehrfach-) Codierung sowie die Gestaltung einzelner Elemente der Informationsausgabe.

Die Art der Informationsausgabe kann optisch, akustisch und/oder taktil erfolgen. Die Fokussierung auf rein visuelle Informationsausgaben, wie sie bei der Gestaltung von Montageanleitungen üblich sind, kann nach der Theorie der multiplen Ressourcen [16] zu Engpässen in der Ressourcenbeanspruchung führen. Vielmehr kann es unter Gesichtspunkten der Reduzierung von Beanspruchung und der Steigerung von Leistung hilfreich sein, eine optische beispielsweise mit einer akustischen Ausgabe zu kombinieren. So könne durch eine „multi-sensorische Kombination von Sinneseindrücken“ häufig die menschliche Leistung positiv beeinflusst werden [17]. Vor diesem Hintergrund sollten Montageinstruktionen künftig in höherem Maße als integraler Bestandteil von multimodal gestalteten Assistenzsystemen angesehen werden.

Aufgrund der großen Informationsmengen, die in der variantenreichen Montage von komplexen Erzeugnissen über Montageinstruktionen an Beschäftigte auszugeben sind, kommt der Strukturierung dieser Informationen eine große Bedeutung zu. Dabei werden Informationen zu Montageauftrag bzw. -aufgabe zumeist in einzelne Ablaufabschnitte, die auch als Sequenzen bezeichnet werden, unterteilt. Der Montagebeschäftigte wird über die Montageinstruktion Sequenz für Sequenz durch den Montageprozess geführt, indem nach Ausführung der Tätigkeiten einer Sequenz ausgegeben wird, was die Inhalte der nächsten Sequenz sind. Der Tätigkeits- und damit Informationsumfang einer Sequenz sollte sich dabei insbesondere an den Anforderungen der Beschäftigten orientieren. Für die Strukturierung der Informationen einzelner Ablaufabschnitte ist bei visuellen Montageinstruktionen ein einheitliches Layout zu entwickeln. Dabei ist zu beachten, dass eine sogenannte Mehrfachcodierung dem Beschäftigten zusätzliche Orientierung bieten kann [12, 13]. Damit ist beispielsweise gemeint, dass zusätzlich zu einer Zeichnungsinformation auch ein kurzer Texthinweis gegeben wird, um mögliche Unsicherheiten des Beschäftigten zur auszuführenden Verrichtung – in diesem Fall durch eine Doppelcodierung – weitgehend auszuschließen. Daher sind bei der Layoutgestaltung unterschiedliche Elemente (z. B. Bild, Texthinweis, Fortschrittsbalken) zu berücksichtigen und in eine feste Anordnung zu bringen.

Darüber hinaus können Codierungen – etwa über Farbe oder Bildzeichen – die Informationsaufnahme erleichtern. So kann beispielsweise die farbliche Codierung des im nächsten Schritt zu montierenden Bauteils – innerhalb einer Schwarz-Weiß-Darstellung des Gesamter-

zeugnisses – helfen, das Bauteil auf Antrieb an der richtigen Stelle zu befestigen [12]. Ebenfalls können Bildzeichen, die auch als Piktogramme bezeichnet werden, die Informationsaufnahme deutlich erleichtern [13, 14] oder die Aufmerksamkeit des Nutzers auf einen wichtigen Sachverhalt lenken [18]. Dabei kann es sinnvoll sein, Bildzeichen mit Textinformationen – im Sinne einer Doppelcodierung – zu kombinieren. Eine Übersicht mit Erklärungen zu allen, in Montageinstruktionen verwendeten Bildzeichen kann helfen, neuen Beschäftigten die Bedeutung einzelner Bildzeichen zu vermitteln. Bildzeichen können beispielsweise das aufzunehmende Werkzeug anzeigen oder einen Warnhinweis übermitteln. Textinformationen sollten möglichst nur in Kombination mit Bild- oder Zeichnungsinformationen verwendet werden. Texte sind verständlich, präzise und mittels einer einheitlichen Terminologie zu formulieren [15]. Dabei sind eine genügend große Schrift und eine gut lesbare Schriftart zu wählen [15]. Im Falle einer digitalen Informationsausgabe sollte der Benutzer die Schriftgröße in der Software ändern können. In Bezug auf die verwendeten Fachbegriffe kann es hilfreich sein, ein Glossar zu erstellen [18].

5. Medium und Technikeinsatz

Montageinstruktionen können über Arbeitspapiere oder in Form von rechnergestützt aufbereiteten Informationen ausgegeben werden. Zudem können ergänzend Baumuster – zum Beispiel einzelner Baugruppen – bereitgestellt werden. Sollen Instruktionen in digitaler Form verwendet werden, so lassen sich als visuelle Ausgabegeräte Bildschirme und Tablets, Licht- und Laserprojektoren sowie Wearables einsetzen. Letztgenannte Ausgabegeräte können wiederum nach den Körperteilen, von denen sie getragen werden, klassifiziert werden. Relevant sind derzeit vor allem Kopf (Augmented Reality-Brille/ Datenbrille/ „smart glasses“), Hand („smart glove“) und Handgelenk („smart watch“). Einzelne Ausgabesysteme sind in der Regel Bestandteil von informatorischen Assistenzsystemen, die derzeit mit einer hohen Dynamik entwickelt werden und die über eine Vielzahl an Funktionen verfügen können (z. B. Quittieren von qualitätskritischen Vorgängen, automatische Fehlererkennung über Bildverarbeitung, Integration von Schraubern) [14]. Die Integration spielerischer Elemente („Gamification“) in den Montageprozess kann die Motivation der Beschäftigten positiv beeinflussen [19].

6. Organisation

Um die mit der Gestaltung von Montageanleitungen verbundenen Ziele zu erreichen (s. Abschnitt 2), sind entsprechende organisatorische Voraussetzungen zu schaffen. Diese unterscheiden sich von Betrieb zu Betrieb. Allerdings lassen sich allgemeingültige Hinweise formulieren.

Montageanleitungen sind einheitlich zu gestalten. Dazu bedarf es eines Leitfadens, in dem sämtliche Richtlinien

zu Inhalten, Informationsaufbereitung und Techniknutzung darzustellen sind. Um eine hohe Gebrauchstauglichkeit der Instruktionen zu gewährleisten, sind in Anlehnung an die in DIN EN ISO 9241-210:2020-03 [20] beschriebene Vorgehensweise insbesondere die Montagebeschäftigten als Nutzer in den Prozess der Entwicklung solcher Standards einzubeziehen. Ausgehend von dem entwickelten Leitfaden ist der Prozess der Gestaltung von Montageanleitungen zu rationalisieren. Zielsetzung muss es dabei sein, Montageanleitungen modular aufzubauen und aus bestehenden Auftrags- und Erzeugnisdaten automatisch zu generieren.

Gleichzeitig bedarf es eines Prozesses, der festlegt, wie die Montagebeschäftigten als Nutzer während ihrer Arbeit Informationen zu den von ihnen identifizierten Fehlern oder Defiziten an einzelnen Instruktionen schnell und unkompliziert – beispielsweise über einen Workflow – an die entsprechende Stelle im Industrial Engineering weiterleiten können [21].

Um einerseits gebrauchstaugliche Instruktionen zu Abläufen in der Montage gestalten zu können und andererseits Prozesse der Generierung und Bereitstellung von Informationen zu automatisieren, bedarf es unterschiedlicher Kompetenzen. Ein Ansatz kann sein, die Funktion des Industrial Engineering, die über ein hohes Maß an Prozesswissen verfügt, stärker interdisziplinär auszurichten und um Kompetenzen aus dem Bereich der Informationstechnik und der kognitiven Ergonomie zu ergänzen. In Unternehmen, die über eine Mehrzahl an Produktionswerken verfügen, kann eine solche Funktion zentral angesiedelt werden, um entsprechende Synergien zu nutzen.

7. Fazit

Viele Informationen sind in der manuellen Montage variantenreicher Erzeugnisse den Beschäftigten über Montageanleitungen bereitzustellen. Diese zielen darauf ab, Informationen zu Montageauftrag und -ablauf so auszuwählen und darzustellen, dass die sensorische Aufnahme und Verarbeitung dieser Informationen zu einer effektiven (Fehlervermeidung) und effizienten (geringe Ausführungszeit) Auftragsbearbeitung führen und gleichzeitig die mittels der Anleitung gewährte Unterstützung von den Beschäftigten in hohem Maße als nützlich empfunden wird (hohe Gebrauchstauglichkeit). Für Industriebetriebe, die über große Bereiche mit manueller Montage verfügen, kann es hilfreich sein, betriebs-spezifische Richtlinien zur Entwicklung von Montageanleitungen und zu ihrer Integration in Assistenzsysteme in einem Leitfaden zusammenstellen. Diese Richtlinien sollten sich auf Inhalte, Informationsbereitstellung, Medium und Technikeinsatz beziehen. Ziel muss es sein, durchgängige digitale Prozessketten – von den Auftrags- und Erzeugnisdaten bis zur Ausgabe von Montageinstruktionen an einzelnen Arbeitsplätzen – zu realisieren. Dafür sind die organisatorischen Voraussetzungen zu

schaffen, indem insbesondere die Kompetenzen des Industrial Engineering weiterentwickelt werden. Auch bedarf es der Etablierung eines kontinuierlichen Verbesserungsprozesses, um bestehende Standards zu Montageinstruktionen regelmäßig unter Einbeziehung der Beschäftigten zu evaluieren und zu optimieren.

Literaturverzeichnis

- [1] M. Bornewasser, D. Bläsing, S. Hinrichsen, Informatrische Assistenzsysteme in der manuellen Montage: Ein nützliches Werkzeug zur Reduktion mentaler Beanspruchung? *Zeitschrift für Arbeitswissenschaft*, 72 (2018) 4, 264–275. <https://doi.org/10.1007/s41449-018-0123-x>
- [2] S. Hinrichsen, M. Bornewasser, Veränderung der Gestaltungsparadigmen industrieller Montagearbeit. In: M. Bornewasser, S. Hinrichsen (Hrsg.): *Informatrische Assistenzsysteme in der variantenreichen Montage – Theorie und Praxis*. Springer 2020, 1–20. https://doi.org/10.1007/978-3-662-61374-0_1
- [3] E. Haller, O. Heer, E. F. Schiller, Innovation in Organisation schafft Wettbewerbsvorteile. *FB/IE*, 48 (1999) 1, 8–17.
- [4] D. Bläsing, Mentale Beanspruchung in der Montage. In: M. Bornewasser, S. Hinrichsen (Hrsg.): *Informatrische Assistenzsysteme in der variantenreichen Montage – Theorie und Praxis*. Springer 2020, 65–87. https://doi.org/10.1007/978-3-662-61374-0_4
- [5] B. Lotter, Manuelle Montage von Kleingeräten. In B. Lotter, H.-P. Wiendahl (Hrsg.): *Montage in der industriellen Produktion*. 2. Aufl. Springer 2012, 109–146. <https://doi.org/10.1007/978-3-642-29061-9>
- [6] H. Wildemann, Produktivitätsverbesserung: Leitfaden zur kurzfristigen Produktivitätssteigerung in kleinen und mittleren Unternehmen. Bd. 52, TCW 2014.
- [7] D. Krause, N. Gebhardt, Zusammenhänge und Auswirkungen der Vielfalt. In: *Methodische Entwicklung modularer Produktfamilien*. Springer Vieweg 2018, 19–65. <https://doi.org/10.1007/978-3-662-53040-5>
- [8] G. Spur, H.-J. Helwig, Einführung in die Montage-technik. In: G. Spur, T. Stöferle (Hrsg.): *Handbuch der Fertigungstechnik*. Bd. 5, Carl Hanser 1986, 591–606.
- [9] B. Brinzer, A. Banerjee, Komplexitätsbewertung im Kontext Cyber-physischer Systeme. *ZWF*, 112 (2017) 5, 341–345. <https://doi.org/10.3139/104.111709>
- [10] S. Mattsson, Å. Fast-Berglund, How to Support Intuition in Complex Assembly? *Procedia CIRP*, 50 (2016) 624–628. <https://doi.org/10.1016/j.procir.2016.05.014>
- [11] Deutsches Institut für Normung e.V. (2012) Erstellen von Gebrauchsanleitungen – Gliederung, Inhalt und Darstellung - Teil 1: Allgemeine Grundsätze

- und ausführliche Anforderungen (IEC 82079-1:2012), DIN EN 82079-1:2013-06.
- [12] W. Hacker, Arbeitspsychologie – Psychische Regulation von Arbeitstätigkeiten. Hans Huber 1986.
- [13] H. Strasser, Kompatibilität. In: T. Hettinger, G. Wobbe (Hrsg.): Kompendium der Arbeitswissenschaft. Kiehl 1993, 228–242.
- [14] S. Hinrichsen, Informatrische Gestaltung der Montage mittels Assistenzsystemen. In: M. Bornewasser, S. Hinrichsen (Hrsg.): Informatrische Assistenzsysteme in der variantenreichen Montage – Theorie und Praxis. Springer 2020, 21-42.
https://doi.org/10.1007/978-3-662-61374-0_2
- [15] L. Kothes, Grundlagen der Technischen Dokumentation – Anleitungen verständlich und normgerecht erstellen. Springer 2011.
<https://doi.org/10.1007/978-3-642-14668-8>
- [16] C. D. Wickens, Multiple Resources and Mental Workload. *Human Factors*, 50 (2008) 3, 449–455.
<https://doi.org/10.1518/001872008X288394>
- [17] C. Schlick, R. Bruder, H. Luczak, Arbeitswissenschaft. 4. Aufl. Springer Vieweg 2018.
<https://doi.org/10.1007/978-3-662-56037-2>
- [18] D. Juhl, Technische Dokumentation – Praktische Anleitungen und Beispiele. 3. Aufl. Springer Vieweg 2015.
- [19] C. Petzoldt, D. Keiser, T. Beinke, M. Freitag, Functionalities and implementation of future informational assistance systems for manual assembly. In: International Conference on Subject-Oriented Business Process Management, December 2020, Springer, 88–109.
- [20] Deutsches Institut für Normung e.V. (2020) Ergonomie der Mensch-System-Interaktion - Teil 210: Menschzentrierte Gestaltung interaktiver Systeme (ISO 9241-210:2019), DIN EN ISO 9241-210:2020-03.
- [21] B. Adrian, S. Hinrichsen, A. Schulz, E. Voß, Low-Code-Programmierung als Ansatz zur Gestaltung bedarfsgerechter informatrischer Assistenzsysteme – eine Fallstudie. In: M. Bornewasser, S. Hinrichsen (Hrsg.): Informatrische Assistenzsysteme in der variantenreichen Montage – Theorie und Praxis. Springer 2020, 173–186.
https://doi.org/10.1007/978-3-662-61374-0_9

Ergonomie 4.0 – Herausforderungen und Chancen der Industrie 4.0 für die Gestaltung menschengerechter Arbeitsplätze

Lars Fritzsche

imk automotive GmbH, Amselgrund 30, 09128 Chemnitz

Eine der großen Herausforderungen in der Industrie 4.0 ist die anforderungsgerechte Integration des Menschen. Es wird auch weiterhin Aufgaben in der Produktion geben, die seine hohe Flexibilität und manuelle Geschicklichkeit erfordern. Dabei kommen zunehmend neuartige technische Assistenzsysteme zum Einsatz, wie z.B. kollaborierende Roboter und Exoskelette. Diese sollen den Menschen im Fertigungsprozess optimal unterstützen. Dazu müssen sie möglichst ergonomisch gestaltet werden, um eine effiziente Mensch-Maschine-Interaktion sicherzustellen. Die menschengerechte Arbeitsgestaltung ist somit weiterhin von großer Bedeutsamkeit, auch wenn sich die Rolle des Menschen im Rahmen von Industrie 4.0 und der voranschreitenden Digitalisierung in der Fertigung verändert. Anhand von drei Beispielen wird aufgezeigt, wie dieses mit Hilfe von digitalen Menschmodellen und 3D-Simulation gelingen kann.

1. Einleitung

Die vierte industrielle Revolution („Industrie 4.0“) ist bereits seit einigen Jahren in vollem Gange und verändert auch die Arbeitswelt im produzierenden Gewerbe enorm. Neue smarte Technologien und deren immanente Vernetzung sollen die Fertigungsprozesse effizienter machen und ersetzen an einigen Stellen den Menschen. Doch der Automatisierungsgrad ist in vielen Bereichen immer noch überschaubar und die menschlichen Stärken wie Flexibilität, Geschicklichkeit und Problemlösefähigkeit sind immer noch gefragt. Daher ist es weiterhin erforderlich, Fertigungsprozesse anhand grundlegender Prinzipien menschengerechter Arbeit zu gestalten und frühzeitig eine ergonomische Interaktion zwischen Mensch und Technik sicherzustellen.

2. 3D-Simulation menschengerechter Arbeit

Bereits seit einigen Jahren kommen in der Planung und Gestaltung von Arbeitsplätzen sog. digitale Menschmodelle zum Einsatz. Sie ermöglichen die risikofreie Überprüfung eines geplanten Arbeitsprozesses hinsichtlich Ergonomie, Effizienz und Sicherheit anhand einer 3D-Simulation.

Die Software „ema Work Designer“, entwickelt von der imk automotive GmbH aus Chemnitz, beinhaltet ein autonom agierendes Menschmodell, mit dem auf Basis sog. Verrichtungsbausteine eine vollständige 3D-Ablaufsimulation in einer virtuellen Arbeitsumgebung erstellt werden kann (Leidholdt, Fritzsche & Bauer, 2016). Darüber hinaus ermöglicht die Software eine automatische Ermittlung von wesentlichen Kennzahlen der ergonomischen Belastung und Produktivität bzw. Fertigungszeit für manuelle Tätigkeiten. Sie kommt bereits in diversen Anwendungsfällen bei verschiedenen Kunden aus der Industrie und Wissenschaft zum Einsatz (Fritzsche, Ullmann, Bauer & Sylaja, 2019).

2.1. Beispiel 1: Partizipative Gestaltung von Arbeitssystemen mit kollaborierenden Robotern

Arbeitssysteme, in denen Mensch und Roboter kollaborativ miteinander interagieren (sog. MRK-Systeme), haben sich in der Praxis bisher kaum durchgesetzt. Eine gute Gestaltung erfordert nicht nur die Betrachtung von technischen und wirtschaftlichen Faktoren, sondern auch die Berücksichtigung der Nutzerperspektive. Das Projekt SOPHIA (“Socio-Physical Interaction Skills for Cooperative Human-Robot Systems in Agile Production“) hat zum Ziel, diese hybriden Arbeitssysteme weiterzuentwickeln, um die Produktivität, Flexibilität und Ergonomie in der Fertigung zu verbessern.

In einem ersten Anwendungsfall wurde ein Arbeitssystem in einem niederländischen KMU betrachtet, das Getriebe für verschiedene Industrieenanwendungen produziert. Mittels eines Quick-Checks

(<https://kompi.org/quickcheck/>) wurde zunächst ermittelt, für welchen Arbeitsplatz die größten Verbesserungspotentiale zu erwarten sind. Anschließend wurde eine 3D-Ablaufsimulation des aktuellen und künftigen Arbeitsprozesses erstellt. Anhand der Visualisierung wurde in einem iterativen Prozess zusammen mit Entwicklern und Betreibern des Systems das konkrete Zusammenwirken von Mensch und Roboter gestaltet (Abb. 1). Mittels der Simulation wurde u. a. sichergestellt, dass die Erreichbarkeit aller Teile gegeben ist und sicherheitskritische Abstände eingehalten werden können. Schließlich wurden die konkreten Verbesserungspotentiale des MRK-Systems hinsichtlich reduzierter Fertigungszeit und ergonomischer Belastung der Mitarbeiter ermittelt. Das Ergebnis wurde als Video veröffentlicht: <https://youtu.be/H-W0yy7LIA8>.

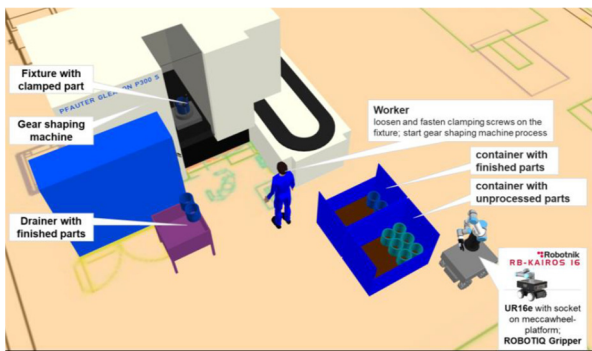


Abbildung 1: 3D-Modell des MRK-Arbeitssystems in SOPHIA

Um weitere Hinweise für die Systemgestaltung zu erhalten, wurde eine Befragung von Mitarbeitenden durchgeführt, die später in dem Arbeitssystem tätig sein werden. Hierbei kamen Ergebnisse der Simulation zum Einsatz, um den künftigen Nutzenden einen Eindruck des Systemkonzeptes zu vermitteln. Ziel dieser frühen Partizipation ist es, bereits während des Entwicklungsprozesses die Erwartungen an das robotische System zu erfassen und zu berücksichtigen (Fritzsche, Funk, Rosen, Spitzhahn & Wischniewski, 2021).

2.2. Beispiel 2: Biomechanische Bewertung der Wirksamkeit von Exoskeletten

In den letzten Jahren wurden von verschiedenen Herstellern Exoskelette für den industriellen Einsatz entwickelt und an den Markt gebracht. Sie können Mitarbeiter bei ergonomisch ungünstigen Arbeitsaufgaben wie Überkopftätigkeiten oder dem Heben schwerer Lasten unterstützen. In Labor- und Feldstudien wurde fortan versucht, die Wirksamkeit von Exoskeletten bezüglich der körperlichen Entlastung zu ermitteln (z.B. de Looze et al., 2016). Dabei gab es sehr unterschiedliche Ergebnisse und es konnten noch keine konkreten Effekte zur Prävention muskuloskeletaler Erkrankungen nachgewiesen werden (Steinhilber et al., 2020).

Biomechanische Simulationen mit Hilfe digitaler Menschmodelle können aufwendige Labor- und Feldstudien ergänzen und erweitern, indem sie auch die internen Kräfte und Gelenkmomente bei der Verwendung von Exoskeletten für eine Analyse zugänglich machen. Dafür kommt z.B. die Software AnyBody Modelling System (Damsgaard et al., 2006) in Frage, die für die Detailanalyse in Kombination mit dem ema Work Designer eingesetzt werden kann.

Im Rahmen des Projektes ANDY („Advancing Anticipatory Behaviors in Dyadic Human-Robot Collaboration“) wurde auf Basis einer umfangreichen Laborstudie (Maurice et al. 2020) eine Simulationsstudie mit der AnyBody Software durchgeführt. Dabei wurde das Exoskelett „Paexo Shoulder“ der Fa. Ottobock verwendet (Abb. 2). Erste Ergebnisse zeigen eine starke Reduktion der Belastungen im Schulterbereich, wo durch die Verwendung des Exoskeletts die Muskelaktivität und die Gelenkreaktionskräfte um bis zu 80 % reduziert werden konnten (Fritzsche, Spitzhahn & Gärtner, 2020). Gleichzeitig waren keine negativen Effekte im Bereich des Rückens und der

Lendenwirbelsäule zu erkennen. Das deutet darauf hin, dass die Kraftumleitung auf die unteren Extremitäten beim Paexo Shoulder sehr gut funktioniert.

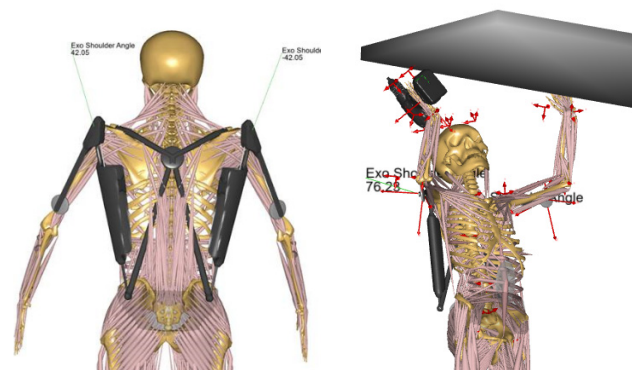


Abbildung 2: 3D-Modell des Exoskeletts „Paexo Shoulder“ in der Software AnyBody zur biomechanischen Detailanalyse

2.3. Beispiel 3: Fähigkeitsgerechte Arbeitsplatzgestaltung und adäquater Mitarbeiterereinsatz

Die steigende Anzahl älterer und leistungsgewandelter Arbeitnehmer ist eine zunehmende Herausforderung für produzierende Unternehmen. Zudem werden diese ebenfalls vor der Herausforderung gestellt, sich an neue Technologien und Assistenzsysteme zu gewöhnen. Zur Sicherstellung eines wertschöpfenden und fähigkeitsgerechten Einsatzes können ebenfalls digitale Menschmodelle verwendet werden. Sie sind insbesondere in der frühen Phase der Prozessplanung und zur präventiven Absicherung ergonomisch günstiger Arbeitsplätze geeignet. Jedoch berücksichtigen die üblichen Menschmodelle bislang nicht die Besonderheiten älterer und möglicherweise körperlich eingeschränkter Mitarbeiter, sondern beschränken sich auf die Abbildung des „Norm-Menschen“.

Im Projekt „Virtual Aging“ wurde daher das ema-Menschmodell um altersunabhängige Leistungsveränderungen und individuelle Fähigkeitsvoraussetzungen ergänzt (Ullmann & Fritzsche, 2020). Beispielsweise lassen sich alterstypische Einschränkungen der Beweglichkeit festlegen und bei der virtuellen Arbeitsgestaltung berücksichtigen. Ebenfalls lassen sich Arbeitsplatzanforderungsprofile erzeugen und mit Mitarbeiterfähigkeitsprofilen vergleichen (Abb. 3). Die Methodik wurde bereits teilweise in die aktuelle Version des ema Work Designers integriert und in Praxisprojekten, vornehmlich in der Automobilindustrie, eingesetzt. Weitere Teile befinden sich im Prototypenstatus und werden in den kommenden Releases veröffentlicht.



Abbildung 3: Workflow zur Verwendung von Mitarbeiterfähigkeitsprofil und Arbeitsplatzanforderungsprofil im ema Work Designer

3. Schlussfolgerung

Die Herausforderungen durch die rasante Veränderung der „Industrie 4.0“ sind vielfältig. Oftmals stehen vor allem die technischen Lösungen im Mittelpunkt der Betrachtung. Aber der Mensch bleibt ein entscheidender Faktor für flexible und qualitativ hochwertige Produktionssysteme. Der Beitrag zeigt Beispiele, wie neue Methoden der 3D-Simulation genutzt werden können, um prospektiv Arbeitsplätze zu gestalten, die zugleich effizient und ergonomisch sind.

Danksagung

Der Autor bedankt sich bei vielen engagierten Mitarbeitern in seinem Team, die zu diesen Arbeiten einen signifikanten Beitrag geleistet haben, besonders Sascha Ullmann, Michael Spitzhirn und Christian Gärtner. Ebenso geht der Dank an die Projektpartner, besonders das Istituto Italiano di Tecnologia, AnyBody Tech. und Ottobock. Die Arbeiten wurde zum Teil finanziert durch Fördergelder der Europäischen Kommission im Programm Horizon 2020 (SOPHIA und ANDY) sowie durch das Bundesministerium für Bildung und Forschung im Programm KMU Innovativ (Virtual Aging).

Literaturverzeichnis

- [1] W. Leidholdt, L. Fritzsche & S. Bauer (2016). Editor menschlicher Arbeit (ema) – Vom digitalen Menschmodell zum virtuellen Facharbeiter. In: A. C. Bullinger-Hoffmann and J. Mühlstedt (Hrsg.), Homo Sapiens Digitalis – Virtuelle Ergonomie und digitale Menschmodelle (pp. 355-362). Wiesbaden: Springer Vieweg.
- [2] L. Fritzsche, S. Ullmann, S. Bauer & V. J. Sylaja (2019). Task-based digital human simulation with Editor for Manual Work Activities – Industrial applications in product design and production planning. In: G. Paul and S. Scataglini (Eds.), DHM and Posturography (pp. 569-575). London, UK: Elsevier.
- [3] L. Fritzsche, M. Funk, P. Rosen, M. Spitzhirn & S. Wischniewski (2021). Simulation and participatory design of HRI work systems in the SOPHIA project. In: Gesellschaft für Arbeitswissenschaft (Hrsg.), Arbeit HUMAINE gestalten (Beitrag B.10.4). Dortmund: GfA-Press.
- [4] M. P. de Looze, T. Bosch, F. Krause, K. S. Stadler & L. W. O'Sullivan (2016). Exoskeletons for industrial

application and their potential effects on physical work load. *Ergonomics*, 59, 671-681.

- [5] B. Steinhilber, T. Luger, P. Schwenkreis et al. (2020). Einsatz von Exoskeletten im beruflichen Kontext zur Primär-, Sekundär-, und Tertiärprävention von arbeitsassoziierten muskuloskelettalen Beschwerden. *Zeitschrift für Arbeitswissenschaft*, 74, 227-246.
- [6] M. Damsgaard, J. Rasmussen, S. Tørholm, E. Surma & M. de Zee (2006). Analysis of musculoskeletal systems in the AnyBody Modeling System. *Simulation Modeling Practice and Theory*, 14, 1100-1111.
- [7] P. Maurice, J. Camernik, D. Gorjan et al. (2020). Objective and Subjective Effects of a Passive Exoskeleton on Overhead Work. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28, 152-164.
- [8] L. Fritzsche, M. Spitzhirn & C. Gärtner (2020). Analyse von Exoskeletten mit biomechanischer Simulation. *ASU Arbeitsmedizin Sozialmedizin Umweltmedizin*, 55, 470-473.
- [9] S. Ullmann & L. Fritzsche (2020). Entwicklung einer Methodik zur fähigkeitsgerechten Arbeitsgestaltung für leistungsgewandelte Mitarbeiter mit Hilfe digitaler Menschmodelle. In: Gesellschaft für Arbeitswissenschaft (Hrsg.), Digitaler Wandel, digitale Arbeit, digitaler Mensch? (Beitrag C.4.6). Dortmund: GfA-Press.

Leitidee Ökologie als Nachhaltigkeitsdimension in Lernprozessen für zukünftige Fachkräfte in der Kunststofftechnik – Beispielhafte Lern- und Arbeitsaufgabe zur Wahl von Fertigungsparametern mit einem Augmented Reality Lernsystem

Marvin Goppold¹, Simon Kreisler¹, Maren Klatt¹, Lukas Dunsche², Tobias Döhring², Sven Tackenberg³, Martin Frenz¹, Verena Nitsch¹

¹Institut für Arbeitswissenschaft der RWTH Aachen University, Eilfschornsteinstr. 18, 52062 Aachen

²Institut für Kunststoffwirtschaft, Johannes-Schuchen-Straße 4, 32657 Lemgo

³Labor für Industrial Engineering der Technischen Hochschule Ostwestfalen-Lippe, Campusallee 12, 32657 Lemgo

Augmented Reality bietet Chancen zur lernwirksamen Nutzung in Lernprozessen der dualen Ausbildung. Die Virtualisierung von simulierten Fehlerfolgen aus Handlungsfehlern erweitert betriebliche Arbeitsprozesse der Kunststofftechnik um eine augmentierte Dimension. Üblicherweise durch Ausbildungspersonal verhinderte Fehlerkonsequenzen lassen sich damit in Lernprozessen nutzen. Der Beitrag zeigt, wie Zielkonflikte didaktischer Leitideen z.B. der Nachhaltigkeit auf Grundlage von Arbeitsanalysen einer Füllstudie beim Spritzgießen für die Ableitung von Zielstellungen und Gestaltung eines betrieblichen Ausbildungsprozesses genutzt werden. Beispielhaft werden die Zielkonflikte anhand der Kühlzeit im Produktionsprozess beschrieben. Darauf aufbauend wird eine konkrete Zielstellung des Lernprozesses diskutiert. Das didaktische Konzept eignet sich sehr gut zur Förderung reflektierte Handlungskompetenz. Es eignet sich zur Fehlerdarstellung in Augmented Reality in betrieblichen Lernprozessen. Ein Transfer auf andere technische Ausbildungsbereiche mit deren Zielkonflikten ist möglich.

1. Einleitung

1.1 Motivation

Fehler ereignen sich wie in fast allen Bereichen des Lebens auch in der Arbeit. In der betrieblichen Ausbildung stellen Fehler gute Lerngelegenheiten für Auszubildende dar, um reale Auswirkungen einer falschen Parameterauswahl beim Spritzgießen zu erleben. Gutes Ausbildungspersonal unterbindet allerdings die Auswahl kritischer Parameter, wenn gefährliche Folgen für Gesundheit, Wirtschaftlichkeit oder Umwelt drohen.

Lernmedien nutzen Augmented Reality (AR) zur Ergänzung des Lernprozesses um eine virtuelle Dimension. Durch eine Erweiterung des realen Arbeitsprozesses um eine virtuelle Dimension kann es möglich werden, folgenschwere Fehler zu Lernzwecken in einer simulierten augmentierten Realität aufzuzeigen, vgl. Abbildung 1.



Abbildung 1: Augmented Reality Lernmedium im Einsatz bei Parametereingaben an einer Spritzgießmaschine

Dieser Beitrag zeigt einen Ausschnitt der didaktischen Konzeption im Rahmen einer so motivierten Lernmedienentwicklung [1] und geht spezifisch auf den Arbeitsprozess des Spritzgießens ein, wobei ein konkreter Parameter vor den didaktischen Vorgaben diskutiert wird und daraus Lernziele abgeleitet werden.

2. Didaktische Konzeption

2.1 Zielgruppe

Die Verwendung des Lernmediums ist primär für die betriebliche Ausbildung von Verfahrensmechaniker/innen für Kunststoff- und Kautschuktechnik als anerkannter Ausbildungsberuf vorgesehen [2], welche Arbeitstätigkeiten mit und an Spritzgießmaschinen ausführen [3]. Der Einsatz des Lernmediums wird für die betriebliche Ausbildung konzipiert.

2.2 Didaktische Leitideen

Angelehnt an den gestaltungsorientierten Didaktikansatz wird Technik als Teil der Gesellschaft und damit Bestandteil kultureller Auseinandersetzungen und Leitideen verstanden [4]. Deshalb müssen Auszubildende befähigt werden, Technik und damit verbunden ihre Arbeit mitzugestalten. Dies ist mit Motivation zur Gestaltung, Verantwortungsbewusstsein und fachlicher Befähigung als Teile beruflicher Handlungskompetenz verbunden [5]. Dabei stehen Auszubildende immer Zielen gegenüber, welche berücksichtigt werden sollen und zu Konflikten führen, sodass eine Lösung gefunden werden muss.

In der Veröffentlichung wird nur auf eine Auswahl an Zielen Bezug genommen. Bekannte Leitideen mit ihren Unterdimensionen, die Zielkonflikte im Produktionskontext darstellen, sind die Trilemmata des Projektmanagements [6] und der Nachhaltigkeit [7].

Ökonomie im Nachhaltigkeitsdreieck lässt sich demnach nochmals in Zeit, Kosten und Qualität unterteilen. Qualität wird typischerweise im Rahmen von Standards definiert [8]. Die verlässliche Erfüllung schwer zu erreichender Qualitätsdefinitionen stellt einen zentralen Wettbewerbsvorteil des Standortes Deutschland dar [9], weshalb diese Zielstellung besonders berücksichtigt werden sollte.

Fachkräfte haben einen erheblichen Einfluss auf die Qualität durch die Wahl der Verfahrensparameter im Fertigungsprozess. Deshalb ist dies ein wichtiger Bestandteil der Ausbildung [3]. Viele Stakeholder besitzen ein generelles Interesse an hoher Qualität, aber zeitgleich nehmen weitere Zielstellungen einen hohen Stellenwert ein.

Dies lässt sich an der gesellschaftlichen Tendenz zu nachhaltigerem Handeln hinsichtlich ökologischer Auswirkungen feststellen, was beispielsweise zu einer Förderung von energieeffizienten Fertigungsprozessen in der Kunststoffindustrie führt [10, p. 6]. Dies dient dem langfristigen Schutz der Umwelt und ist Bestandteil des Ausbildungsberufsbildes der Zielgruppe [3]. Die Befähigung zur Übernahme von ökologischer Verantwortung [11] bildet ein weiteres Bildungsziel der betrachteten dualen Ausbildung ab.

2.3 Didaktische Konzeption des Lernprozesses

Die Entwicklung von Handlungskompetenz stellt ein zentrales Anliegen in der beruflichen Bildung und somit auch in der entwickelten Lern- und Arbeitsaufgabe dar [12]. Um die intendierten Zielvorstellungen vor diesem Hintergrund umzusetzen, bedarf es einer genauen Organisation von Lernprozessen und der Gestaltung von Rahmenbedingungen, in denen Lernen stattfinden kann [13]. Als Basis der entwickelten Lern- und Arbeitsaufgabe dient das Hamburger-Lehr-Lern Modell [13]. Ferner ist der Lernprozess auf Grund der Lernförderlichkeit von Handlungsfehlern [14, 15] geprägt durch einen offenen und konstruktiven Umgang mit Fehlern.

Das Ziel der Lern- und Arbeitsaufgabe besteht darin, vor der Leitidee der Nachhaltigkeit und Ökonomie einen Zielkonflikt zu lösen: Es muss ein Lernprozess gestaltet werden, in welchem gleichzeitig die Nachhaltigkeit sowie Qualität der Teile, Zeitaufwand und Kosten im Fokus stehen. Ein solcher Konflikt ist besonders lernwirksam, da dieser erfordert, dass sich die Auszubildenden bei jeder Entscheidung bewusst machen müssen, welche Auswirkungen diese auf die unterschiedlichen Arbeitshandlungen besitzt. Schließlich müssen die Auszubildenden die Auswirkungen gegeneinander abwägen und innerhalb diesen Zielkonflikts eine Entscheidung treffen.

Durch die Nutzung von AR-Technologie können die Auszubildenden während des Lern- und Arbeitsprozesses Erfahrungen anhand von Fehlern machen, ohne dass das Ausbildungspersonal eingreift. Kostenintensive oder gefährliche Fehlerkonsequenzen werden dabei visualisiert. Dies wird dadurch ermöglicht, dass das AR-System über verschiedene Schnittstellen und Sensoren verfügt und somit die Handlungen der Auszubildenden erfassen, bei der Detektion von folgenschweren Fehlern die Maschine stoppt und die Fehlerkonsequenzen in AR darstellt. Fahren Auszubildende beispielsweise das Werkzeug in der Spritzgießmaschine zusammen, ohne vorher die Auswerfstifte einzufahren, können diese einen kostenintensiven Schaden erzeugen. In diesem Fall erkennt das AR-System die Fehlhandlung, stoppt die reale Maschine und zeigt eine Visualisierung von den abgebrochenen Auswerferstiften. Auszubildende können somit einen Fehler begehen und die Fehlerkonsequenzen erleben, ohne dass dabei reale und kostenintensive Konsequenzen entstehen.

Zusätzlich werden die erfassten Daten für das an die Durchführung anschließende Reflexionsgespräch automatisch aufbereitet. Anhand dieser Daten kann das Ausbildungspersonal im Reflexionsgespräch exakt nachvollziehen, welche Handlungen die Auszubildenden vorgenommen haben. Dies bietet die Möglichkeit individuell auf Fehlhandlungen einzugehen und diese mit den Auszubildenden aufzuarbeiten. Hierbei sollen die Auszubildenden nicht nur darin unterstützt werden, die vielfältigen Interdependenzen der Fertigungsparameter zu verstehen. Bei dieser Reflexion wird im Sinne der Leitideen insbesondere auf die Auswirkungen der Parametereinstellungen auf ökologische und ökonomische Faktoren eingegangen.

Dafür wird es notwendig, sich mit den Arbeitsprozessen beim Spritzgießen im Detail auseinanderzusetzen, um eine lernwirksame Lern- und Arbeitsaufgabe zu gestalten. Darauf aufbauend müssen Zielkonflikte im Arbeitsprozess ermittelt und konkretisiert werden, um die didaktischen Leitideen mit dem Arbeitsprozess zu verknüpfen und konkrete Lernziele formulieren zu können.

3. Analyse eines Arbeitsprozesses

Im Rahmen von Fallerhebungen [16] hat sich in Rücksprache mit Ausbildungspersonal der Arbeitsprozess Füllstudie als potentiell geeignet ergeben. Füllstudien sind Bestandteil der Tätigkeiten von Fachkräften bei der Fertigung von Kunststoffspritzteilen. Dabei werden experimentell Grundeinstellungen von Prozessgrößen bei der Nutzung eines Werkzeugs zur Herstellung eines Spritzteils mit dem Fertigungsverfahren Spritzgießen ermittelt [10].

3.1 Ablauf der Analyse

Der Arbeitsprozess „Füllstudie an einer Spritzgießmaschine“ zeichnet sich durch eine beobachtbare ausführende Handlung aus, welche jedoch viele analysierende

und planende Teilhandlungen benötigt. Der Arbeitsprozess wird zunächst mittels Arbeitsanalysen erhoben und aufbereitet [16]. Daraus ergeben sich Informationen und Anforderungen für die konkrete Lern- und Arbeitsaufgabe, die technische Konzeption des Lernmediums sowie zu Teilhandlungen des Lern- und Arbeitsprozesses. Durch vergleichende Analyse der empirischen Erhebungen mit einer literaturbasierten Füllstudienbeschreibung [10] können die notwendigen Arbeitsschritte und damit verbundenen Prozessgrößen identifiziert werden.

3.2 Parameteridentifikation

Aus der Arbeitsprozessanalyse geht hervor, dass zuerst allgemeine Maschinengrundeinstellungen vorgenommen werden. Anschließend erfolgt die experimentelle Bestimmung weiterer verfahrensbezogener Parameter. Zwischen den Parametern kommen Abhängigkeiten vor, die berücksichtigt werden müssen. Insgesamt müssen 20 in den einzelnen Arbeitshandlungen einzustellende Parameter bestimmt werden, welche nach einer erfolgreichen Füllstudie hinsichtlich der Zielstellungen weiter optimiert werden können [17]. Wie empirische Studien [18] bei einem Transfer auf die Facharbeit zeigen, sind dabei nicht triviale Interdependenzen der Parameter zu berücksichtigen, um Anforderungen an den Arbeitsprozess sowie dessen Ergebnis zu gewährleisten.

Ein Dilemma hinsichtlich der in Kapitel 2 beschriebenen Zieldimensionen Qualität und Ökologie kann bei der Wahl verschiedener Parameter beobachtet werden. Dazu wird zunächst keine weiteren Zieldimensionen betrachtet, um die Komplexität der Beispieldarstellung gering zu halten und im folgenden Unterkapitel ein Beispiel erläutert.

3.3 Ableitung eines Zielkonflikts

Als ausgewählter Parameter dient die Kühlzeit, welche die Zeitspanne von der Abkühlung bis zur Erreichung von ausreichend Formstabilität des Spritzteils umfasst [17]. Ein Fehler bezogen auf die Qualität des Spritzteils kann durch äußere Belastungen wie Auswerferstifte oder Stöße beim Entformvorgang entstehen, wenn die erforderliche Kühlzeit unterschritten wird, was materialwissenschaftlich durch temperaturabhängige Größen wie die Zugfestigkeit [19] oder der Schubmodul erklärbar ist [20]. Fehlerfolgen durch dauerhafte Verformungen des Spritzteils bedeuten Produktionsausschuss, was negativen Einfluss auf alle Zielstellungen hat.

Aus ökologischer Perspektive ist die Kühlzeit eng mit dem Energiebedarf des Produktionsprozesses verbunden [21, 22]. Dies bedeutet, dass der Energiebedarf mit der Länge der Kühlzeit ansteigt, da u.a. statische Kräfte auf das Werkzeug aufgebracht werden müssen und sich die Gesamtzykluszeit der laufenden Maschinen verlängert. Dabei wird mit längerer Kühlzeit auch kumuliert mehr Leerlaufleistung benötigt, die nicht wertschöpfend umgesetzt werden kann, sodass hier die ökologische Bilanz verschlechtert wird.

Es gilt beim beschriebenen Dilemma, ein Optimum zwischen zu kurzer und zu langer Kühlzeit als Kompromiss aus Ökologie und Qualität im Arbeitsprozess sowie bezogen auf das Handlungsergebnis zu finden. Die bereits eingeführte Zeitdimension ist aus wirtschaftlicher Hinsicht von großer Bedeutung, da die ausbringbare Stückzahl mit einer längeren Kühlzeit sinkt. Dies ist eine wichtige wirtschaftliche Kenngröße, welche der Qualitätsdimension entgegensteht.

4. Didaktische Diskussion

Der dargelegte Parameter Kühlzeit im Arbeitsprozess Füllstudie an einer Spritzgießmaschine ist sehr gut geeignet, um im Lernprozess verschiedene Kompetenzen zu fördern. Der Parameter eignet sich besonders, um die folgende Beschreibung als Erkenntnisziel in den Lernprozess einzubinden.

Zur Ermittlung der erforderlichen Kühlzeit vor dem Hintergrund der genannten Leitideen, sollen die Auszubildenden von den Möglichkeiten der Simulationstechnik [10] Gebrauch machen. Liegen keine Lösungen aus Simulationen vor, besteht die Möglichkeit, Näherungsformeln basierend auf den geometrischen Voraussetzungen oder Diagramme zur Ermittlung der Kühlzeit genutzt zu nutzen [17]. Mathematische Problembeschreibungen verdeutlichen, dass das Zusammenwirken verschiedener Parameter, aber auch, dass deren jeweiliger Einfluss auf die Kühlzeit sehr genau beschrieben werden kann. Unter der Annahme mittlerer Masse-, Werkzeug- und Entformungstemperaturen kann die erforderliche Kühlzeit auch grafisch als Funktion der Wandstärke dargestellt werden [17].

Die Auszubildenden lernen durch den Zielkonflikt in einer entsprechenden Lern- und Arbeitsaufgabe insbesondere, dass die einzelnen Parameter unterschiedliche Anforderungen an den Fertigungsprozess stellen: Während sich zu kurze Kühlzeiten negativ auf die Qualität der Spritzteile auswirken können, sind zu lange Kühlzeiten aus Gründen der ökologischen Nachhaltigkeit zu vermeiden. Eine zu lang eingestellte Kühlzeit ist im Vergleich zu einer zu kurzen Kühlzeit weniger problematisch für den Anfahrvorgang [23], sollte jedoch mit Blick auf die didaktische Leitidee der ökologischen Nachhaltigkeit vermieden werden. Da die Kühlzeit unter Verwendung von Formeln lediglich näherungsweise ermittelt wird, ist diese im Rahmen der Füllstudie bei Bedarf anzupassen [10]. Die Fachkraft muss aufgrund dieser gegenläufigen Beziehung einen geeigneten Wert ermitteln, der beide Aspekte berücksichtigt.

Die univariate Betrachtung eines isolierten Parameters im komplexen Zusammenhang eines Fertigungsverfahrens mit nichtlinearen Parameterabhängigkeiten kann nur als Beispiel und Ausgangslage oder Reflexionsaufgabe für die zu durchlaufenden Lernprozesse in der betrieblichen Ausbildung dienen. Die dabei erlangten Erkenntnisse müssen im Verlauf der Ausbildung fortlau-

fend erweitert und hinterfragt werden, um ein die Realität abbildendes Verständnis vom Fertigungsprozess Spritzgießen zu erlangen.

5. Fazit

Ein offener und konstruktiver Umgang mit Fehlern und Fehlentscheidungen eignet sich vor allem dann, wenn man sich in Dilemma-Situationen befindet, in welchem mehrere Ziele verfolgt werden sollen, die im Konflikt zueinanderstehen. Gerade noch unerfahrene Auszubildende können im Rahmen einer Füllstudie im Fertigungsverfahren Spritzgießen durch das AR-System neue Erfahrungen machen. Das dargelegte Lernkonzept eignet sich, um Verfahrensgrößen verantwortlich zu wählen, ohne dabei einzelne Zieldimensionen unbewusst zu vernachlässigen. Dies ist verkürzt an der Beispielgröße Kühlzeit hinsichtlich ökonomischer und ökologischer Ziele dargelegt worden. Aufgrund der Adressierung von Zielkonflikten ist ein Transfer auf weitere technische Ausbildungsbereiche mit vielfältigen Zieldimensionen denkbar, die auch weniger komplex ausfallen, als die Parameterabhängigkeiten im Spritzgießen. Das vorgestellte Beispiel eignet sich auch sehr gut, um Zielkonflikte der Nachhaltigkeit auf verschiedenen gesellschaftlichen Ebenen zu adressieren und eine reflektierte Handlungskompetenz in der Ausbildungsphase zu fördern, etwa an weiteren Lernorten wie Berufskollegs.

Danksagung

Der Beitrag entstammt dem Projekt „FeDiNAR – Fehler didaktisch nutzbar machen mit AR“. Es wird mit Mitteln des Bundesministeriums für Bildung und Forschung (BMBF) innerhalb des Fachprogramms „Digitale Medien in der beruflichen Bildung“ gefördert und vom DLR Projektträger unter den FKZ 01PV18005A betreut.

Literaturverzeichnis

- [1] A. Atanasyan, D. Kobelt, M. Goppold, T. Cichon, and M. Schluse, "The FeDiNAR Project: Using Augmented Reality to Turn Mistakes into Learning Opportunities," in *Springer Series on Cultural Computing, Augmented Reality in Education: A New Technology for Teaching and Learning*, V. Geroimenko, Ed.: Springer, 2020, pp. 71–86.
- [2] Bundesinstitut für Berufsbildung, "Bekanntmachung des Verzeichnisses der anerkannten Ausbildungsberufe und des Verzeichnisses der zuständigen Stellen nach § 90 Absatz 3 Nummer 3 des Berufsbildungsgesetzes (BBiG) vom 23. März 2005 (BGBl. I S. 931)," Bundesinstitut für Berufsbildung, 2020.
- [3] Bundesministerium für Wirtschaft und Technologie, "Verordnung über die Berufsausbildung zum Verfahrensmechaniker für Kunststoff- und Kautschuktechnik und zur Verfahrensmechanikerin für Kunststoff- und Kautschuktechnik: KStoffVerfMAusbV 2012," *Bundesgesetzblatt I*, vol. 64, no. 23, pp. 1168–1200, 2012.

- [4] F. Rauner, "Zur Konstitution einer neuen Bildungs-idee: "Befähigung zur Technikgestaltung"," in *Forschungsreihe des Forschungsschwerpunkts "Arbeit und Bildung"*, vol. 6, *Ende der Aufklärung?: Zur Aktualität einer Theorie der Bildung*, R. Drechsel, Ed., Bremen: Univ, 1987, pp. 266–297.
- [5] R. Bader, "Entwicklung beruflicher Handlungskompetenz durch Verstehen: Ein Beitrag zum systemtheoretischen Ansatz in der Technikdidaktik," *Die berufsbildende Schule*, vol. 43, 7-8, pp. 441–458, 1991.
- [6] R. Atkinson, "Project management: cost, time and quality, two best guesses and a phenomenon, its time to accept other success criteria," *International Journal of Project Management*, vol. 17, no. 6, pp. 337–342, 1999.
- [7] Deutscher Bundestag, "Abschlußbericht der Enquete-Kommission „Schutz des Menschen und der Umwelt - Ziele und Rahmenbedingungen einer nachhaltig zukunftsverträglichen Entwicklung": Konzept Nachhaltigkeit Vom Leitbild zur Umsetzung," Drucksache 13/11200, Bonn, 1998.
- [8] *Qualitätsmanagementsysteme - Grundlagen und Begriffe*, DIN EN ISO 9000, DIN e.V., Berlin, 2015.
- [9] F. Johannaber and W. Michaeli, *Handbuch Spritzgießen*. München: Carl Hanser Verlag, 2004.
- [10] A. Schötz, *Abmusterung von Spritzgießwerkzeugen: Strukturierte und analytische Vorgehensweise*, 3rd ed. München: Hanser, 2019.
- [11] Kultusministerkonferenz, "Rahmenvereinbarung über die Berufsschule: Beschluss-Nr. 323," Berlin, 2019.
- [12] Kultusministerkonferenz, "Verfahrensmechaniker für Kunststoff- und Kautschuktechnik/ Verfahrensmechanikerin für Kunststoff- und Kautschuktechnik," Kultusministerkonferenz, Mar. 2012.
- [13] P. T. Tramm and N. Naeve, "Auf dem Weg zum selbstorganisierten Lernen – Die systematische Förderung der Selbstorganisationsfähigkeit über die curriculare Gestaltung komplexer Lehr-Lern-Arrangements," *bwp@*, no. 13, 2007.
- [14] A. A. Cattaneo and E. Boldrini, "Learning from errors in dual vocational education: video-enhanced instructional strategies," *Journal of Workplace Learning*, vol. 29, no. 5, pp. 357–373, 2017, doi: 10.1108/JWL-01-2017-0006.
- [15] J. Bauer and R. H. Mulder, "Modelling learning from errors in daily work," *Learn Health Soc Care*, vol. 6, no. 3, pp. 121–133, 2007, doi: 10.1111/j.1473-6861.2007.00150.x.
- [16] M. Goppold, L. Nobis, M. Frenz, and V. Nitsch, "Erhebung und Modellierung von Arbeitsprozessen in der metall- und elektrotechnischen Ausbildung als Basis einer didaktischen Konzeption „Lernen mit Fehlern in einer AR angereicherten Lernumgebung"," in *Digitaler Wandel, digitale Arbeit, digitaler Mensch?*, Berlin, 2020, B.17.1.
- [17] Demag Plastics Group, *Spritzgießen - kurz und bündig*, 5th ed. Schwaig: Demag Ergotech GmbH, 2004.

- [18] R. Schiffers, K. Morik, A. S. Struchtrup, P.-J. Honysz, and J. Wortberg, "Anomaly detection in injection molding process data based on unsupervised learning," *Zeitschrift Kunststofftechnik/Journal of Plastics Technology*, vol. 2018, no. 1, pp. 301–347, 2018, doi: 10.3139/O999.02052018.
- [19] C. Bonten, *Kunststofftechnik: Einführung und Grundlagen*, 3rd ed., 2020.
- [20] E. Baur, G. Harsch, and M. Moneke, *Werkstoff-Führer Kunststoffe: Eigenschaften - Prüfungen - Kennwerte*, 11th ed., 2019.
- [21] I. Meekers, P. Refalo, and A. Rochman, "Analysis of Process Parameters affecting Energy Consumption in Plastic Injection Moulding," *Procedia CIRP*, vol. 69, pp. 342–347, 2018, doi: 10.1016/j.procir.2017.11.042.
- [22] J. B. Tranter, P. Refalo, and A. Rochman, "Towards sustainable injection molding of ABS plastic products," *Journal of Manufacturing Processes*, vol. 29, pp. 399–406, 2017, doi: 10.1016/j.jmapro.2017.08.015.
- [23] C. Jaroschek, *Spritzgießen für Praktiker*, 4th ed. München: Hanser Carl, 2019.

Standardisiertes Übungsszenario für HMD-basierte virtuelle Lernumgebungen

Leif Goldhahn, Robert Eckardt,
Katharina Müller-Eppendorfer

Hochschule Mittweida, Fakultät Ingenieurwissenschaften,
InnArbeit - Zentrum für innovative Arbeitsplanung und Arbeitswissenschaft,
Technikumplatz 17, 09648 Mittweida

Mittels einer entwickelten virtuellen Lernumgebung inklusive integrierten Trainingsszenarien kann methodisches Wissen zur Steigerung der Bediensicherheit von Fertigungsanlagen anschaulich vermittelt werden. Dabei stehen vor allem die typischen Bedienhandlungen zum Betreiben oder Rüsten einer Werkzeugmaschine, wie das Betätigen von Tasten und Drehschaltern an Bedienpulten oder Arbeitsstationen im Vordergrund. Als Hardware zur Darstellung der virtuellen Szene kommt für ein voll-immersives VR-Erlebnis ein Head Mounted Display (HMD) mit zwei handgehaltenen Controllern (Eingabegeräte mit Tracking) zum Einsatz. In der erstellten Lernumgebung „Trainingsfabrik 4.0“, welche auch als reale Trainingsfabrik 4.0 an der Fakultät Ingenieurwissenschaften der Hochschule Mittweida existiert, konnten IngenieurInnen und Studierende die Lernumgebung mit den integrierten Bedienhandlungen testen. Deutlich wurde dabei, dass Vorkenntnisse vor allem im Hardware-Umgang begünstigende Voraussetzung sind, um das eigentliche Trainingsziel „Erlernen von Bedienhandlungen“ erreichen zu können. Viele Testpersonen arbeiteten mit den Controllern in der virtuellen Szene zunächst nicht zielführend oder gar intuitiv. Das Navigieren in der Umgebung aber auch das Ausführen der Bedienhandlungen (Steuerung) wurden dadurch stark behindert, was den gewünschten Lernerfolg sowie die Akzeptanz der Technik und der Herangehensweise klar minderte. Daraus leitet sich die Notwendigkeit ab, den Lernenden zunächst den Umgang mit der HMD-Technologie sowie die Interaktion mit der VR-Softwareumgebung in einem virtuellen Übungsszenario gezielt zu veranschaulichen. Dafür wurde ein standardisiertes 2-Phasenkonzept erstellt.

1. Motivation

Digitale Technologien wie Virtual Reality (VR) eröffnen neue Potentiale und Ansätze für selbstgesteuerte und selbstorganisierte Lernprozesse zur Wissensvermittlung komplexer Prozesse. Bekannt und anerkannt ist, dass für das Erlernen der Flugzeugbedienung, welche in der Realität nur schwer zu simulieren ist (z. B. Landung bei Wind) bereits VR-basierte Trainingssimulatoren existieren. Mit Hilfe dieser Simulatoren und den darin durchzuführenden Szenarien kann das Verhalten der Piloten analysiert und dadurch evtl. Fehlverhalten aufgedeckt und prospektiv abgestellt werden. [1] Fehlerquoten aber auch der Schulungsaufwand im realen Betrieb werden damit deutlich reduziert. [2]

Für den Umgang mit Werkzeugmaschinen, wie CNC-gesteuerten Fräsmaschinen und charakteristischen Bedienhandlungen daran, gibt es solche VR-basierte Simulatoren gegenwärtig kaum. [3] Das Interagieren und Arbeiten mit VR-basierten Simulatoren wird durch die Entwicklung und Verwendung anwendungsbezogener virtueller Lernszenarien möglich. Virtuelle Lernumgebungen mit zielorientierten Lernszenarien, zum Beispiel für das Inbetriebnehmen oder das Rüsten einer Werkzeugmaschine, orientieren sich am sensomotorisch- und visuell basierten Wissensaufbau und unterstützen damit die Vermittlung prozeduraler und handlungsbezogener Fertigkeiten in einer virtuellen Trainingswelt. Daraus ergibt sich der Vorteil, das Schadenspotential für Menschen,

Maschinen oder Werkzeuge bei Fehlbedienungen im Prozess enorm zu reduzieren bzw. sogar auszuschließen. Auch unproduktive Stillstandzeiten der Werkzeugmaschinen, welche bspw. durch unsicheres Agieren, Fehlbedienungen oder Trainingsprozesse erzeugt werden, können dadurch reduziert werden. Die am Markt verfügbaren und kommerziell nutzbaren VR-Systeme, wie Head Mounted Displays (HMD) sind zwar kostengünstig in der Anschaffung, haben aber den grundlegenden Nachteil, dass die Navigation sowie die Steuerung vor allem auf die Adaption für VR-basierte Simulatoren im industriellen Bereich wie hier von Werkzeugmaschinen und deren komplexen Bedienhandlungen nicht intuitiv ist. Daher muss der Umgang mit dieser Technologie zunächst geübt werden, um sie adäquat für das voll-immersiv unterstützende Lernen nutzen zu können. Dafür wurde ein virtuelles Übungsszenario entwickelt, welches im Beitrag vorgestellt wird.

2. Grundlagen

2.1. Virtual Reality

Die Literatur bezeichnet VR als die „visuelle Präsentation und Manipulation von 3D-Daten in Echtzeit“. [4]

Nach [5] wird VR durch die drei nachfolgenden Eigenschaften charakterisiert:

- **Immersion**, die realitätsnahe Darbietung und Wahrnehmung

- **Interaktion**, die Beeinflussung und Veränderung in Echtzeit sowie
- **Imagination**, das Erleben als quasi reale Welt.

Mit Hilfe von VR werden die visuellen, akustischen und/oder haptischen Sinne des Menschen angesprochen. Publikationen wie [6] untersetzen, dass sich die auf dem Markt vorhandenen VR-Systeme (z. B. HMD oder Powerwall) durch den Grad der Trennung des Anwendenden von der realen Welt sowie der Einbindung in eine virtuelle Umgebung unterscheiden. Ein HMD wie die HTC vive pro, mit einem hohen Grad der Immersion [6], ermöglicht das freie Interagieren mit beiden Händen mittels Controllern an einem virtuellen Arbeitsplatz, wodurch Arbeitsaufgaben inkl. Bedienhandlungen gezielt sowie wiederholt trainiert werden können. Die Visualisierung und perspektivischen Veränderungen unterstützen einen realitätsnahen Tiefeneindruck.

2.2. Bedienhandlungen, Lernszenarien und -umgebung

Eine VR-basierte Bedienhandlung bezeichnet einen ganzheitlichen virtuellen Bedienvorgang zur Veranlassung, Ausführung und Abschluss einer Funktion an einer virtuellen technischen Einrichtung, welche sich unverzüglich, auch mehrstufig, auf die virtuelle technische Einrichtung auswirken kann. [7]

Bei einem virtuellen Lernszenario erfolgt, im Sinne von Problemlösungs- oder Handlungsstrategien, die Bereitstellung von mehreren kontextbezogenen Handlungssituationen. [8] Die Literatur spricht bei dieser Nutzung

der VR in Lernszenarien (auch „lernbezogene Interaktivität“) von einem hohen Praxisbezug und „guten Resultaten im Hinblick auf die Lernwirksamkeit“. [4] Virtuelle Lernszenarien unterstützen daher die Vermittlung prozeduraler und handlungsbezogener Fertigkeiten.

Sie sind in komplexere Lernumgebungen (Trainingswelten) einzubinden. Diese virtuellen Umgebungen stellen „begehbare“ und realitätsnahe Modelle bereit, welche die Lernmotivation und Akzeptanz fördern und auch der späteren Orientierung im realen Raum dienen. [9]

Die virtuellen Lernszenarien nutzen die realen Betriebszustände der Werkzeugmaschinen (z. B. Normalbetrieb, An- und Abfahren, Inbetrieb- und Außerbetriebnahme) und die für die auftragsspezifische Bearbeitung erforderlichen Tätigkeiten wie Rüsten, Einlegen, operative Tätigkeiten usw.

Arbeitsaufgaben wie die Inbetriebnahme einer Maschine können in einem virtuellen Lernszenario abgebildet und nachfolgend innerhalb einer virtuellen Lernumgebung (vgl. Abbildung 1) trainiert werden. [4] In der Abbildung 1 muss die Testperson die Bedienhandlung „Druckluftversorgung herstellen“ umsetzen. Dazu bekommt sie die Informationen, wie das Ventil aussieht, wie sie zur Handlungsstelle gelangt (Fußspuren) und wie sie die Bedienhandlung umsetzen muss. Dies kann die Anzahl und Schwere der Fehler senken, die während der realen Herstellung der Betriebsbereitschaft oder des Betriebs der Anlage auftreten und zu einer Gefährdung des Menschen sowie der Beschädigung der realen Anlagenkomponenten führen könnten.

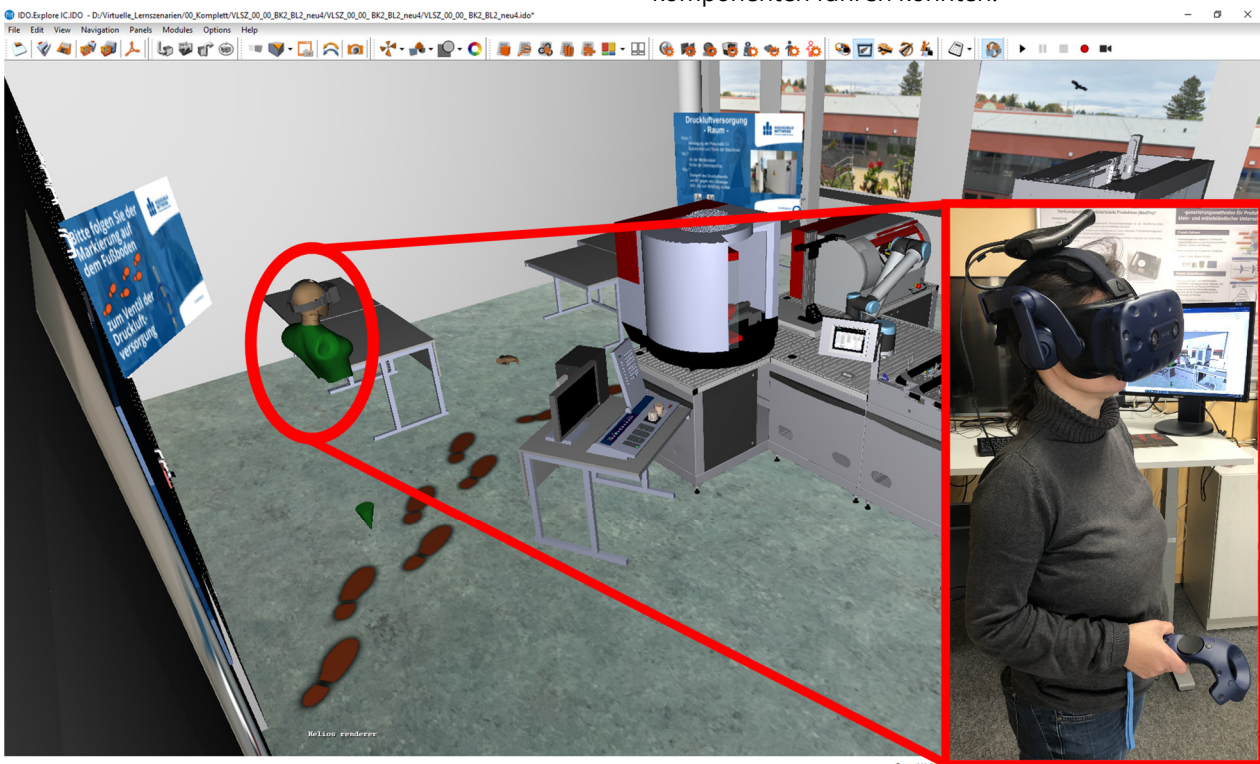


Abbildung 1: Virtuelle Lernumgebung mit Orientierungselementen (Fußspuren) sowie Anwenderin mit HMD-Brille und Controller

3. Entwicklung eines standardisierten Übungsszenarios in 2 Phasen

3.1. Allgemeines & Aufbau virtuelles Übungsszenario

Voraussetzung für die Akzeptanz eines virtuellen Lernszenarios ist, dass ein hoher Wiedererkennungswert der virtuellen Umgebung im Vergleich zur realen Umgebung vorhanden sein muss. Darüber hinaus ist es wesentlich, dass nach dem Betätigen einer Taste auf der Steuerung der virtuellen Werkzeugmaschine (Bedienhandlung) eine möglichst realitätsnahe Reaktion im virtuellen System erfolgt. Daher ist es wesentlich, das Nutzungserlebnis bewerten zu können. Erste Versuche mit den erarbeiteten virtuellen Lernszenarios zeigten, dass zunächst nicht das Umsetzen der geforderten Bedienhandlungen die Testpersonen vor Probleme stellte, sondern der Umgang mit der voll-immersiven VR-Technologie selbst. 8 von 10 Testpersonen kamen mit der Steuerung und Navigation innerhalb der virtuellen Lernumgebung nicht zurecht, wodurch das eigentliche Training des Lernszenarios in diesen Fällen nicht durchgeführt werden konnte. Die Abbildung 2 zeigt eine Übersicht, welche Fehler, in welcher Häufigkeit bei den gescheiterten Versuchen aufgetreten sind. Die Abbildung macht deutlich, dass vor allem die Navigation beim Bewegen durch die virtuelle Lernumgebung erhebliche Probleme verursachte.

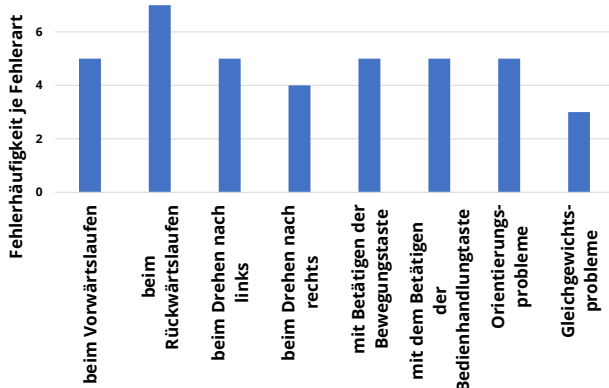


Abbildung 2: Überblick Fehlerarten und -häufigkeiten während der Durchführung des virtuellen Lernszenarios (8 Testpersonen)

Auf der Grundlage dieser Versuchsreihen wurde ein zusätzliches und voranzuschaltendes **virtuelles Übungsszenario** erarbeitet. Mit diesem sollen die Testpersonen sowohl die Navigation als auch die Steuerung im Sinne spezifischer Bedienhandlungen im Kontext eines zielführenden Umgangs mit der VR-Hardware sowie den charakteristischen Aufgaben an Werkzeugmaschinen zukünftig standardisiert üben können. Damit werden zunächst technische Versuche zur Machbarkeit und Akzeptanz vorgenommen. Die Basis für das virtuelle Übungsszenario bildet die virtuelle Lernumgebung (vgl. Abbildung 3), wodurch ein hoher Wiedererkennungswert erreicht werden konnte. Mit der Hilfe der Informationstafeln im Übungsszenario wird der Testperson schrittweise beschrieben, wie sie durch die Lernumgebung navigieren kann sowie Bedienhandlungen ausführen (steuern) muss. Verbal unterstützt wird dies von außen durch den Laboringenieur.

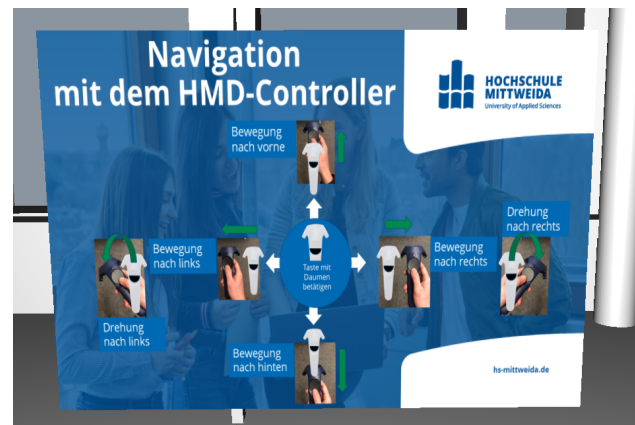


Abbildung 3: Informationstafel zum Erlernen der Navigation im virtuellen Übungsszenario

3.2 Beschreibung 2-Phasenkonzept

Als Konsequenz der beschriebenen Probleme wurde ein 2-Phasenkonzept entwickelt, welches in Abbildung 4 auszugsweise dargestellt ist. Die 2 Phasen bauen dabei aufeinander auf und sind auch unabhängig voneinander anwendbar. Dies ist vor allem bei unterschiedlichen Wissensständen im Hinblick auf die Navigation und Steuerung durch die künftigen Testpersonen essentiell.

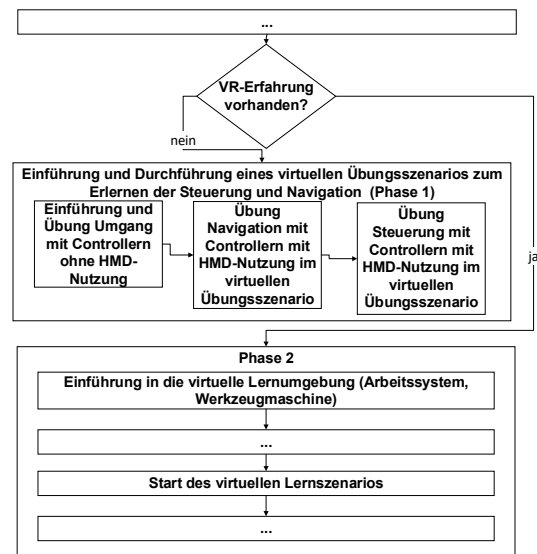


Abbildung 4: Ausschnitt methodischer Ansatz zur Entwicklung eines virtuellen Übungs- und Lernszenarios mit integriertem 2-Phasenkonzept

Die **Phase 1 – Umgang mit der Hardware HMD** zielt dabei darauf ab, zunächst ein Gefühl für die Controller und deren Tasten zu bekommen (vgl. Abbildung 1). Vor allem das gezielte Auswählen der unteren Taste mit dem Zeigefinger sowie der oberen Taste für den Daumen müssen für eine Interaktion durch die Testperson nutzbar sein (Hand-Auge-Koordination). Dazu erhalten die Testpersonen zunächst eine Einführung in die Controller, ohne dass ihnen dabei das HMD die Sicht auf die Controller nimmt. Im Anschluss wird das Controller-Handling, während die HMD den Testpersonen die Sicht auf die Controller nimmt, geübt, ohne dass eine Hand-Auge-Koordination möglich ist. Sobald der Umgang mit der Hardware sicher ist, erfolgt die Bewegung

und Interaktion im virtuellen Übungsszenario (vgl. Abbildung 5). Im Übungsszenario wird das Zusammenspiel zwischen Hardware (Controller und Brille) sowie Software (VR-Umgebung) geübt. Dazu wird die Testperson mit Hilfe der Informationstafeln durch die (vgl. Abbildung 5) virtuelle Welt geleitet und vollzieht kleine Aufgaben wie das Drehen eines Schalters (Steuerung). Nach Abschluss des Trainings der Navigation und der Steuerung (Testpersonen können sich von Bedienstelle zu Bedienstelle selbstständig bewegen sowie Bedienhandlungen durchführen) in der Phase 1 folgt die Phase 2. Die **Phase 2** umfasst das **selbstständige virtuelle Lernen** und beinhaltet die Bedienhandlungen, welche zum Erlangen der Bediensicherheit von Werkzeugmaschinen dienen (vgl. Abschnitt 2.2).

3.3. Beschreibung Versuchsreihe virtuelles Übungsszenario

Drei Testpersonen konnten bisher entsprechend den Ausführungen der Phase 1 mit den Navigations- und Steuerungsgegebenheit des HMD vertraut gemacht werden. Dabei wurde je Testperson die dokumentiert, zu welcher die Testpersonen dem eigenen Empfinden nach, sicher im Umgang mit der Hardware waren Zeit (im Durchschnitt 9min und 13s). Um die Steuerung, das heißt das Ausführen von Bedienhandlungen zu üben, wurde den Testpersonen das in der Abbildung 5 dargestellte Druckluftventil vorgestellt. Hier musste das Berühren, das Festhalten sowie das Drehen des Ventils mit unterschiedlichen Ventilgrößen geübt werden. Je kleiner das Ventil, desto schwieriger die Bedienhandlung des Drehens. Ziel war es, das Bedienen des Ventils in Originalgröße zu üben.



Abbildung 5: Druckluftventil zum Erlernen der Steuerung im virtuellen Übungsszenario

4. Auswertung und Zusammenfassung

Drei von drei Testpersonen erachteten die Durchführung des virtuellen Übungsszenarios als hilfreich für die anschließende Durchführung des virtuellen Lernszenarios. Weiterhin kann gesagt werden, dass das Übungsszenario eine wesentliche Unterstützung darstellt, den anvisierten Erfolg für das virtuelle Lernen realer Bedienhandlungen an maschinellen Anlagen sicherzustellen. Der Umgang mit dem HMD im Zusammenspiel mit dessen Controllern erfolgte nach Durchführung des Übungsszenarios sicher und selbstständig durch die jeweilige Testperson. Die ersten Daten der Versuchsreihe zeigen, dass nach der Durchführung des virtuellen Übungsszenarios die Testpersonen

das anschließende virtuelle Lernszenario erfolgreich absolvieren konnten. Darüber hinaus zeigen die Versuche, dass der Zeitbedarf zur Ausführung des virtuellen Lernszenarios (vgl. Abschnitt 2.2.) im Mittel bei 4min und 23s lag. Die beiden Testpersonen, die ohne die Durchführung des Übungsszenarios das virtuelle Lernszenario durchführten, benötigten dafür im Mittel 4min und 50s.

Danksagung

Die Autoren danken dem Bundesministerium für Bildung und Forschung für die finanzielle Unterstützung sowie dem Projektträger Jülich für die Betreuung des Forschungsprojektes „Smart University Grid Saxony5 - Wissensströme intelligent vernetzen (Saxony5)“.



Literaturverzeichnis

- [1] Dörner, R., Geiger, C., Oppermann L., Paelke V., Interaktionen in Virtuellen Welten in Virtual und Augmented Reality (VR/AR), Springer-Verlag Berlin, Heidelberg, (2013), S. 157ff.
- [2] Wiendahl, HP., Fiebig, T. Virtual factory design – a new tool for a co-operative planning approach Nr. 7: (2003), S. 535-540
- [3] Klimant, P., Klimant, F., Putz, M., Allmacher, C., Trends der virtuellen und erweiterten Realität mit dem Fokus auf virtueller Inbetriebnahme. VAR² 2019 – Realität erweitern. Chemnitz, (2019), S. 7-26
- [4] Hirt C, Spahn M, Kompis Y, Jetter D, Kunz A Alles rund um die Maschine – Begehbare virtuelle Schulung an Werkzeugmaschinen. In: VAR² 2019 – Realität erweitern, Hrsg. Putz M, Klimant P, Klimant F, Chemnitz, ISBN: 978-3-00-064420-7, (2019), S. 43 – 58
- [5] Goldhahn, L. Montageplanung und -ausführung – virtuell und real. 16 th International Scientific Conference Mittweida IWKM. Scientific Reports Nr. 2, (2003), S. 46 – 50
- [6] Engel, B., Immersion oder Versinken in der virtuellen -Realität – auch ein Thema für die Arbeitsmedizin? In: Zeitschrift für medizinische Prävention, vol. 54, n° 09, Enlign: <https://www.asu-arbeitsmedizin.com/node/30413/print>. (2019), S. 604–607
- [7] Goldhahn, L, Eckardt, R., Pietschmann, C., Roch, S. Lernszenarien und Virtual Reality-basierte Bedienhandlungen zur Mitarbeiterqualifizierung. Dortmund: GfA Press, ISBN 978-3-936804-27-0, B1.1. (2020), S. 1-6
- [8] Hirschle, T. Lernszenario, Zugang Jan 07. <https://bildungsserver.berlin-brandenburg.de/fileadmin/bbb/unterricht/faecher/gesellschaftswissenschaften/ethik/ Lernszenario.pdf> (2020)
- [9] Schwan, S., Buder, J., Virtuelle Realität und E-Learning, S. 7-9, Zugang Jan 7, 2020, <https://www.e-teaching.org/didaktik/gestaltung/vr/vr.pdf> (2006)

Auswahl und Einsatz von Smart Devices in der Arbeitsanalytik

Torsten Merkel

Westsächsische Hochschule Zwickau, Kornmarkt 1, 08056 Zwickau

Smart Devices versprechen die ständige Erhebung und Kopplung von Daten mit vielfältigen Sensoren. Mit Geräten der Gruppe für „Vital- und Ambiente-Monitoring“ lassen sich formal große Teile der Arbeitsanalytik abdecken. Auf Grund ihrer Mobilität, der Möglichkeit Daten aus verschiedenen Quellen zusammenzuführen und für Auswertungen über eigene Rechenkapazität und individualisierte Anzeigemöglichkeiten zu verfügen, erscheinen „Smart Devices“ als ideale Ergänzung für die Analyseaufgaben für arbeitsmedizinische oder arbeitshygienische Aufgabenstellungen. Der Beitrag stellt die Ergebnisse einer entsprechenden Studie vor. Es wurden Geräteklassifikationen erarbeitet, welche die Anwendungen für verschiedene Einsatzszenarien unterstützen. Als klassische Kriterien werden die Einordnung in Abhängigkeit des Mobilitätsgrades oder die verbaute Sensortechnologie und deren Messbereich betrachtet.

Smart devices promise the constant collection and coupling of data with a wide range of sensors. Devices of the group for "vitality and ambience monitoring" offer the most relevant parts of the work analysis. Their mobility, the possibility of merging data from different sources and their own computing capacity and individual display allows options for evaluations, de-coding and so on. Devices ideal for the analysis tasks for occupational health or occupational hygiene tasks.

The paper presents the results of a corresponding study. Device classifications have been developed that support applications for different application scenarios. As classic criteria, the classification is considered depending on the degree of mobility or the installed sensor technology and its measuring range.

1. Zielstellung und Inhalt der Studie

Arbeitsmedizin und Arbeitshygiene setzen sich mit der Analyse menschlicher Arbeit auseinander, um daraus Maßnahmen, Regeln und Empfehlungen zur Gestaltung der Arbeit abzuleiten, welche Gesundheit und Leistungsfähigkeit der Betroffenen erhalten.

Eine Analyse von Arbeitssystemen ist komplex und wird durch zahlreiche Einflussfaktoren bestimmt. Hinzukommen individuelle Unterschiede bezüglich des Trainings, der Erfahrung, der Arbeitsmethode, aber auch Alter, Geschlecht und Gesundheitszustand. Die folgende Grafik versucht die Vielfalt der zu berücksichtigenden Einflussfaktoren zu umreißen.



Abbildung 1: Komplexität der Beurteilung von Arbeitssystemen

Sowohl Tätigkeitsabläufe als auch die Umgebungsparameter unterliegen einer gewissen Dynamik, deshalb erfolgt die Erfassung und Beschreibung in der Regel entweder durch zeit- und arbeitsaufwendige Arbeitsstudien oder durch punktuelle Untersuchungen, welche durch subjektiv geprägte Berichte der Stelleninhaber unteretzt werden.

Der Einsatz von smarten Geräten verspricht für diese Aufgabenstellung eine deutlich aufwandsreduzierte und gleichzeitig qualitativ hochwertigere Analyse als diese durch Beobachtung oder punktuelle Kontakte, z. B. als Teil von Begehungen oder im Rahmen einer Sprechstunde möglich ist.

Smarte Geräte bieten im Allgemeinen eine breite Palette von Mehrwerten gegenüber herkömmlicher Technik mit vergleichbaren Basisfunktionen.

Ein maßgeblicher Funktionsunterschied von smarten Geräten besteht darin, dass sich die Systeme weitgehend vernetzen lassen und damit Funktionen, wie Gerät A löst bei Gerät B eine Aktion aus. Die Daten verschiedener Sensoren oder Ergebnisse von Algorithmen können ausgelesen und übertragen werden, ohne unmittelbar vor Ort zu sein. Ein Transfer von Daten und Informationen kann von einem smarten Messgerät mit und ohne Anzeige automatisch in eine Datenbank oder auf ein großes Display mit anderen Darstellungsmöglichkeiten übertragen werden.

Für eine begleitende Arbeitsanalytik bedeutet dies beispielsweise den Einsatz miniaturisierter Messgeräte,

welche ohne größere Beanspruchung von Personen getragen werden, kann. Die dabei erfassten Daten sind durch Dritte, z. B. betriebliche Arbeits-mediziner orts- und zeitunabhängig auswertbar. Ebenso besteht die Möglichkeit, dass gleichzeitig mehrere Personen verfolgt und deren Arbeitsparameter analysiert werden.

In den letzten Jahren wurden zahlreiche smarte Geräte entwickelt, welche ein ganzheitliches Konzept einer tätigkeitsbegleitenden Analytik in den Bereichen der Arbeitsmedizin und Arbeitshygiene ermöglichen. Algorithmen ermöglichen es, beim Erreichen definierter Grenzbereiche andere Aktivitäten auszulösen. Beispielsweise kann bei einer einseitigen körperlichen Belastung ein Tätigkeitswechsel angeregt werden. Ebenso lässt sich beim Erreichen von ungünstiger Raumluftparameter ein Lüftungsprozess initiiert werden. Die Geräte sind klein und leicht, so dass auch mehrere dieser Geräte von der zu untersuchenden Person mitgeführt werden können, ohne dass es zu einer signifikanten Erhöhung der Beanspruchung kommt. Ergänzend lassen sich in den relevanten Arbeitsbereichen Sensoren installieren, die über Netzwerke kommunizieren und mit der Zentraleinheit für Datenaufbereitung und Steuerungsprozesse verbunden sind.

Will man die vorhandenen Möglichkeiten in ihrer Vielfalt nutzen, so entstehen schnell komplexe Gebilde. Erschwerend kommt hinzu, dass es kaum einheitliche Standards der Hersteller gibt. Damit besteht unter anderem ein großes Problem in der zeitlichen Synchronität der verwendeten Geräte und der Übergabe von Daten. Die angebotenen Geräte weisen zudem große Unterschiede in der Messgenauigkeit auf. Ein Vital-Tracker, welcher vorzugsweise die Motivation seines Trägers zu einer Verbesserung der Aktivitätsrate unterstützen soll, ist in der Regel für eine medizinische Interpretation der ermittelten Vitaldaten nicht oder nur bedingt geeignet. In die andere Richtung betrachtete, sind die Geräte für den Anwendungszweck zu teuer. Infolgedessen erweisen sich zahlreiche Analyseansätze praktisch als nicht sinnvoll umsetzbar.

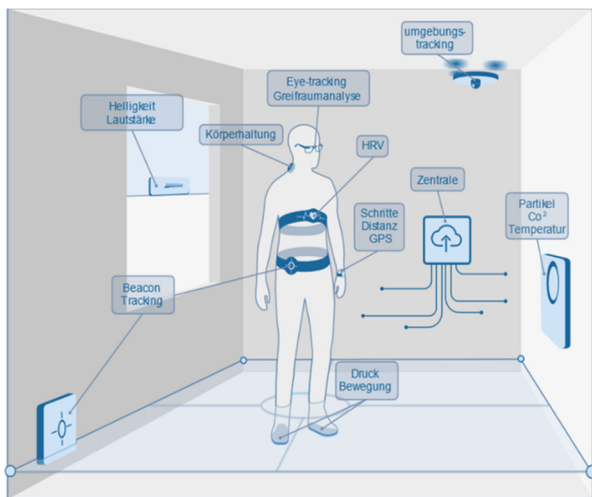


Abbildung 1: Theoretische Möglichkeit einer umfassenden Analyse und Steuerung mittels smarterer Geräte

Für die Auswahl und den Einsatz von Smart Devices bedarf es deshalb einer umfassenden Vorplanung, welche die Anforderungen einer geplanten Arbeitsstudie definiert und die dazu notwendigen Geräte in einer kompatiblen Gerätegruppe zusammenstellt. Ohne einen nachprüfbaren Test sollte man sich auf die Qualität der bisherigen Studienergebnisse nicht verlassen. Wie bei allen Messgeräten kann eine Eichung/ Kalibrierung und die Abstimmung der Geräte untereinander notwendig sein.

Empfehlenswert ist es, die Zahl eingesetzter smarterer Geräte für eine Studie so gering wie möglich zu halten. Vor Beginn der Studie aufgestellte Thesen zu den Korrelationen mehrere Parameter können für eine solche Vereinfachung hilfreich sein. Sind eine

Vielzahl sich gegenseitig beeinflussender Parameter zu erwarten, so ist ein iteratives Vorgehen sinnvoll.

Unbedingt sollten die Sicherheit der Daten und die rechtliche Konformität der geplanten Studie im Vorfeld geklärt werden.

Sind diese Vorarbeiten erfolgreich abgeschlossen, so bieten smarte Geräte zahlreiche Möglichkeiten einer umfassenden, arbeitsbegleitenden Analyse von Belastung und Beanspruchung einschließlich der direkten Einflussnahme auf den Arbeitsprozess. Der größte Vorteil besteht in der Vereinfachung einer lückenlosen Begleitung, da der Arbeitsanalytiker nicht permanent vor Ort sein muss. Für die Untersuchung psychisch bedingter Beanspruchungen ist dies von besonderem Wert, da durch die durchlaufende Begleitung der die Studie durchführenden Personen eigene Effekte entstehen. Insbesondere Langzeitstudien von Tätigkeiten lassen sich durch den Einsatz smarterer Messtechnik sinnvoll und effizient umsetzen. So kann eine einzelne Person Studien der Arbeit von mehreren Personen über Monate begleiten und muss nur punktuell die Funktionsfähigkeit der Geräte und die Konsistenz des Analyseansatzes z. B. durch eine Multimomentstudie prüfen.

2. Erstellen einer Auswahlmethodik

Beschäftigt man sich mit der Beschaffung von smarten Geräten, dann lassen sich die folgenden Auswahlkriterien anwenden:

- Funktionsumfang und Zugriff auf relevante Parameter (wie Schritte, Herzfrequenz, Herzfrequenzvariabilität, Energieumsatz, Umgebungsparameter),
- Kabelgebunden oder Akku-Versorgung (Laufzeit und selbständige Anwendung der benötigten Messaufgaben über einen oder mehrere Arbeitstage)
- Tragekomfort
- Zugänglichkeit der Daten/Sicherung des Datenschutzes
- Preis-Leistungs-Verhältnis

Danksagung

Ein besonderer Dank gilt Herrn Markus Weber für die Erstellung der Grafiken.

Literaturverzeichnis

- [1] Weber B.; Ellegast R.; Schellewald V., Weber A.; Röhrig M.; Friemert D.; Hartmann, U.: „Messung der physischen Aktivität mit Wearables“ 6. DGUV Fachgespräch Ergonomie – Zusammenfassung der Vorträge vom 2./3. November 2016

Konzept zur Simulation des Entscheidungsverhaltens von Arbeitspersonen bei der Produktion kundenindividueller Produkte

Sven Tackenberg¹, Jan-Phillip Herrmann¹, Marvin Goppold²

¹Labor für Industrial Engineering der Technischen Hochschule Ostwestfalen-Lippe, Campusallee 12, 32657 Lemgo

²Institut für Arbeitswissenschaft der RWTH Aachen University, Eilfschornsteinstr. 18, 52062 Aachen

Der Einsatz von neuen Informationstechnologien führt zu einer signifikanten Veränderung der individuellen Arbeitsorganisation der Arbeitspersonen auf dem Shopfloor eines Unternehmens. Durch den Einsatz von mobilen Assistenzsystemen werden neue Möglichkeiten geschaffen, Aufgaben an die Mitarbeiter situationsbezogen zu delegieren und Bearbeitungsstände zu überwachen. Trotz einer verstärkt algorithmenbasierten Planung der Arbeit, bestehen im Rahmen einer Werkstattorganisation weiterhin vielfältige Freiheitsgrade hinsichtlich der Ausgestaltung der individuellen Arbeitsorganisation. Eine Beobachtung von realen Arbeitsabläufen führte daher zu der Entwicklung eines Bayesschen Netz zur Beschreibung des Verhaltens einer Arbeitsperson in Abhängigkeit der bereitgestellten Informationen. Das entwickelte Netz berechnet hierbei die Wahrscheinlichkeit, dass eine Arbeitsperson aus einem Aufgabenpool eine Aufgabe auswählt und eine spezifische Handlung ausführt. Für das Bayessche Netz wird ein aktororientiertes Simulationsmodell zur Beschreibung einer individuellen Arbeitsorganisation einführungsvorgestellt.

1. Einleitung

Durch die fortschreitende Individualisierung der Produktion kommt es zu einem verstärkten Einsatz von wandlungsfähigen Produktionssystemen mit hochgradig flexiblen Arbeitssystemen. Um die geforderte Flexibilität auf dem Shopfloor zu erreichen, wird der Mensch zunehmend durch Assistenzsysteme unterstützt. Neben Assistenzsystemen zur Anleitung der durchzuführenden Montageschritte existieren Applikationen zur Arbeitsorganisation. Bei diesen Anwendungen werden den Arbeitspersonen die auszuführenden Aufgaben und ggf. die Reihenfolge der geplanten Ausführung angezeigt. Dies hat zur Folge, dass der Entscheidungsspielraum der Arbeitspersonen durch die vorgegebene starre Aufgabensequenz oftmals sehr eingeschränkt ist und die Unzufriedenheit mit der Arbeitssituation ansteigt.

Im Folgenden wird daher ein Bayessches Netz zur Simulation des individuellen Entscheidungsverhaltens von Arbeitspersonen bei der Auswahl von Aufgaben aus einem Aufgabenpool vorgestellt. Hierzu werden für das Bayessche Netz Wahrscheinlichkeiten berechnet, die eine Simulation des Verhaltens einer Arbeitsperson im Arbeitssystem ermöglichen. Gelingt es, die Entscheidungen hinsichtlich der Auswahl von Aufgaben aus einem Aufgabenpool a-priori zu beschreiben, bietet dies vielfältige Potentiale für die Planung der individuellen Arbeitsorganisation von Arbeitspersonen bei schwach strukturierten Produktionsprozessen.

2. Arbeitssysteme

Arbeitssysteme dienen der Beschreibung von Arbeitssituationen und Arbeitsplätzen. Ein Arbeitssystem ist ein generisches, zeitdifferentes, flexibles, soziotechnisches, menschliches Handlungssystem [1] und lässt sich durch

mengenalgebraische Systemdefinitionen beschreiben [1], [2].

Das in diesem Beitrag herangezogene Arbeitssystem besteht aus einem Führungs- und Planungssystem sowie einem Wirksystem, vgl. Abbildung 1 [3]. Das Führungs- und Planungssystem umfasst die Planung einer Aufgabenausführung und auch eine Entscheidung darüber, wie ein definiertes Ziel erreicht werden kann. Als Information zwischen Führungs- und Planungssystem und Wirksystem fungiert die Arbeitsaufgabe. In diesem Zusammenhang ist zu beachten, dass die Planung und Ausführung einer Aufgabe entsprechend einer Arbeitsteilung durch verschiedene Akteure erfolgen kann. Somit kann die Planung der Reihenfolge einer Aufgabenbearbeitung durch eine andere Arbeitsperson oder einem Algorithmus vorgenommen werden und einer weiteren Arbeitsperson mittels eines Assistenzsystems – z. B. einem Smartphone – angezeigt werden. Eine solche, im Wirksystem verortete Arbeitsperson, wählt eine der verfügbaren und im Assistenzsystem angezeigten Aufgaben für die Bearbeitung aus. Für die ausgewählte Aufgabe versucht die Arbeitsperson durch die Ausführung einer Handlung das gewünschte Arbeitsergebnis in Form eines Arbeitsobjektes zu erzielen. Liegt dieses vor, wird neben dem Arbeitsobjekt eine Information über die vollständig bearbeitete Aufgabe in das Assistenzsystem eingegeben und steht dem Planungs- und Führungssystem zur Verfügung.

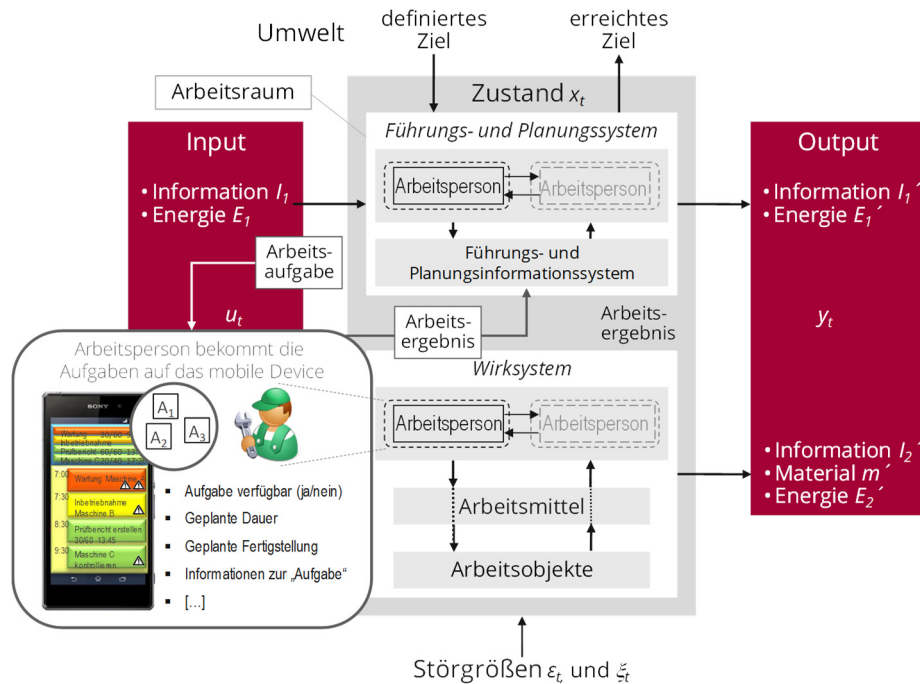


Abbildung 1: Erweitertes Arbeitssystem mit dezentraler Aufgabendisposition i. A. a. [3]

Die Darstellung des Arbeitssystems zu verschiedenen (Zeit-)Zuständen ist durch einen sog. Deltaoperator beschrieben, der die Überführungs- (Relation Input und Zustand) und Markierungsfunktionen (Relation Zustand und Output) im Sinne einer allgemeinen Systemdefinition übernimmt [1], [3].

3. Bayessches Netz für die Aufgabenauswahl

Bei einem Bayesschen Netz handelt es sich um einen gerichteten, azyklischen Graphen aus Knoten und Kanten. Hierbei repräsentieren die Knoten Hypothesen über spezifische Sachverhalte, für die jeweils eine Wahrscheinlichkeit angegeben werden kann. Die Beziehungen zwischen den einzelnen Sachverhalten und somit den Hypothesen werden durch gerichtete Kanten abgebildet.

Die Beobachtung von realen Arbeitsabläufen in einer Produktion mit schwach strukturierten Arbeitsprozessen hat zu der Entwicklung des im Folgenden beschriebenen Bayesschen Netz geführt: Eine Arbeitsperson hat eine Aufgabe aus einem auf einem Smartphone angezeigten Aufgabenpool für die Bearbeitung auszuwählen. Da mehrere Aufgaben von der Arbeitsperson ausgewählt werden können, ist für die weitere Planung der Produktion von Interesse, ob eine spezifische Aufgabe ausgewählt wird. Um eine solche Einschätzung vornehmen zu können, sind entsprechend der Temporal Motivation Theory zumindest die Faktoren Wichtigkeit der Aufgabe sowie deren Dringlichkeit zu berücksichtigen: Je größer die Wichtigkeit bzw. je höher die Dringlichkeit der Aufgabe im Assistenzsystem angezeigt wird, desto größer ist die Wahrscheinlichkeit, dass die Arbeitsperson die Aufgabe für die Bearbeitung auswählt. Andererseits

kann es bei verschiedenen Arbeitspersonen aufgrund von Persönlichkeitsmerkmalen sowie Erfahrungswissen zu einer unterschiedlichen Reihenfolge hinsichtlich der Auswahl möglicher Aufgaben aus dem Aufgabenpool kommen. Daher ist neben der objektiven, auf dem Smartphone vom Planungsalgorithmus angezeigten Wichtigkeit und Dringlichkeit einer Aufgabe, auch die von der Arbeitsperson wahrgenommene subjektive Wichtigkeit und subjektive Dringlichkeit von Relevanz.

Die für ein Arbeitssystem auftretenden Zusammenhänge lassen sich für eine Aufgabe mittels des Bayesschen Netzes in Abbildung 2 beschreiben. Es werden fünf Knoten eingeführt, die jeweils die folgenden Hypothesen repräsentieren:

- Die Arbeitsperson wählt die Aufgabe für die Bearbeitung aus bzw. nicht aus (Knoten A).
- Die Aufgabe hat eine hohe, mittlere, niedrige Dringlichkeit (Knoten D).
- Die Aufgabe hat eine hohe, mittlere, niedrige Wichtigkeit (Knoten W).
- Die Aufgabe hat eine hohe, mittlere, niedrige subjektiv von der Arbeitsperson empfundene Dringlichkeit (Knoten sD).
- Die Aufgabe hat eine hohe, mittlere, niedrige subjektiv von der Arbeitsperson empfundene Wichtigkeit (Knoten sW).

Zwischen den fünf Knoten werden die in Abbildung 2 gezeigten vier Kanten eingeführt. Die Kanten beschreiben, dass die subjektiv wahrgenommene Wichtigkeit und Dringlichkeit einen Einfluss auf die Einschätzung hat, ob die Arbeitsperson eine spezifische Aufgabe aus dem Auf-

gabenpool auswählt. Ferner ist die Auswahl einer Aufgabe wahrscheinlicher, wenn deren im Assistenzsystem angezeigte Wichtigkeit und Dringlichkeit höher ist.

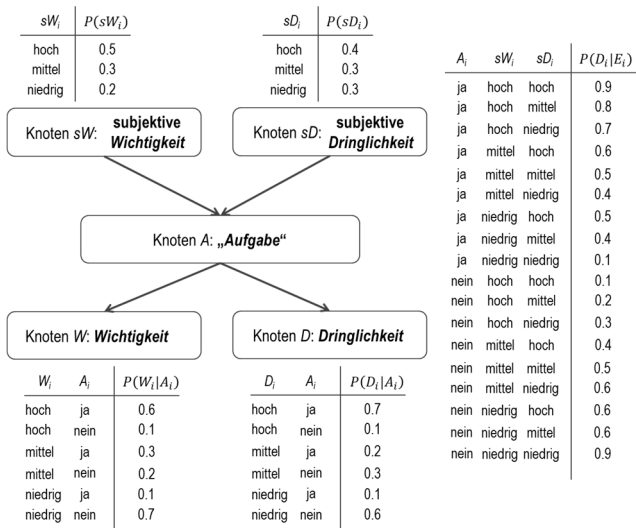


Abbildung 2: Bayessches Netz – a-priori Wahrscheinlichkeiten

Ausgangspunkt sind a-priori festgelegte Wahrscheinlichkeiten für die einzelnen Knoten des Bayesschen Netzes, vgl. Abbildung 2. Für die Knoten ohne Vorgänger werden die Wahrscheinlichkeiten je zulässiger Ausprägung (Hypothese) angegeben, während für die anderen Knoten bedingte Wahrscheinlichkeiten in Abhängigkeit der Ausprägung der jeweiligen Vorgänger festzulegen sind. Die Angabe der initialen Wahrscheinlichkeiten basiert zunächst auf dem Erfahrungswissen über das Arbeitssystem. Mit weiteren Beobachtungen über das Verhalten der Arbeitspersonen unterliegen die Werte einer kontinuierlichen Anpassung. Für eine detaillierte Beschreibung der Initialisierung und Aktualisierung von Bayesschen Netzen sei verwiesen auf [4].

Aus den initialen Wahrscheinlichkeiten des Netzes sind je Knoten die folgenden drei Wahrscheinlichkeiten zu bestimmen:

- **π-Wert:** Gegenwärtige Stärke der kausalen Unterstützung einer Hypothese durch die Vorgängerknoten,
- **λ-Wert:** Gegenwärtige Stärke der diagnostischen Unterstützung einer Hypothese durch die Nachfolgerknoten,
- **BEL-Wert:** Der Belief-Wert gibt das Gesamtvertrauen in eine Hypothese an.

Der BEL-Wert einer Hypothese K_j für einen Knoten K berechnet sich dabei aus dem π -Wert und dem λ -Wert ergänzt um einen Normierungsfaktor α , der so gewählt wird, dass gilt:

$$BEL(K_j) = \alpha \cdot \lambda(K_j) \cdot \pi(K_j)$$

$$\sum_{j=1}^h BEL(K_j) = 1$$

Das Bayessche Netz in Abbildung 3 beschreibt das Arbeitssystem, bevor eine Aktualisierung der Informationen zu den Aufgaben im Aufgabenpool erfolgt: Zunächst wird geschätzt, wie die Wichtigkeit und Dringlichkeit der Aufgabe von der Arbeitsperson wahrgenommen wird und es wurden Annahmen darüber abgeleitet, wie wahrscheinlich es ist, dass die Arbeitsperson die Aufgabe auswählt. Im Folgenden wird nun das gleiche Arbeitssystem zu einem späteren Zeitpunkt betrachtet. Der Aufgabenpool auf dem Smartphone der Arbeitsperson wurde vom übergeordneten Manufacturing Execution System (MES) aktualisiert, so dass ein veränderter Wert für die Dringlichkeit der hier betrachteten Aufgabe vorliegt. Die Auswahl einer Aufgabe durch die Arbeitsperson steht aber noch aus.

Bei der durch das MES veränderten objektiven Dringlichkeit handelt es sich um einen direkt beobachtbaren Sachverhalt. Durch die nun im Vergleich zu anderen Aufgaben zu erwartende hohe Dringlichkeit der Aufgabe verändert sich die Wahrscheinlichkeiten der drei Hypothesen im Knoten D zu:

- $P(D_1) = P(D = \text{hoch}) = 0,7,$
- $P(D_2) = P(D = \text{mittel}) = 0,2,$
- $P(D_3) = P(D = \text{niedrig}) = 0,1.$

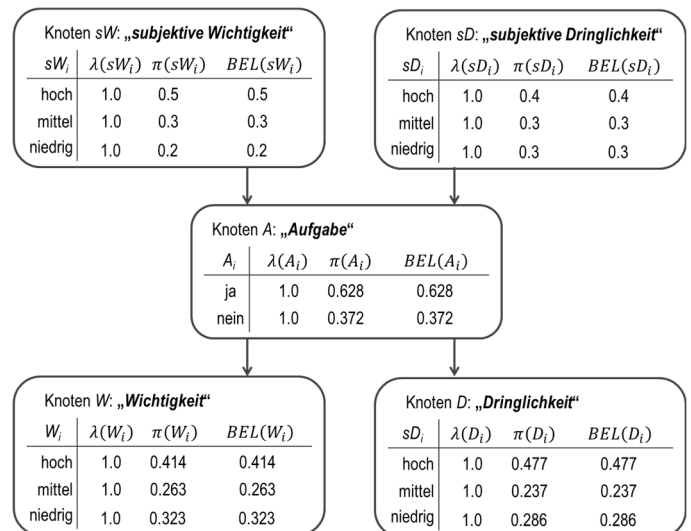


Abbildung 3: Bayessches Netz - nach Initialisierung der Knoten

Die neue Information hat nicht nur Auswirkungen auf den Knoten D , sondern auch auf die anderen Knoten des Netzes. Wird der Arbeitsperson eine hohe Dringlichkeit für eine Aufgabe auf dem Smartphone angezeigt, so ist es nicht mehr plausibel anzunehmen, dass die Wahrscheinlichkeiten für die beiden Hypothesen, dass die Arbeitsperson die Aufgabe auswählt bzw. nicht auswählt annähernd gleich groß sind. Folglich führt die neue Information zu einer Aktualisierung des Knoten A , indem sich die Wahrscheinlichkeit für die Hypothese A_1 – Auswahl der Aufgabe – erhöht und sich die Wahrscheinlichkeit der Hypothese A_2 – keine Auswahl der Aufgabe – im gleichen Maße reduziert. Die Auswirkungen einer neuen Information sind nicht nur auf die direkten Vorgängerknoten beschränkt. Vielmehr propagiert die aktualisierte

Information durch das Netz und die Wahrscheinlichkeiten in den einzelnen Knoten müssen angepasst werden. Die Aktualisierung der einzelnen Knoten erfolgt entsprechend der Darstellung nach [4].

Das nach der Aktualisierung aller Knoten vorliegende Netz ist in Abbildung 4 dargestellt. Die Wahrscheinlichkeit der Auswahl der Aufgabe hat sich durch die im Assistenzsystem angezeigte Dringlichkeit von 62,8% auf 82,8% erhöht. Aufgrund der vom Knoten A enthaltenen λ -Nachrichten wurden innerhalb der Knoten sD und sW jeweils die λ - und BEL-Vektoren modifiziert. So ist davon auszugehen, dass die objektive, beobachtete Dringlichkeit auch zu einer Erhöhung der durch die Arbeitsperson empfundenen subjektiven Dringlichkeit führt.

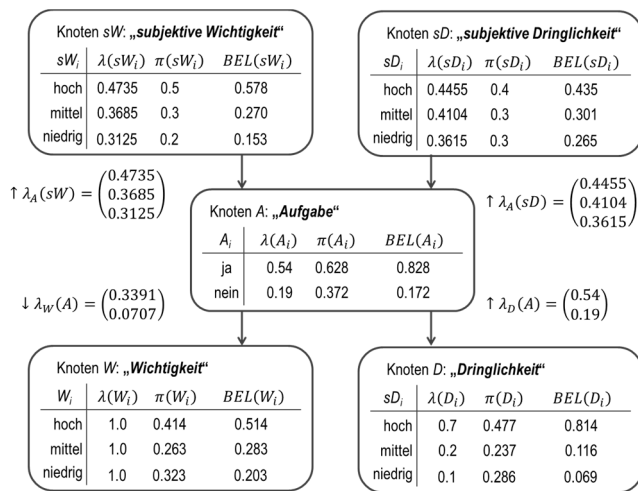


Abbildung 4: Bayessches Netz - nach Aktualisierung der Knoten

5. Simulationsumgebung

Das Simulationsmodell basiert auf den vorangegangenen Veröffentlichungen zum aktororientierten, personenzentrierten Simulationswerkzeug [5]. Im Folgenden wird sich daher auf die Darstellung der simulierten Aufgabenauswahl durch eine Arbeitsperson beschränkt. Hierbei bleibt das Grundkonzept bestehen, dass die Simulation der individuellen Arbeitsorganisation der Arbeitsperson das dynamische Verhalten des Arbeitssystems bestimmt, vgl. Abbildung 5.

Die Zustände des Arbeitssystems entstehen durch einen iterativen Wechsel der Arbeitsperson zwischen ihren Status „Sperrzeit“, „Organisieren“ und „Bearbeiten“. Der Status „Sperrzeit“ bildet die Nicht-Verfügbarkeit der Arbeitsperson, z.B. aufgrund einer Pause ab. Im Status „Organisieren“ berechnet die Arbeitsperson für die ausführbaren Aufgaben entsprechend der Temporal Motivation Theory personenbezogene Werte für die subjektive Dringlichkeit und Wichtigkeit der Aufgaben. Die jeweils ermittelten Werte bilden die Grundlage für eine Zuordnung der Aufgabe in die Kategorien „niedrig“, „mittel“ und „hoch“ der Variablen sD, sW, D und W sowie der Zuweisung von Wahrscheinlichkeiten zu den einzelnen Knoten. Werden zudem neue Informationen zu der „objektiven“ Dringlichkeit und Wichtigkeit einer Aufgabe

vom MES bereitgestellt, so erfolgt nach der Aktualisierung der Wahrscheinlichkeiten im Knoten A die Auswahl einer Aufgabe durch die Arbeitsperson entsprechend der berechneten Wahrscheinlichkeiten. Mit der Entscheidung über die Auswahl einer Aufgabe erfolgt der Wechsel vom Planungs- und Führungssystem in das Wirksystem des Arbeitssystems. D. h., für die ausgewählte und eindeutig definierte Arbeitsaufgabe wird im Status „Bearbeiten“ ein Arbeitsergebnis durch die Arbeitsperson erzeugt. Während der simulierten Aufgabebearbeitung erhöht sich der Bearbeitungsgrad der Aufgabe. Ferner prüft die Arbeitsperson, ob die Arbeitsaufgabe fertig gestellt ist oder ob eine personenindividuell festgelegte Zeitspanne abgelaufen ist. In beiden Fällen wechselt die Arbeitsperson in den Status „Organisieren“, priorisiert den Aufgabenpool entsprechend der neu berechneten Wahrscheinlichkeiten des Netzes und wählt eine neue oder die gleiche Aufgabe für die Bearbeitung aus. Bei einer vollständigen Bearbeitung einer Aufgabe wird ein Arbeitsobjekt als Output erzeugt, welches den erforderlichen Input für die initiale Bearbeitung der nachfolgenden Aufgaben darstellt.

Der implementierte Algorithmus durchläuft während einer Simulation immer die gleiche Abfolge der beschriebenen Status. Die jeweiligen Bearbeitungsstände der einzelnen Aufgaben im Aufgabenpool in Verbindung mit den vom MES-System bereitgestellten Informationen ergeben die Wahrscheinlichkeiten für die verschiedenen Bayesschen Netze des Arbeitssystems.

6. Fazit

Das Konzept zur Simulation des Entscheidungsverhaltens von Arbeitspersonen bei schwach strukturierten Arbeitsprozessen ermöglicht eine prospektive Beschreibung des individuellen Verhaltens bei der Auswahl von Aufgaben. Gelingt es, basierend auf den von einem MES bereitgestellten aufgabenbezogenen Informationen, die Entscheidung einer Arbeitsperson in Bezug auf die Aufgabenauswahl zu prognostizieren, können auch schwach strukturierte Produktionsprozesse von Manufakturen effizient geplant werden. So könnte durch eine gezielte Steuerung der für eine Aufgabenauswahl relevanten Einflussfaktoren das Entscheidungsverhalten so beeinflusst werden, dass hieraus eine für das Arbeitssystem optimale Aufgabensequenz entsteht, ohne dass der von der Arbeitsperson erlebte Entscheidungsspielraum sich verringert.

Das beschriebene Konzept des Bayesschen Netzes muss noch in das bestehende aktororientierte Simulationsmodell implementiert und anschließend verifiziert werden. So lag der Schwerpunkt der bisherigen Arbeiten ausschließlich auf der Erprobung der Algorithmen zur Abbildung eines deterministischen Entscheidungsver-

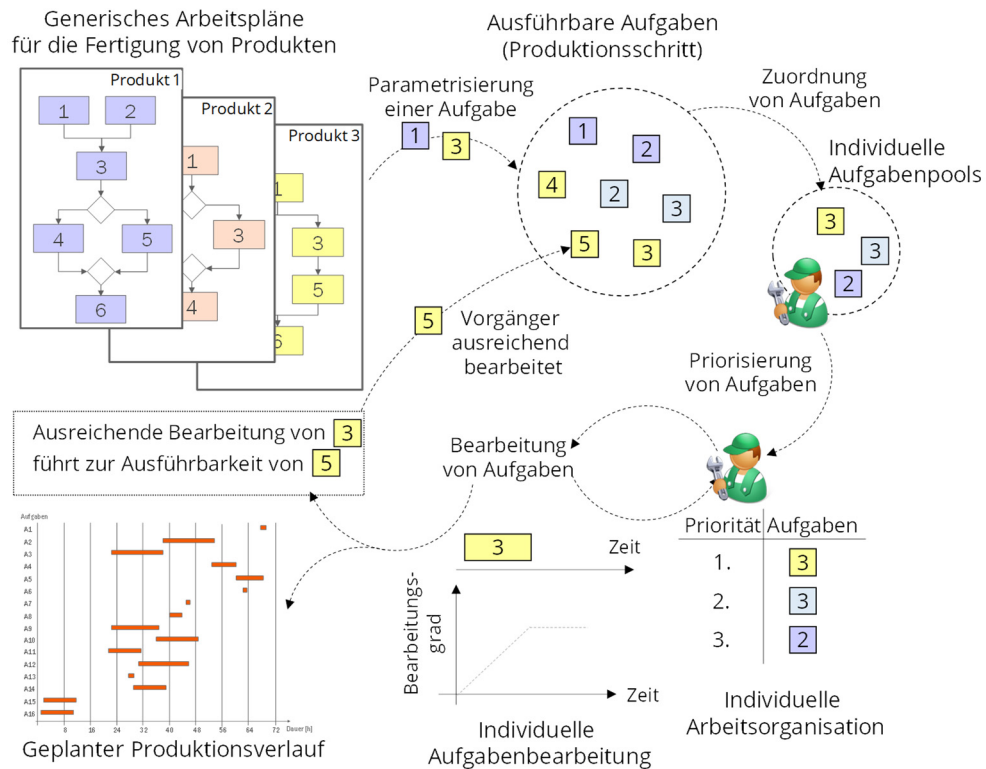


Abbildung 5: Konzept für ein aktor-orientiertes Simulationsmodell

haltens von Arbeitspersonen in einem Arbeitssystem sowie der Analyse von stochastischen Zeitaufwänden. Ferner bedarf es empirischer Studien, um den Einfluss von Persönlichkeitsmerkmalen auf die subjektiv erlebte Dringlichkeit und Wichtigkeit einer Aufgabe nachzuweisen und so die Vorhersagegüte des Modells zu erhöhen.

Literaturverzeichnis

- [1] G. Ropohl, Allgemeine Technologie: Eine Systemtheorie der Technik, Universitätsverlag Karlsruhe, KIT Scientific Publishing, 2009.
- [2] V. Hubka, Theorie technischer Systeme: Grundlagen einer wissenschaftlichen Konstruktionslehre, Berlin, Heidelberg: Springer, 1984.
- [3] C. M. Schlick, R. Bruder, H. Luczak H, Arbeitswissenschaft, Berlin, Heidelberg: Springer, 2010.
- [4] B. Kipper, T. Brants, M. Plach, R. Schäfer, Bayessche Netze: Ein einführendes Beispiel. Technical Report 4, Graduiertenkolleg Kognitionswissenschaft, Saarbrücken, Universität des Saarlandes, 1995.
- [5] S. Tackenberg, Konzeption und Entwicklung einer mehrkriteriellen evolutionären Metaheuristik zur Planung wissensintensiver Dienstleistungen, Aachen: Shaker Verlag, 2016.

Dynamische Losbildung für eine dezentral geplante und gesteuerte Produktion

Martin Krockert, Marvin Matthes, Torsten Munkelt
HTW Dresden, Friedrich-List-Platz 1, 01069 Dresden

Um die Produktion bestmöglich auszulasten, sowie schnell und termintreu zu produzieren, muss die Produktionsplanung und -steuerung anstehende Arbeitsgänge zu Losen zusammenzufassen, wodurch sie die Anzahl der benötigten Rüstvorgänge und somit auch die kumulierte Rüstzeit reduziert. Während für zentral geplante und gesteuerte Produktionen bereits viele Heuristiken zur Losbildung existieren, gibt es für dezentral geplante und gesteuerte Produktionen noch keine zufriedenstellende Lösung zur Losbildung. Wir stellen ein neues Verfahren zur dynamischen Losbildung für eine dezentral geplante und gesteuerte Produktion vor, dass unter Berücksichtigung der Durchlaufterminierung und des Kapazitätsangebotes Arbeitsgänge mit gleichem Rüstbedarf dynamisch zu Losen zusammenfasst. Das Verfahren ergänzt unsere sich selbst organisierenden Produktion um die Losbildung und zeigt in Simulationen seine Überlegenheit gegenüber gängigen Verfahren.

1. Einleitung

Unser Projektpartner aus der metallverarbeitenden Industrie steht vor der Herausforderung, viele Produkte mit unterschiedlichen Kundenspezifika zumeist auf Abruf herzustellen. Bei dem Bearbeiten dieser Kundenaufträge entfallen durchschnittlich 17% der benötigten Zeit auf Rüstvorgänge. Durch die hohe Diversität der Kundenaufträge eignen sich klassische Losgrößenverfahren nicht zum Zusammenfassen gleichartiger Produkte und damit auch nicht zur Reduktion der benötigten Rüstvorgänge. Um dem Problem der hohen Rüstzeiten zu begegnen, stellen wir ein neues Verfahren zur dynamischen Losbildung für die Anwendung in einer dezentralen Produktion vor. Der Ansatz basiert auf einer Losbildung über Arbeitsgänge mit gleichem Rüstbedarf unter Berücksichtigung der Durchlaufterminierung und des zugrundeliegenden Kapazitätsangebots. Die entstehenden Lose von Arbeitsgängen mit gleichem Rüstbedarf werden in diesem Verfahren als Buckets bezeichnet. Buckets werden während der Laufzeit einer Produktion auf Basis von Ereignissen dynamisch erstellt, erweitert oder aufgelöst. Der neue Ansatz einer dynamischen Losbildung für eine dezentral geplante und gesteuerte Produktion erweitert unsere bisherige Forschung an einer sich selbst organisierenden Produktion. Die sich selbst organisierende Produktion zeigte sich in vorherigen Forschungsergebnissen einer zentral geplanten und gesteuerten Produktion überlegen, da der Produktionsplan in der sich selbst organisierenden Produktion durch lokale Neuplanungen ereignisgesteuert angepasst wird [1]. Die Potenziale einer dezentral geplant und gesteuerten Produktion werden insbesondere bei einer Produktion mit einem hohen Bedarf an Flexibilität deutlich, bei der Unsicherheit in Form von variierenden Zwischenankunftszeiten der Kundenaufträge und Schwankungen der Bearbeitungszeiten von Arbeitsgängen auf die Produktion Einfluss nehmen. In der abschließend

vorgestellten empirischen Studie wird dazu das bisher genutzte Verfahren zur Zusammenfassung von Arbeitsgängen mit dem neu entwickelten Verfahren simulativ verglichen.

2. Problembeschreibung

Die Deutschen Akademie der Technikwissenschaften stellen fest, dass die „[...] Art und Weise der Veredelung und Herstellung von Produkten ... sich noch stärker an den Kundenbedürfnissen ausrichten müssen [...]“ [2]. Unternehmen sind deshalb gefordert, flexibel auf das sich verändernde Nachfrageverhalten von Kunden zu reagieren. Die Bedarfsmengen der Kunden entwickeln sich zu immer kleineren Mengen, die in unregelmäßigen Abständen abgerufen werden. Obwohl die Vorproduktion von größeren Mengen Kosteneinsparungen in der Produktion durch größere Synergieeffekte und weniger Umrüstvorgängen auf den Maschinen ermöglicht, entstehen jedoch zwei potenzielle Probleme: (1.) entstehen Lagerkosten für die vorzeitige Bereitstellung der Produkte und (2.) können bei einer Vorratsproduktion, aufgrund Veränderungen in den Anforderungen der Kundenbestellung, manche Produkte nur mit Verlust verkauft oder müssen vollständig abgeschrieben werden. Diesen Problemen kann begegnet werden, indem die Produktion flexibel auf Änderungen reagiert und nur geringe Stückmengen der jeweiligen Produkte fertigt. Aufgrund der Ungewissheit über Abrufmengen und -zeitpunkte, erzielen Losbildungsverfahren über Materialien hinweg keine wesentliche Effizienzverbesserung der Produktion. Bisher werden Materialien mit gleichem Rüstbedarf für den nächsten Arbeitsschritt vor einer Maschine gesammelt, nach Priorität geordnet und sequenziell abgearbeitet, bis alle Arbeitsgänge mit gleichem Rüstbedarf erschöpft sind. In der Literatur werden Arbeitsgänge mit gleichen Voraussetzungen hinsichtlich des benötigten Rüstbedarfs unter dem Begriff der *Group Technologies*

(kurz GT) geführt. [3,4] Durch das Zusammenfassen von Arbeitsgängen der gleichen GT, findet während der sequenziellen Bearbeitung von Arbeitsgängen der gleichen GT kein Rüstvorgang auf einer Maschine statt. Somit reduziert sich unmittelbar die Rüstzeit für die Maschine. GT wenden ein zweistufiges Verfahren an, bei der zuerst die Gruppe von Arbeitsgängen mittels einer Prioritätsregel ausgewählt und im zweiten Schritt die Bearbeitungsreihenfolge der Arbeitsgänge innerhalb der Gruppe mittels einer Prioritätsregel bestimmt wird [5]. Das bisherige Verfahren setzt dabei auf eine Warteschlange, für welche die Prioritätsregel nach der Zielgröße in der Produktion ausgewählt wird. Ist das Produktionsziel die Sicherstellung einer hohen Termintreue werden terminbezogene Prioritätsregeln, wie z.B. die kürzeste Schlupfzeit, eingesetzt. Weiterhin wird bei der Anwendung von GT in zwei Verfahrenstypen unterschieden:

Exhaustiv Verfahren: Nach Auswahl der GT werden alle Arbeitsgänge der GT bearbeitet, bei denen alle Vorbedingungen erfüllt sind.

Nicht-exhaustiv Verfahren: Nach Auswahl der GT werden sukzessive Arbeitsgänge mit gleichem Rüstbedarf bearbeitet, bei denen alle Vorbedingungen erfüllt sind. Im Gegensatz zu exhaustiv Verfahren kann jedoch auch zwischenzeitlich auf andere GT gewechselt werden.

Frazier zeigt in [5], dass exhaustiv Verfahren zur Bearbeitung von Arbeitsgängen in einer Flow-Shop-Fertigung bereits massive Rüstkosten reduzieren und in vielen Fällen den nicht-exhaustiv Verfahren überlegen sind. Exhaustiv Verfahren führen jedoch tendenziell zu größeren Losen, welche wiederum Maschinen auf einem Rüstzustand halten. In dieser Arbeit soll deshalb ein Verfahren zur dynamischen Losbildung entwickelt werden, welches trotz tendenziell kleinerer Lose den Rüstaufwand reduziert und dadurch in einer dezentral geplanten und gesteuerten Produktion weiterhin die notwendige Flexibilität gewährleistet.

3. Dynamische Losbildung

Das Verfahren für eine dynamische Losbildung basiert auf drei Bestandteilen: den Arbeitsgängen, den verfügbaren Maschinen und dem Bucket-Manager zur Zuordnung von Arbeitsgängen in Buckets. Der *erste Bestandteil* umfasst die Arbeitsgänge aller Fertigungsaufträge. Das Verfahren betrachtet ausschließlich sequenzielle

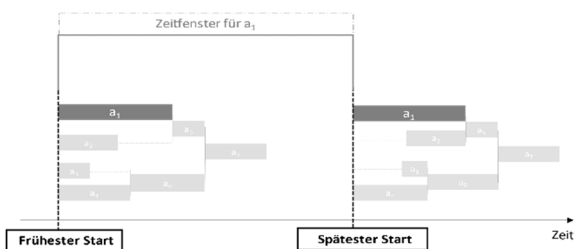


Abbildung 1: Zeitfensters mittels Durchlaufterminierung

Arbeitsgänge. Für jeden Arbeitsgang liegen die folgenden Informationen vor: voraussichtliche Bearbeitungszeit, angestrebter Fertigstellungstermin für den Fertigungsauftrag in der Produktion, Übergangszeit, benötigter Rüstzustand auf einer Maschine sowie das Zeitfenster zur Bearbeitung, welches sich aus frühestem Startzeitpunkt und spätestem Startzeitpunkt der Vorwärts- und Rückwärtsterminierung des Fertigungsauftrags ergibt (siehe Abb. 1). Alle in der Produktion verfügbaren Maschinen bilden den *zweiten Bestandteil*. Jede Maschine ist einer Maschinengruppe zugeordnet und kann mit einer Reihe von Werkzeugen gerüstet werden. Eine Maschine kann jedoch nur mit einem mit einem Werkzeug gleichzeitig gerüstet sein. Jede Kombination einer Maschine mit einem der ihr zur Verfügung stehenden Werkzeuge steht für einen Rüstzustand. Wird eine Maschine mit einem Werkzeug gerüstet, kann die Maschine alle Arbeitsgänge, die diesem Rüstzustand zugeordnet sind, durchführen. Da jeder Maschine ihr eigenes Set an Werkzeugen zur Verfügung steht, können Maschinen derselben Maschinengruppe zur selben Zeit mit dem gleichen Werkzeug ausgerüstet sein. Somit können bei Bedarf Arbeitsgänge mit dem gleichen benötigten Rüstzustand an unterschiedlichen Maschinen parallel bearbeitet werden. Der *dritte Bestandteil* umfasst die Bucket-Manager. Bucket-Manager stellen eine Menge von virtuellen Entitäten dar, welche auf Basis der existierenden Rüstzustände erstellt werden.

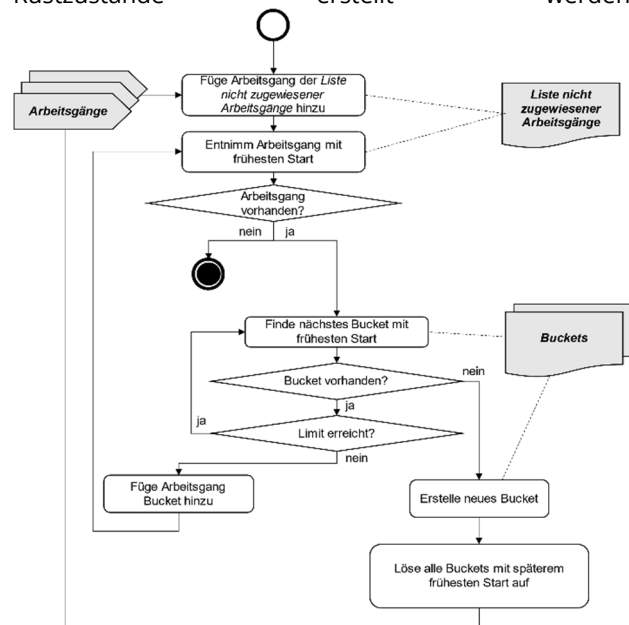


Abbildung 2: Einordnung von Arbeitsgängen in Buckets

Jeder Bucket-Manager organisiert alle Arbeitsgänge, die genau seinem Rüstzustand zugeordnet sind. Der Algorithmus für die Zuordnung von Arbeitsgängen in Buckets ist in Abbildung 2 dargestellt. Nach der Freigabe eines Fertigungsauftrags in die Produktion und der darauffolgenden Stücklistenauflösung, werden alle Arbeitsgänge des Fertigungsauftrags in die Produktion überführt. Mit jedem neu eintreffenden Arbeitsgang wird der Algorithmus zur Einordnung von Arbeitsgängen in Buckets aus-

gelöst, um für den jeweiligen Arbeitsgang ein passendes Bucket zu finden oder zu erzeugen. Im ersten Schritt wird jeder Arbeitsgang der *Liste nicht zugewiesener Arbeitsgänge* hinzugefügt. Die *Liste der nicht zugewiesener Arbeitsgänge* ist sortiert nach dem frühesten Startzeitpunkt, resultierend aus der Vorwärtsterminierung eines jeden Arbeitsganges. Für die Zuordnung zu einem Bucket wird der erste Arbeitsgang aus der *Liste der nicht zugewiesenen Arbeitsgänge* entnommen. Wenn die Liste leer ist und somit alle Arbeitsgänge bereits einem Bucket zugeordnet sind, terminiert der Algorithmus an dieser Stelle. Andernfalls wird für den entnommenen Arbeitsgang zuerst ein passendes Bucket gesucht. Dazu wird aus der Liste der existierenden Buckets jenes Bucket mit dem frühesten Startzeitpunkt ausgewählt. Der früheste Startzeitpunkt eines jeden Buckets entspricht dem frühesten Startzeitpunkt des Arbeitsganges, welcher zuerst dem Bucket hinzugefügt wurde und das Bucket somit erstellt hat. Wird ein Bucket gefunden, wird im nächsten Schritt geprüft, ob das maximale Limit für das Bucket erreicht wurde. Die Bestimmung des Limits erfolgt auf Basis des Kapazitätsangebots nach Formel (1).

$$l_B \geq \sum_{j=1}^{|A_B|} d(a_j) + d(a_i) \quad (1)$$

$$l_B = \min(l_{scope}, l_{setup}) \quad (2)$$

$$l_{scope} = rts_{a_{1B}} - vts_{a_{1B}} \quad (3)$$

$$l_{setup} = f \cdot \frac{|W_{mg}|}{|M_{mg}|} \cdot \frac{\sum_{j=1}^{|A_w|} d(a_j)}{\sum_{j=1}^{|A_{mg}|} d(a_j)} \quad (4)$$

mit:

- rts ... Startzeitpunkt aus der Rückwärtsterminierung
- vts ... Startzeitpunkt aus Vorwärtsterminierung
- d ... Fertigungsdauer in Minuten
- $a \in A$... Arbeitsgang ist ein Spaltenvektor mit $\{d, rts, vts, w, mg\}^T$
- a_i ... aktueller Arbeitsgang
- B ... Bucket ist ein Spaltenvektor mit $\{A, w, mg\}^T$
- f ... vordefinierter Bucket-Faktor
- w ... Werkzeuge
- m ... Maschine
- mg ... Maschinengruppe ist ein Spaltenvektor mit $\{W, M\}^T$
- P ... Produkt (Baugruppe)

Das Verfahren setzt als Bucket-Limit zuerst das ermittelte Zeitfenster des ersten Arbeitsgangs im Bucket nach Formel (3) an. Darüber hinaus wird in Formel (4) eine zweite Grenze für das Limit des Buckets in Abhängigkeit von der Verwendungshäufigkeit des betrachteten Rüstzustands einbezogen. Der Parameter f in Formel (4) beschreibt einen Faktor, der sich aus der zugrundeliegenden Produktion ergibt. Im Rahmen der Untersuchungen wurde ein Faktor in Abhängigkeit von üblichen Schichtzeiten (8 Stunden bzw. 480 Minuten) eingesetzt. In Formel (2) wird das Minimum der beiden zuvor ermittelten Werte mit der Summe über alle Arbeitsgänge im Bucket

und dem noch nicht zugewiesenen Arbeitsgang verglichen. Wird das Limit unterschritten, wird das Arbeitsgang dem Bucket hinzugefügt. Andernfalls wird das nächste Bucket geprüft. Wird kein Bucket ermittelt oder ist bei allen Buckets das Limit erreicht, wird für den Arbeitsgang ein neues Bucket erstellt und der Arbeitsgang dem neu erstellten Bucket hinzugefügt. Bei der Erstellung eines neuen Buckets wird darüber hinaus geprüft, ob weitere Buckets mit dem gleichen Rüstbedarf aber einem späteren „Frühesten Start“ existieren. Buckets mit einem späteren „Frühesten Start“ werden aufgelöst und alle Arbeitsgänge der aufgelösten Buckets werden wiederum über den Algorithmus zur Einordnung von Arbeitsgängen in Buckets einem Bucket hinzugefügt.

4. Empirische Studie

4.1 Experiment Design

Der entwickelte Algorithmus wurde in unsere existierende Simulation für eine sich selbst organisierenden Produktion implementiert und in einer empirischen Studie untersucht. Der Aufbau und die Funktionsweise der sich selbst organisierenden Produktion ist in der vorherigen Veröffentlichung [1] dargestellt. Für die empirische Studie werden Maschinen, Stücklisten und Arbeitspläne erzeugt, die hinsichtlich ihrer statistischen Kenngrößen einer realen Produktion gleichen, aber nicht mit ihr identisch sind. Das Produktionsmodell ist von einer Produktion eines mittelständigen Unternehmens aus der metallverarbeitenden Industrie abgeleitet und in Tabelle 1 dargestellt.

| MASCHINENGRUPPE | TRENNEN | UMFORMEN | FÜGEN | MONTIEREN |
|-------------------------|---------|----------|-------|-----------|
| ANZAHL MASCHINEN | 1 | 2 | 4 | 4 |
| ANZAHL WERKZEUGE | 3 | 4 | 5 | 5 |
| Ø FERTIGUNGSZEIT IN MIN | 25 | 26 | 5 | 18 |
| Ø RÜSTZEIT IN MIN | 20 | 108 | 2 | 6 |
| ANTEIL AN GESAMTZEIT | 33,78% | 35,14% | 6,76% | 24,32% |

Tabelle 1: Produktionsmodell und Arbeitsplankennwerte

Die Produktion unterteilt sich in vier Maschinengruppen. Die Maschinengruppen sind nach dem Flow-Job-Prinzip organisiert. Für jede Maschinengruppe ist eine Anzahl an Maschinen und Werkzeugen definiert. Die Anzahl der Werkzeuge ergibt die Anzahl der möglichen Rüstzustände. Für Fertigungs- und Rüstvorgänge fällt jeweils eine Zeitspanne an, welche durch eine durchschnittliche Fertigungs- und Rüstzeit je Maschinengruppe festgelegt wird. Der Anteil jeder Maschinengruppe an der Gesamtzeit wird über die durchschnittliche Fertigungszeit festgelegt. Auf Basis der Realdaten des Unternehmens wurden Kenngrößen für den Mehrfachverwendungsgrad (MV) von 4,14 nach Formel (5) und Komplexitätsgrad (KG) von 1,57 nach Formel (6) der Produktion ermittelt [6]. Für die Studie wurde aufbauend auf den Kennwerten für die Mehrfachverwendung und den Komplexitätsgrad eine bis zu drei Stufen tiefe Produktstruktur von 1000 Endprodukten bestehend aus 344 unfertigen

$$MV = \frac{\sum_{p \in P} |succ(p)|}{|P \setminus P_{Endprodukt}|} \quad (5)$$

$$KG = \frac{\sum_{p \in P} |pred(p)|}{|P \setminus P_{Kaufteil}|} \quad (6)$$

Zwischenerzeugnissen und 144 Kaufteilen generiert. Die unfertigen Erzeugnisse und Endprodukte werden jeweils in durchschnittlich fünf Arbeitsgängen gefertigt, denen gleichverteilt eine Maschinengruppe und ein Werkzeug der jeweiligen Maschinengruppe zugeordnet wird. Während der Simulation werden über einen Zeitraum von 4 Wochen Kundenaufträge mit einer exponentialverteilten Zwischenankunftszeit generiert. Für jeden Auftrag wird der Liefertermin individuell, d.h. anhand der Produktionsdauer des Auftrags, bestimmt.

4.2 Ergebnisse der empirischen Studie

Die Ergebnisse der empirischen Studie hinsichtlich der Auslastung und der Rüstzeiten beider Verfahren sind in Tabelle 2 dargestellt. Mit dem Verfahren zur dynamischen Losbildung konnten die Rüstzeit und folglich auch die Gesamtauslastung der Produktion deutlich reduziert werden. Von ursprünglich 17,35%, konnte das Verfahren zur dynamischen Losbildung die Rüstzeiten auf 13,09% reduzieren. Im Verhältnis entspricht das wiederum einer Reduzierung von 24,52%. Die Rüstzeit konnte insbesondere bei denjenigen Maschinengruppen mit einem hohen Rüstaufwand und einer hohen Auslastung reduziert werden. Im Verlauf der empirischen Studie wurden neben der Auslastung und Rüstzeit, weitere Leistungsparameter ausgewertet und in Tabelle 3 dargestellt.

| | EXHAUSTIV LOSBIILDUNG | DYNAMISCHE LOSBIILDUNG | DIFFERENZ |
|---------------------|--------------------------|---------------------------|-----------|
| LIEFERTERMINTREUE | 98,35% (2) | 95,45% (5) | -2,94% |
| Ø VERSPÄTUNG | 125 MIN | 210 MIN | 67,71% |
| Ø TERMINABWEICHUNG | -744 MIN | -750 MIN | 0,81% |
| Ø DURCHLAUFZEIT | 592 MIN | 586 MIN | -1,01% |
| LAGERKOSTEN/PRODUKT | 1,66 EUR | 0,93 EUR | -43,97% |

Tabelle 3: Vgl. Exhaustive- und Dynamische Losbildung
Obwohl beide Verfahren durchschnittlich 750 Minuten vor dem Termin abschließen, kann die Exhaustive Losbildung eine geringfügig höhere Liefertermintreue erreichen. Ebenso erhöht sich die durchschnittliche Verspätung von der Exhaustive Losbildung (125 Minuten) im Vergleich zur Dynamischen Losbildung (210 Minuten). In den terminbezogenen Leistungsparametern zeigt sich

| MASCHINENGRUPPE | EXHAUSTIV LOSBIILDUNG | | DYNAMISCHE LOSBIILDUNG | | VERÄNDERUNG DURCH - DYNAMISCHE LOSBIILDUNG | |
|-----------------|-----------------------|------------------|------------------------|------------------|---|------------------|
| | RÜSTZEIT | GESAMTAUSLASTUNG | RÜSTZEIT | GESAMTAUSLASTUNG | RÜSTZEIT | GESAMTAUSLASTUNG |
| TRENNEN | 14,21% | 67,96% | 14,80% | 68,54% | +3,94% | +0,83% |
| UMFORMEN | 52,79% | 80,95% | 36,96% | 65,11% | -42,84% | -24,34% |
| FÜGEN | 0,58% | 3,28% | 0,10% | 2,78% | -460,00% | -18,11% |
| MONTIEREN | 1,80% | 11,56% | 0,51% | 10,29% | -250,29% | -12,36% |
| GESAMT | 17,35% | 40,94% | 13,09% | 36,68% | -24,52% | -10,41% |

Tabelle 2: Vergleich der Maschinengruppen der Exhaustiv- und Dynamische Losbildung

das bisherige Verfahren der Exhaustive Losbildung somit überlegen. In den beiden zuletzt betrachteten Kennwerten für die durchschnittliche Durchlaufzeit je Auftrag und die Lagerkosten je Produkt ist die Dynamische Losbildung überlegen. Während sich die Durchlaufzeit nur geringfügig verbessert, können die Lagerkosten je Produkt von 1,66 Euro/Stück auf 0,93 Euro/Stück um fast 44% reduziert werden.

5. Zusammenfassung und Ausblick

Im Rahmen des Forschungsvorhabens wurde auf Basis von Arbeitsgängen mit gleichem Rüstbedarf ein Verfahren zur dynamischen Losgrößenbildung konzipiert, prototypisch implementiert und simulativ gegen die exhaustiv Losbildung evaluiert, das in der realen Produktion verwendet wird. Die dynamische Losgrößenbildung und die Evaluationsstudie sind Ergebnisse unserer Forschungsarbeit. Die Ergebnisse zeigen, dass das entwickelte Verfahren besonders bei hohem Rüstaufwand und einer hohen Auslastung der Maschinen die Rüstkosten erheblich senkt und gleichzeitig die Flexibilität der Produktion gewährleistet.

Danksagung

Das Forschungsvorhaben wird im Rahmen des Förderprogramms „Forschung an Fachhochschulen“ durch das Bundesministerium für Bildung und Forschung gefördert (Zeichen: 13FH133PX8).

Literaturverzeichnis

- [1] T. Munkelt, M. Krockert, in 25. *Interdisziplinäre Wissenschaftliche Konferenz Mittweida 2018* (Eds: Hochschule Mittweida)
- [2] G. Schuh, R. Anderl, J. Gausemeier, W. Wahlster, *Industrie 4.0 Maturity Index: Die digitale Transformation von Unternehmen gestalten*, München 2017.
- [3] I. Ham, K. Hitomi, and T. Yoshida, *Group technology: Applications to production management, ser. International series in management science/operations research*. Boston: Kluwer-Nijhoff, 1985.
- [4] A. Klausnitzer, J. S. Neufeld, and U. Buscher, *Scheduling dynamic job shop manufacturing cells with family setup times: a simulation study*. U. Logist. Res., 2017.
- [5] Frazier, G. V. *An evaluation of group scheduling heuristics in a flow-line manufacturing cell*. International Journal of Production Research, (1996) 34(4):959–976.
- [6] C. E. Heinrich, *Mehrstufige Losgrößenplanung in hierarchisch strukturierten Produktionsplanungssystemen*. Springer Berlin Heidelberg, Berlin, Heidelberg 1987.

Entwicklung eines Systems zur Planung und Fertigung patientenspezifischer Schädelimplantate unter Verwendung additiver Fertigungsverfahren und Techniken des maschinellen Lernens

Michael Kaiser, Patrick Knorr

Westsächsische Hochschule Zwickau, Kornmarkt 1, 08056 Zwickau

Mit Hilfe von Schädelimplantaten werden in der sogenannten Kranioplastik Fehlbildungen sowie durch Trauma oder Krankheit entstandene Schädeldefekte behandelt. Die derzeitige Entwicklung des Implantats basiert oftmals auf Spiegelung der unbeschädigten Schädelhälfte, was durch die Patienten im Nachhinein als negativ betrachtet wird. Andere Schädelimplantate werden gar erst im OP angepasst. Mittels eines patientenspezifisch, additiv hergestellten Implantats, basierend auf den umliegenden Schädelkrümmungen sollen diese Defizite beseitigt werden. Des Weiteren ist die Implantatentwicklung gegenwärtiger Systeme zu langwierig. Mittels einer Cloud-basierten Prozesskette soll die Abstimmung zwischen Konstrukteur, Fertiger und Chirurg verbessert sowie die Entwicklungszeit deutlich verkürzt werden.

Ein Schädeldefekt, unabhängig davon ob es sich um eine angeborene Fehlbildung, die Folge eines Unfalls (Trauma) oder eine Krankheit (Tumor etc.) handelt, gefährdet die empfindlichen Strukturen des Hirns und muss daher umgehend behandelt werden. Derartige operative Eingriffe am Gehirn und Schädel sind, noch weit mehr als andere Operationen, mit großen Ängsten der Patienten und Angehörigen verbunden. Diese Ängste sind verständlich, weil es sich beim Gehirn um ein außerordentlich sensibles Organ handelt, welches entstandene Schäden in der Regel nicht rückbilden kann. Mithilfe der Kranioplastik, welche routinemäßig in der Neurochirurgie durchgeführt wird, ist es möglich großflächige Kopfverletzungen zu behandeln indem die ursprüngliche Kontur des Schädels durch ein Implantat wiederhergestellt wird. (Abbildung 1)

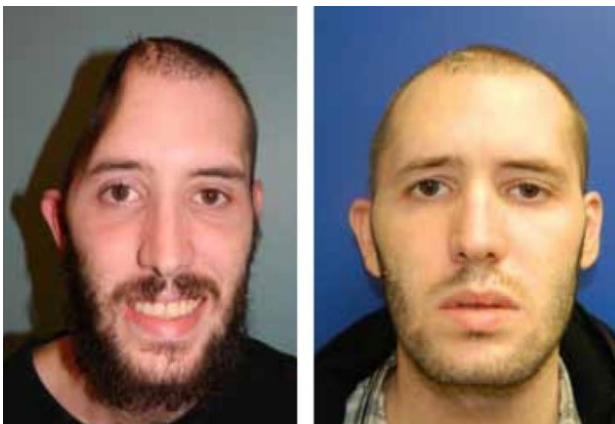


Abbildung 1: Patient vor und nach der Kranioplastik [1]

Der Einsatz des Schädelimplantats hat sowohl kosmetische als auch schützende Wirkung. Um die knöcherne Struktur des Schädels nach einer Fraktur wiederherzustellen, werden derzeit Titanetze verwendet. Die Größe wird entsprechend einer Abschätzung basierend auf CT-Bildern (Computertomografie) aus einer ebenen Titanetz-Platte manuell durch den Chirurgen zugeschnitten und vorgeformt. Diese Implantate sind aufgrund ihres

Aufbaus und der relativ hohen Steifigkeit nur über herkömmliche Zugangswege im Patienten positionierbar. Die Gefahr der Verletzung anatomischer Risikostrukturen, wie beispielsweise des Gehirns ist relativ hoch. Für bisherige Schädelimplantate werden die Materialien Titan(-legierungen) oder Hydroxylapatit verwendet. Beide haben hohen Infektionsraten und damit chirurgische Revisionseingriffe zur Folge. Es ist von einer postoperativen Komplikationsrate von bis zu 40% auszugehen. Der Erfolg von Schädelimplantaten hängt von einer Vielzahl an Faktoren ab, wie z.B. der Größe und Position der Kranioplastik sowie dem Zustand des Patienten, dem verwendeten Material und der Implantationszeit. [2]

Anspruch dieses Projektes ist es einerseits die hohe Komplikationsrate durch Auswahl und Verwendung geeigneter Materialien zu senken und andererseits die Zeit bis zur Fertigstellung des patientenspezifischen Schädelimplantats zu minimieren. Basierend auf diesem Anspruch lassen sich folgende Projektziele definieren.

Das erste Projektziel ist die Entwicklung einer Prozesskette zur Herstellung patientenspezifischer Schädelimplantate mit Hilfe der neuen 3D-Drucksysteme, mit denen Schädelimplantate aus verschiedenen Materialien wie Polyetheretherketon (PEEK)/ Polyetherketonketon (PEKK), Biokeramik und Titan individuell für jeden Patienten hergestellt werden können. Durch anschließende Beschichtung der additiv hergestellten Implantate soll zusätzlich die Biokompatibilität und das Einwachsverhalten verbessert werden.

Zweites Projektziel ist die Entwicklung einer Cloud-basierten Planungssoftware mit Anwendung von Methoden des maschinellen Lernens. Diese soll relevante anatomische Strukturen innerhalb der CT-Bilder segmentieren und automatisch verschiedene Designversionen des Implantats erstellen, welche dem Chirurgen zur Verfügung gestellt werden. Die zu entwickelnde Prozesskette soll eine Implantatlieferung innerhalb von 48h ermöglichen (momentan etwa 2-6 Wochen). Für die CT-basierte

präoperative Planung des patientenspezifischen Schädelimplantats wird derzeit eine Software eingesetzt, welche mehrere Stunden benötigt, um die relevanten anatomischen Strukturen zu segmentieren und das CAD-Modell des Implantats zu erstellen. Zudem soll mittels Cloud die Interaktion zwischen Chirurgen und Ingenieuren verbessert und intensiviert werden, um bestmögliche Implantate zu gewährleisten.

Die aus sieben Teilen bestehende Prozesskette zur Herstellung der patientenspezifischen Schädelimplantate unterscheidet sich annähernd in jedem Teil von der einer herkömmlichen Implantatentwicklung. Neben der standardmäßigen Bildgebung durch CT- und MRT-Bilder wird die Bildgebung durch einen intraoperativen 3D-Scan erweitert um beispielsweise Schädelfragmente erfassen zu können. Die durch künstliche Intelligenz (KI) unterstützte Planungssoftware ermöglicht eine weitestgehend automatisierte Segmentierung des Knochendefekts. Die Konstruktion des Schädelimplantats orientiert sich an den Krümmungen der umliegenden, unbeschädigten Schädelbereiche um ein optimal an den Patientenschädel angepasstes Implantat zu gewährleisten. (Abbildung 2)

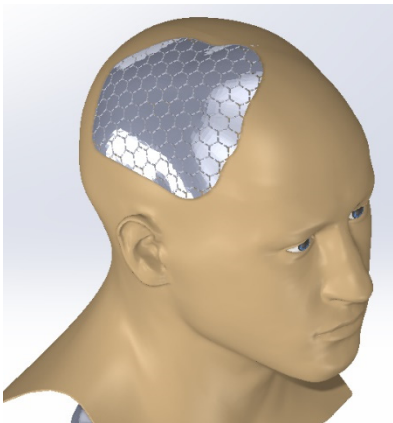


Abbildung 2: Skizze eines patientenspezifischen Implantats (perfekte Kontureinpassung)

Nachteil der sonst üblichen Spiegelung der intakten Schädelhälfte ist die Vereinfachung der Schädelgeometrie, welche die Asymmetrie der anatomischen Details außen vorlässt. [3] Dies führt dazu, dass das Implantat die ursprüngliche Schädelform des Patienten nicht exakt rekonstruiert, was durch den Betroffenen häufig als störend empfunden wird. Im Bereich der Fertigung soll sowohl der 3D-Druck von hochfesten Kunststoffen wie PEEK als auch Titanlegierungen wie etwa Ti6Al14V durch SLM-Verfahren (Selective Laser Melting) Anwendung finden. Wesentlicher Vorteil dieses Verfahrens ist die Strukturvariabilität welche es ermöglicht hochporöse und gleichzeitig mechanisch stabile Strukturen zu realisieren. [4] Insbesondere die Herstellung von Titan-implantaten erfordert eine anschließende Beschichtung, um eine bestmögliche Biokompatibilität zu gewährleisten. [5] Es ist geplant zusätzlich ausgewählte Implantate mit einer Mikrostruktur zu versehen, um das Einwachsverhalten positiv zu beeinflussen.

Im Teilprojekt „Konstruktive Implantatentwicklung mit Variantenerstellung und Analyse prozessbedingter Abweichungen“ wird zu Beginn eine Defizitanalyse bestehender Systeme hinsichtlich verwendeter Materialien und Komplikationen durchgeführt. Die angestrebten Verbesserungen werden in einem Lastenheft definiert. Im Anschluss erfolgt die Erstellung eines ersten CAD-Modells des Implantat-Prototyps. Die Basis dafür bilden die CT-Aufnahmen der Klinik unter Berücksichtigung der intraoperativ erstellten 3D-Scans des Defekts und eventuellen Schädelfragmenten. Nach der additiven Fertigung der Prototypen aus Titan-(legierung) und PEEK werden die Bauteile mittels μ -CT aufgenommen. Durch den Vergleich dieser Aufnahmen mit dem CAD-Datensatz werden die fertigungsbedingten Abweichungen erfasst und für den weiteren Entwicklungsvorgang dokumentiert. In einem weiteren Arbeitspaket werden unterschiedliche Implantatmodelle erstellt. Dabei sollen die Dicke des Schädelimplantats, die Anzahl und Ausführung der Befestigungslaschen sowie die Fase entlang der Kontur variiert werden. Letzteres soll sowohl die Positionierung, als auch das Einwachsverhalten durch eine Oberflächenvergrößerung, verbessern. Die dabei generierten Implantatvarianten werden in eine eigens entwickelte Cloud geschoben und somit der medizinischen Planungssoftware zur Verfügung gestellt. Dem Chirurgen ist es dadurch möglich in kürzester Zeit die verschiedenen Implantatmodelle in der Software zu positionieren und seine Vorzugsvariante zu definieren. Ein zusätzlicher Schwerpunkt des hier beschriebenen Teilprojektes bildet die Mikrostrukturierung und Gewichtsoptimierung des Implantats. Vor der Strukturierung werden dafür die Einwachsflächen der gedruckten Schädelimplantate mittels Mikroskop aufgenommen und die Rauheiten über der Fläche als Basis ermittelt. Im Anschluss an die Mikrostrukturierung, beispielsweise durch Laser, wird dieser Vorgang wiederholt und die Werte verglichen. Ziel der Oberflächenbearbeitung ist einerseits die Vergrößerung der Oberfläche und andererseits die Erhöhung der Rauheit, beide sollen sich positiv auf das Einwachsverhalten auswirken. Die Gewichtsoptimierung mittels FEM-Simulation (insbesondere bei Titanimplantaten) ermöglicht es, das Implantatgewicht an das Gewicht der humanen Schädeldecke anzupassen. Die während des gesamten Entwicklungsprozesses dokumentierten FuE-Ergebnisse inklusive Bewertung münden in einer gemeinschaftlichen wissenschaftlichen Veröffentlichung zusammen mit den Projektpartnern.

Danksagung

Ein besonderer Dank gilt dem Fördermittelgeber ZIM für die finanzielle Realisierung des Projekts.

Literaturverzeichnis

- [1] Archavlis E, Carvi Y Nievas M: The impact of timing of cranioplasty in patients with large cranial defects after decompressive hemicraniectomy. Acta Neurochir (Wien) 154:1055–1062, 2012

- [2] https://www.researchgate.net/profile/Chi_Hieu_Le3/publication/235250511_Medical_rapid_prototyping_applications_and_methods_Assembly_Automation_254_284-292/links/564b600e08ae4ae893b7c1ac/Medical-rapid-prototyping-applications-and-methods-Assembly-Automation-254-284-292.pdf
- [3] <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3159354/>
- [4] <https://www.liebertpub.com/doi/full/10.1089/neu.2011.1794>
- [5] https://www.lazoi.com/DiagnosticLab/ViewDiseaseArticle?D_Id=20180106003&DH=Cranioplasty

KI und Maschinelles Lernen

Sorting of Single-Molecule Trajectories by means of Machine Learning- a status update on the annotation procedure

Lisa Krenkel¹, Tobias Schlosser², Danny Kowerko², Richard Börner¹

¹ Laserinstitut Hochschule Mittweida, University of Applied Sciences Mittweida, 09648 Mittweida, Germany, ² Department of Informatics, Technical University Chemnitz, 09111 Chemnitz, Germany

We use machine learning for the selection and classification of single-molecule trajectories to replace commonly used user-dependent sorting algorithms. Measured fluorescence time series of labelled single molecules need to be sorted into 'good molecules' and 'bad' molecules before further kinetic and thermodynamic analysis. Currently, processing, sorting and analysis of the data is mainly done with the help of laboratory specific programs. Although there are freely available programs for processing smFRET data, they do not offer 'molecular sorting' or it is purely empirical. Only recently, new approaches came up to solve this problem by means of machine learning. Here, we describe a sound terminology for molecular sorting of smFRET data and present an efficient workflow for manual annotation followed by the training of the ML algorithm. Descriptive statistics of our generated dataset are provided and will serve as the basis for supervised ML-based molecular sorting algorithms yet to be developed.

Introduction

In recent years single-molecule fluorescence resonance energy transfer (smFRET) has been established as a mature and adaptable method regarding the study of biomolecular structures as well as their structural dynamics¹. Förster resonance energy transfer describes the weak dipol-dipol coupling of a donor and an acceptor fluorophore with overlapping emission/excitation spectra. The efficiency of this energy transfer is strongly distance dependent and used to calculate the inter-dye distance of a specific donor-acceptor-pair. The experimental measure FRET is calculated according to:

$$FRET = \frac{F_A}{F_A + F_D} \quad (1)$$

where F_A refers to the fluorescence emission intensity of the acceptor and F_D to the emission of the donor fluorophore.

In total internal reflection fluorescence (TIRF) microscopy experiments, the (bio-)molecules of interest are labelled with fluorophores and immobilised on a surface before their fluorescence signal is recorded via EMCCD or sCMOS camera resulting in a single-molecule video (SMV) [6]. The analysis of SMV's usually requires two steps: starting with video processing, including the separation of the field of view (FOV) into donor and acceptor channel and the detection of single-molecule spots; consisting of the calculation of biomolecule coordinates and their fluorescence intensities [2]. The subsequent second step is the trace processing including fluorescence intensity correction and reliable state detection [7]. However, the fluorescence trajectories which are extracted

from the SMV contain not only useful fluorescence intensity fluctuations e.g., from structural changes in the biomolecule, but photophysical artefacts, such as photobleaching, blinking etc. Because of this, the intensity trajectories are sorted into 'good' and 'bad' molecules, i.e., fluorescence intensity trajectories fulfilling a number of criteria, before further analysis. This molecular sorting (or classification) is currently done mostly by hand or through semi-automated algorithms based on threshold criteria still requiring profound user interference. Admittedly, this is not only problematic in terms of the required expertise and time, but also inevitably introduces user bias. Most of the freely available tools for sorting of fluorescence trajectories rely on purely heuristically approaches like intensity thresholds etc. which requires a rather advanced understanding of FRET [8]. New tools have emerged implementing machine learning approaches to classify smFRET trajectories such as FRET-board or deepFRET while minimising user bias and necessary time [4, 9].

At its simplest, machine learning describes the design of algorithms with the ability to optimize their performance based on example data or prior experience [10]. The chosen approach in case of the sorting algorithm is based on supervised learning, meaning that the algorithm is trained with annotated training data to later sort similar, un-annotated data [11]. Since the annotation is done per hand, based on user experience and/or prior knowledge, the intensity-time traces of a biomolecule are generally depicted in form of graphs containing both the donor and acceptor intensities as well as the FRET

¹ Also known as Förster resonance energy transfer. The effect was discovered and described first theoretically by Theodor Förster in 1948[5].

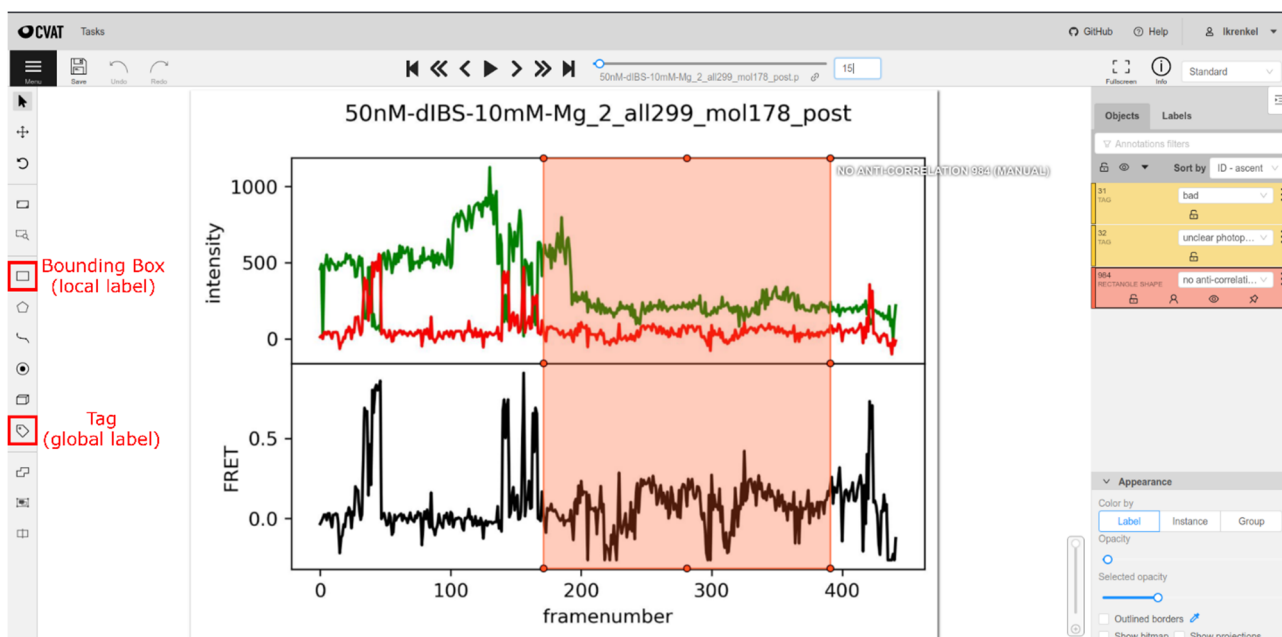


Fig. 1: CVAT Annotation tool: Both, the global and local labels, were assigned via their respective tabs in the CVAT annotation tool as marked on the (left). On the (right) all of the assigned labels are displayed. The annotation progress can be viewed at the top.

trajectory (Fig. 1). Thus, the annotated data is used as training data set.

The field of automated and semi-automated visual inspection utilises classification approaches which can be distinguished according to their functionality: filter-based, projection-based, and hybrid approaches. These are usually accompanied by clustering approaches, mostly as additional classification steps, but also to generate interpretable rule sets for the classification [17]. The problem with these approaches lies within their limited ability to classify novel signatures and patterns as well as the likely need for manual adaptation and parameterisation through the user. Because of this more recent machine learning approaches to the problem of molecular sorting have included support vector and multilayer perceptron-based classifiers. Following the contributions of *de Lannoy et al.* [3], *Thomsen et al.* [4], and *Li et al.* [9] even the latest deep learning motivated research utilized classical single deep and convolutional neural networks. These approaches are assumed to result in improved training and test rates, which results in the definition of an optimum which allows for both, a timely classification as well as a solution independent classification of regions of interest.

Following the beforehand highlighted and emphasised principles of automated and semi-automated visual inspection, this contribution proposes an evaluation within a range different approaches in comparison to contrast and accentuate their respective advantages and differences. For this purpose, said approaches are separated into the following three categories: baseline ML models, DL-based models, as well as more novel,

temporal information and region-based convolutional (R-CNN) and deep neural networks in general. Furthermore, this also includes commonly deployed DL-based principles such as pre-trained and optimised building blocks for DNNs (e.g., DenseNets) [18], building blocks with different characteristics, i.e. kernel sizes and connectivities (inception modules) [19], approaches with residual learning (ResNets) [20], as well as combinations of these approaches, i.e. Inception-ResNets [21].

In the last decade, numerous annotation tools for images and videos have emerged in the computer vision community. We use CVAT² (Computer Vision Annotation Tool), which supports many easily machine readable annotation formats, developed or originally used in computer vision challenges [12]. As such CVAT fulfils the annotation requirements for this project which are mainly bounding box and global labels, the rich text-based export format support (json, csv), collaborative labelling via web (application) and preserving data privacy via hosting it as own web service/web application. Considering the chosen approach of supervised learning, the first step in developing a neural network for molecular sorting is the annotation of fluorescence trajectories to build a data set for training, validation, and testing.

Herein, we give an overview of the different label categories used, the percentage of the labels within the test data set, as well as a short outlook how this work will continue.

Methods

The video processing and trace generation to obtain fluorescence intensity and thus FRET trajectories for a

² <https://github.com/opencv/opencv/cvat>

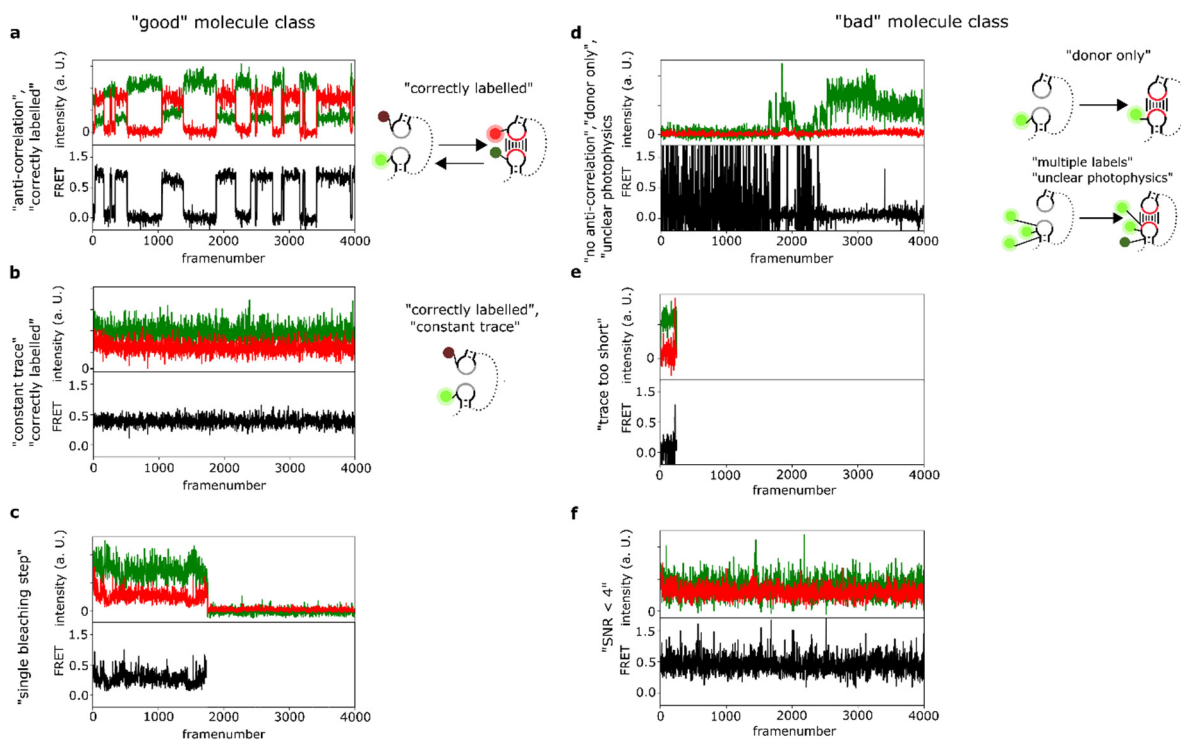


Fig. 2: Label classifiers: (left column) Good molecules (right column) Bad molecules. (a) Correctly labelled, anti-correlated signal of single-molecule fluorescence intensity trajectory. (b) Correctly labelled, static single-molecule intensity trajectory. (c) Correctly labelled, anti-correlated signal of a single-molecule intensity trajectory - both donor and acceptor signal display single bleaching steps. (d) Incorrectly labelled (multiple donor labels) signal for a single-molecule intensity signal with no anti-correlation but unclear photo physics. (e) Too short trace of a single-molecule intensity trajectory. (f) Single-molecule intensity trajectory with a too low signal to noise ratio, resulting in a very 'noisy' signal.

given experimental or simulated SMV is described elsewhere [13]. An annotation for the training dataset was achieved by converting the fluorescence intensity and FRET traces into graphs, which were labelled using the online image/video annotation tool CVAT (Fig. 1). An important distinction was made between local and global labels. Global labels are defined by tags, e.g. bad or good, and were selected based on the combination of assigned local labels (compare Tab. 1). The local labels in turn are assigned using bounding boxes to mark their occurrence in the trace, i.e. by assigning a certain range of frame numbers, as such all traces require a minimum number of two labels - one global and one local. Examples for "good" and "bad" molecules are given in Fig. 2.

The declaration as a "good" or "bad" molecule was made after the local labels were assigned. An important factor for the declaration of a "good" or "bad" molecule in smFRET experiments is the anti-correlation of donor and acceptor signals, inherently linked to the FRET process, where a high donor signal corresponds to a low acceptor signal and *vice versa*. Thus, "anti-correlated" signal traces are labelled accordingly as "good" molecules (Fig. 2a). If the intensity trace of a given molecule does not display anticorrelation, but a constant signal well above the background, the trace is labelled as constant or static (Fig. 2b). This is the case if no structural changes take place during the measurement time. In some instances, a fluorescence signal can be observed to de-

crease instantaneously to a minimal level, i.e. the background signal. This is referred to as single-bleaching step, an event in which a fluorescent dye is (photo)chemically destroyed. A single bleaching step is usually not considered to be an indicator for a bad molecule, rather these traces can be used to calculate correction factors (Fig. 2c) [14]. In contrast, multiple bleaching steps confirm multiple dyes in the spot (Fig. 2d) and are labelled as "bad" molecules. Usually, it is not possible to determine whether these multiple bleaching steps are a result of the labelling method itself, the immobilization procedure, or an experimental artefact. As such, they were annotated as "unclear photophysics" (Fig. 2d). If the fluorescence signal of a donor-dye remains at the background signal level for the whole trace, the molecule was labelled as "donor only" (Fig. 2d) and consequently as "bad" molecule. In contrast, a signal for both, donor and acceptor well above the background signal, is an indicator for a correctly labelled molecule, and designated as "good" (Fig. 2a, b, and c). Further, smFRET traces vary in their length due to the statistical nature of photobleaching or potential molecular desorption. Heuristically, traces are categorized as "trace too short" to give any significant information (< 10 frames, Fig. 2e). Lastly, some smFRET traces contain important information regarding the different states of the biomolecule but are obscured by experimental noise. Here, observed fluctuations have near to no correlation to actual state changes of the molecules - the noise simply masks the fluctuations due to state changes. Because of this, traces with a low signal

to noise ratio SNR < 4 can generally be considered as "bad" molecules (Fig 2f) [7].

Results

The training data set, at the current stage, contains 2409 molecules, of which 74.89 % are designated as "good" and 26.24 % as "bad". The percentages of all defined labels are listed in Tab. 1. It is important to note that some labels like "anti-correlation" were used multiple times on a single trace but count only once. Further, traces which exhibit multiple features were labelled multiple times (compare Tab. 1, right column).

Tab. 1: Label distributions across training data set

| label | fraction of molecules | labels per trace | fraction of molecules |
|-----------------------|-----------------------|------------------|-----------------------|
| good | 74.89 % | 1 | 0.00 % |
| bad | 26.24 % | 2 | 11.50 % |
| single bleaching step | 2.95 % | 3 | 84.85 % |
| SNR < 4 | 13.08 % | 4 | 3.28 % |
| donor only | 7.85 % | 5 | 0.30 % |
| trace too short | 0.00 % | | |
| correctly labelled | 71.19 % | | |
| no anti-correlation | 8.09 % | | |
| anti-correlation | 78.41 % | | |
| unclear photophysics | 12.41 % | | |

Discussion and Outlook

A known issue of training data sets for ML approaches is the existence of class imbalance. Here, we observe a higher number of "good" in comparison to "bad" molecules. If the distribution of classes, i.e. labels, is highly imbalanced, classification learning algorithms usually display low predictive accuracy for the infrequent classes [15]. The effect on the classification performance of the resulting neural network is generally detrimental and solved e.g., by under/oversampling which might lead to overfitting of the algorithm [16]. As raw smFRET data sets are rare and if published, do not contain "bad" molecules, raw SMV data needs to be processed to minimise the class imbalance in the training data set. Therefore, we will annotate more smFRET traces especially for the classification of bad molecules.

MASH-FRET as a tool for the processing of smFRET data also includes the option to simulate traces and even the option to simulate whole SMVs - including "bad" molecules [2]. The annotation of simulated data will help to overcome class imbalance in our training data set. Further, simulated data will be of particular interest during the evaluation process of our algorithm - since all relevant parameters are defined in this simulated data sets they are considered as ground-truth [7].

Only recently, deepFRET [4] and AutoSiM [9] introduced AI-assisted automated sorting/classification of smFRET trajectories. But despite the very positive results of both studies, they show a lack of cross-sample and cross-laboratory variability. As such not only data sets have to be continuously updated but algorithms need to be trained with community driven annotated data sets. Therefore, we will pay special attention to use single-molecule data sets of different labs to introduce lab specific variances in the training data set.

Acknowledgements

The authors thank the organizers of the *Interdisziplinäre Wissenschaftliche Konferenz an der Hochschule Mittweida* (IWKM) for giving them the opportunity to present their research. We thank Susann Zelger-Paulus for sharing SMV. We further thank Victoria Birkedal and Thomas Villmann for insightful discussions; we are looking forward to further collaborate with them.

References

- [1] Juetten, M. et al.: Nature Methods, 13 (2016), 341–344
- [2] Börner, R. et al.: PLOS ONE, 13 (2018), e0195277
- [3] de Lannoy, C. et al.: bioRxiv, 15 (2020), 2020.08.28.272195
- [4] Thomsen, J. et al.: eLife, 9 (2020), e60404
- [5] Lerner, E. et al.: eLife, just accepted (2021)
- [6] Hellenkamp, B. et al.: Nature Methods, 15 (2018), 669–676
- [7] Hadzic, M. C. A. S. et al.: The Journal of Physical Chemistry B, 122 (2018), 6134–6147
- [8] White, D. S. et al.: eLife, 9 (2020), e53357
- [9] Li, J. et al.: Nature Communications, 11 (2020), 5833
- [10] Alpaydin, E.: MIT Press, 1 (2010)
- [11] Bonaccorso, B.: Packt Publishing, 1 (2017)
- [12] Liu, L. et al.: International Journal of Computer Vision, 128 (2020), 261–318
- [13] Hadzic, M. C. A. S. et al.: International Society for Optics and Photonics, 9711 (2016), 971119
- [14] Zelger-Paulus, S. et al.: Methods in Molecular Biology (Clifton, N.J.), 2113 (2020), 1–16
- [15] Ling, C. X. et al.: Springer US, 1 (2010)
- [16] Buda, M. et al.: Neural Networks, 106 (2018), 249–259
- [17] Lakkaraju, H. et al.: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2016, 1675–1684
- [18] Huang, G. et al.: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, 2261–2269
- [19] Szegedy, C. et al.: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015, 1–9
- [20] He, K. et al.: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, 770–778
- [21] Szegedy, C. et al.: AAAI Conference on Artificial Intelligence (AAAI-17), 2016, 4278–4284

Analyse von Thorax-Aufnahmen zur Erkennung von Virusinfektionen

Lena Kopp, Roman Radtke, Alexander Jesser
Hochschule Heilbronn, Daimlerstraße 22, 74653 Künzelsau

Ende Dezember 2019 infizierte sich in der chinesischen Stadt Wuhan vermutlich zum ersten Mal ein Mensch mit dem SARS-CoV-2 Virus. Um möglichst schnell nach einer Infektion reagieren zu können sind schnelle Diagnosemöglichkeiten von größter Bedeutung, um frühzeitig medizinische Maßnahmen ergreifen zu können. Zu diesem Zweck wird ein Bildverarbeitungsalgorithmus vorgestellt, der anhand von Röntgen-Thorax-Aufnahmen ermitteln kann, ob eine Virusinfektion oder eine bakterielle Infektion der Lunge vorliegt. Dieser soll vor allem zu einer schnelleren Diagnose von SARS-CoV-2 Virus-Infektionen in medizinisch unterversorgten Gebieten beitragen.

At the end of December 2019, a person in the Chinese city Wuhan was probably infected for the first time with the novel SARS-CoV-2 virus. In order to be able to react as quickly as possible after infection, rapid diagnostic measures are of the utmost importance so that the medical measures can be taken at an early stage. An image processing algorithm is presented using chest X-rays to determine whether a lung infection has a viral or a bacterial cause. This should contribute to a faster diagnosis of SARS-CoV-2 virus-infection, especially in medically underserved areas.

1. Einleitung

Seit Ende 2019 breitet sich das neuartige Virus SARS-CoV-2 immer weiter aus und hat mittlerweile zu einer weltweiten Pandemie geführt. Das Krankheitsbild der sogenannten COVID-19-Erkrankung ist unspezifisch und ähnelt stark dem einer Grippe oder Erkältung. Neben symptomlosen und leichten Krankheitsverläufen gibt es schwere Verläufe mit Lungenentzündungen bis hin zum Lungenversagen. Um möglichst prompt nach einer Infektion reagieren zu können sind schnelle Diagnosemöglichkeiten von größter Bedeutung, um sofort medizinische Maßnahmen ergreifen zu können. Vor allem in Entwicklungsländern sind einfache Hygienemaßnahmen nur schwer umzusetzen und moderne Technik für eine schnelle Diagnose kaum zugänglich. Deshalb ist die Entwicklung von einfachen und kostengünstigen Diagnosetools vor allem für diese Länder von großer Bedeutung.

2. Diagnosemöglichkeiten

Trotz der weltweiten Forschungsanstrengungen existieren noch sehr große Wissenslücken in Bezug auf die vom SARS-CoV-2 ausgelöste Krankheit COVID-19. Die Krankheitsverläufe sind meist unspezifisch und ähneln einer Grippe oder Erkältung. Laut dem Robert-Koch-Institut (RKI) [1] verlaufen 80% der Erkrankungen mild bis moderat, das heißt die Betroffenen zeigen lediglich grippeähnliche Symptome wie Fieber, Halsschmerzen und Husten. In 20% der Fälle kann es zu Atemnot infolge einer Lungenentzündung (Pneumonie) kommen, die sich typischerweise 7-10 Tage nach einer Infektion bemerkbar macht. Etwa 5% der an COVID-19 erkrankten Patienten müssen intensivmedizinisch behandelt werden, da die Gefahr eines Lungenversagens besteht (Stand Oktober 2020). Um eine COVID-19 Erkrankung schnellstmöglich

zu Diagnostizieren und die Zahl der weiteren Ansteckungen so gering wie möglich zu halten, bieten sich neben molekularbiologischen Verfahren auch bildgebende Verfahren an, bei denen die Ergebnisse direkt vorliegen und anhand derer die Diagnose einer COVID-19 Erkrankung beschleunigt werden kann. Laut dem RKI werden bei 50-60% der Erkrankten Veränderungen bei einer konventionellen Röntgenaufnahme des Thorax sichtbar. Bei einer CT-Untersuchung der Lunge werden bei ca. 80% der Erkrankten Veränderungen, die sich durch milchglasartige Infiltrate oder lokalen Verdichtungen bemerkbar machen, detektiert.[1]

Da die radiologischen Befunde einer COVID-19 Erkrankung sich mit denen von Patienten mit SARS oder MERS ähneln, erfordert eine eindeutige Diagnose eine Bestätigung durch einen positiven Nachweis von Virusgenen anhand eines mikrobiologischen Tests. In Entwicklungsländern und medizinisch unterversorgten Regionen stehen moderne Labore für mikrobiologische Tests meist nicht mit ausreichender Kapazität zur Verfügung. Da radiologischen Befunde jedoch direkt vorliegen, kann eine Verdachtsdiagnose anhand bildgebender Diagnoseverfahren gestellt werden. Um möglichst schnell nach einer Infektion reagieren zu können sind schnelle Diagnosemöglichkeiten von größter Bedeutung, um anhand dieser frühzeitig medizinische Maßnahmen ergreifen zu können und Kontakte zu reduzieren.

3. Analyse von Thorax-Aufnahmen

Konventionelle Radiographie ist meist auch in medizinisch unterversorgten Regionen zugänglich. Darüber hinaus sind schnelle Diagnoseverfahren hier oft besonders wichtig, da viele Menschen auf engstem Raum leben. Zu diesem Zweck wurde ein Tool entwickelt, das anhand einer Thorax-Röntgenaufnahme mit möglichst großer Wahrscheinlichkeit autonom bestimmt, ob es sich

um eine bakterielle oder virale Erkrankung des Lungengewebes handelt. Das Tool basiert auf klassischen Bildverarbeitungsalgorithmen und unterscheidet anhand der befallenen Lungenfläche ob eher eine bakterielle Infektion oder eine Virusinfektion vorliegt. Da typischerweise am Anfang einer Pandemie nur wenige Daten vorliegen, wurde bewusst auf den Einsatz von künstlicher Intelligenz verzichtet und ausschließlich klassische Bildverarbeitungsalgorithmen angewandt.

Bei der Aufnahme von Röntgenbildern wird der Körper mithilfe von Röntgenstrahlen durchstrahlt. Ein Röntgenfilm oder ein digitaler Detektor fängt die aus dem Körper austretenden Strahlen auf und erzeugt ein Röntgenbild, welches die unterschiedliche Absorption der verschiedenen Gewebe widerspiegelt. Dabei erscheint dichtes Gewebe wie Knochen auf dem Röntgenbild heller, da die Strahlung durch die hohe Dichte stärker abgeschwächt wird. Gewebe mit einer geringeren Dichte, wie zum Beispiel Muskeln oder das ventilierte Gewebe in der Lunge, haben einen kleineren Absorptionskoeffizienten, weshalb diese Regionen eher dunkel auf dem Röntgenbild erscheinen. Bei der Röntgenaufnahme einer gesunden Lunge sind die anderen Organe klar abgegrenzt und die Lunge selbst ist gut zu erkennen. Bei einer Erkrankung der Lunge sind die inneren Organe auf der Aufnahme weniger gut voneinander abgegrenzt. Das erkrankte Gewebe besitzt meist eine höhere Dichte als das gesunde Gewebe und wird deshalb auf einer Röntgenaufnahme heller als das umliegende gesunde Gewebe abgebildet. Form und Abgrenzung des erkrankten Gewebes kann Aufschluss ergeben, um welche Art der Erkrankung es sich handelt. Bei einer COVID-19 Erkrankung sind auf den Röntgenaufnahmen meist milchglasartige Infiltrate in beiden Lungenflügeln erkenntlich, wobei sich diese über mehrere Lungenabschnitte erstrecken. In Abbildung 1 sind Röntgenaufnahmen einer gesunden Lunge im Gegensatz zu einer mit COVID-19 erkrankten Lunge dargestellt.

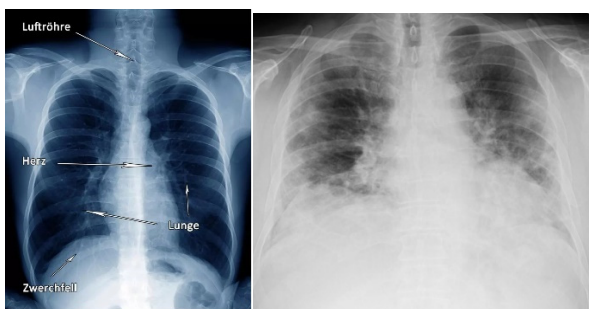


Abbildung 1: Röntgenaufnahme einer gesunden Lunge (links) im Vergleich zu einer erkrankten Lunge (rechts) [2]

In der rechten Abbildung ist erkennbar, dass beinahe die gesamte Lunge einen vergleichsweise hohen Absorptionskoeffizienten besitzt (helle Bereiche), was auf krankhaft verändertes Gewebe hindeutet. Um eine Unterscheidung zwischen bakteriellen Erkrankungen und Virusinfektionen anstellen zu können ist es wichtig die Thorax-Aufnahmen auf die spezifischen Merkmale der

Erkrankungen hin zu untersuchen. Bakterielle Infektionen zeichnen sich meist dadurch aus, dass nur ein Lungenflügel von der Infektion betroffen ist und sich das befallene Gewebe deutlich abgrenzt. Bei Virusinfektionen sind meist beide Lungenflügel betroffen und die pathologisch auffälligen Bereiche sind räumlich weniger begrenzt.

3.1. Algorithmus zu Analyse der Thorax-Aufnahmen

Der entwickelte Algorithmus basiert auf klassischen Bildverarbeitungsalgorithmen und unterscheidet anhand des Verhältnisses der pathologisch hellen Bereiche der Lungenfläche, sowie der Länge der Lungenflügel, ob eher eine bakterielle oder virale Infektion vorliegt. Die Röntgenaufnahme des Thorax wird in den Algorithmus eingelesen und mithilfe folgender Schritte analysiert:

- Einlesen und formatieren der Bilddatei
- Umwandlung des Grauwertbild in einen Farbraum
- Kantendetektion zur Segmentierung der Lunge
- Separate Analyse der Lungenflügel
- Vergleich der Ergebnisse beider Lungenflügel
- Auswertung der Ergebnisse

Im ersten Schritt wird die Röntgenaufnahme eingelesen und auf einheitliche Größe formatiert, um die Bilder vergleichen zu können. Anschließend wird der Kontrast anhand einer Histogramm Auswertung des Intensitätsbereichs des Bildes optimiert. Da es für das menschliche Auge sehr schwer ist in einem Grauwertbild die Unterschiede zwischen den einzelnen Graustufen zu erkennen, wird das Bild zusätzlich in einen Farbraum überführt. Der linke Teil von Abbildung 2 stellt eine derart eingefärbte Thorax-Aufnahme eines COVID-19 Patienten dar; das normal ventilierte, gesunde Gewebe ist hier dunkelblau dargestellt

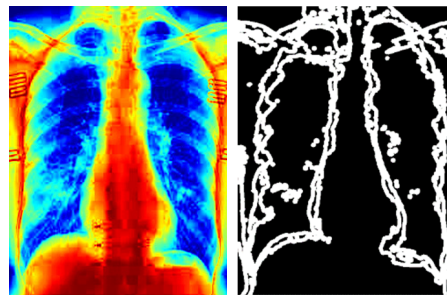


Abbildung 2: Einfärben und Segmentierung der Lunge

Die infizierten Teile des Gewebes sind hier hellblau dargestellt. Die Lungenflügel können anschließend mithilfe einer Kantendetektion segmentiert werden. Nach erfolgreicher Transformation liegt ein Binärbild des Lungeninhaltes vor. Das mutmaßlich gesunde Gewebe der Lunge wird darin weiß dargestellt. Die schwarzen Bereiche, die von gesunden weißen Bereichen umgeben sind, können als infiziertes Gewebe betrachtet werden.

Anschließend wird der Flächeninhalt jedes Lungenflügels berechnet und zusätzlich dessen Länge ermittelt. Der Lungeninhalt und die Länge der Lungenflügel nach

der Segmentierung lassen Rückschlüsse auf die Art der Erkrankung zu. Ist sowohl der Flächeninhalt und die Länge beider Lungenflügel vergleichbar groß, so legt die Vermutung nahe, dass es sich hierbei eher um eine Virusinfektion handelt. Sind die Ergebnisse beider Lungenflügel jedoch stark unterschiedlich, so lässt sich eine bakterielle Erkrankung der Lunge vermuten. Die Grenzwerte zur Unterscheidung zwischen bakteriellen Infektionen und Virusinfektionen wurden empirisch anhand mehrerer Fallstudien ermittelt. Ist dabei ein Lungenflügel um über 30% länger als der andere Lungenflügel lässt sich eine bakterielle Erkrankung vermuten. Haben die Lungenflügel jedoch eine vergleichbare Länge und der ermittelte Lungeninhalt stimmt zu 90% überein kann eine Virusinfektion vermutet werden. In Abbildung 3 sind die Ergebnisse einer Virusinfektion (rechts) denen einer bakteriellen Infektion (links) gegenübergestellt.



Abbildung 3: Ergebnis nach Segmentierung der Aufnahmen einer bakteriellen Infektion und einer Virusinfektion

An der Aufnahme eines an COVID-19 erkrankten Patienten (rechts) sind die infizierten Bereiche innerhalb der Lunge zu erkennen: beide Lungenflügel weisen die durch die typische infektionsbedingte Verteilung der Herde vergleichbare Verschattungen auf. Durch die vertikale Teilung des Bildes, der separaten Berechnung des Flächeninhaltes und der Länge der Kontur, können die Verhältnisse zwischen den beiden Lungenflügeln ermittelt werden. Anhand dieser Verhältnisse kann anschließend auf die Symmetrie des Befalls der Lungenflügel geschossen werden. Ist das Verhältnis zwischen den beschatteten Bereichen beider Lungenhälften größer als der empirisch ermittelte Schwellwert, so kann von einer bakteriellen Erkrankung ausgegangen werden, da diese meist nur einen der beiden Lungenflügel erfasst. Ist das Verhältnis in beiden Lungenflügel nahe eins, so kann davon ausgegangen werden, dass beide Lungenflügel in vergleichbarem Umfang befallen sind. Dies legt die Vermutung nahe, dass es sich um eine Viruserkrankung handelt.

Der Algorithmus erkennt ebenso die bakterielle Infektion durch das verringerte Lungenvolumen im rechten Lungenflügel in der linken Aufnahme. Die Auswertung der analysierten Bilder basiert also auf den Verhältnissen der vermutlich anhand der Abschattung als infiziert zu bewertenden Bereichen beider Lungenflügel zueinander, um so eine Verdachtsdiagnose erstellen zu können.

4. Ergebnisse

Der hier entwickelte Algorithmus liefert für viele der vorliegenden Röntgenbilder gute Ergebnisse, ist jedoch derzeit nicht robust genug, um bei jedem Röntgenbild befriedigende Ergebnisse zu liefern. Dies ist zum Teil der Tatsache geschuldet, dass die Röntgenbilder unter unterschiedlichen Bedingungen aufgenommen wurden und keine Informationen über die genauen Settings bei der Erstellung vorhanden sind. Außerdem ist die Größe der Lunge, sowie das exakte Größenverhältnis der Lungenflügel aufgrund anatomischer Individualitäten nicht immer gleich, was einen Vergleich des Verhältnisses der Flächeninhalte der Lungenflügel erschwert. Zudem liegen die Röntgenaufnahmen in komprimierten Bilddateien vor, wodurch Bildinformationen verloren gehen. Da die Größe der Bilder unterschiedlich war, mussten diese, um einen Vergleich zwischen den Bildern anstellen zu können, auf ein einheitliches Bildformat skaliert werden. Bei diesem Vorgang muss entweder für eine Bildvergrößerung zwischen den Pixeln interpoliert werden, oder es gehen aufgrund der Größenreduzierung Bildinformationen verloren. In einem Versuch mit 15 zufällig ausgewählten Thorax-Aufnahmen unterschiedlicher Qualität wurden diese zu 47% korrekt ausgewertet. Bei den übrigen Aufnahmen konnte durch den meist zu geringen Kontrast zwischen Lungengewebe und Knochen die Kanten nicht vollständig detektiert werden.

Um die Stabilität des Algorithmus zu verbessern sind weitere Tests mit Thorax-Aufnahmen unterschiedlichster Krankheitsbilder durchzuführen. Nichtsdestotrotz konnte mit der diskutierten Methode bei vielen Aufnahmen der Inhalt der Lungenflügel separiert und eine Vermutung über den Ursprung der Erkrankung getroffen werden. Wie bei allen bildgebenden Verfahren ist zu bedenken, dass keine finale COVID-19 Diagnose erstellt werden kann. Die Methode ist vielmehr, gerade in Ländern bei denen viele Menschen keinen Zugang zu fachkundiger medizinischer Versorgung haben, ein adäquates Mittel um schneller Verdachtsdiagnosen erstellen zu können.

5. Danksagung

Abschließend möchten wir uns bei Dr. Konrad Appelt für die Unterstützung bei medizinischen Fragen bedanken.

Literaturverzeichnis

- [1] Ständiger Arbeitskreis der Kompetenz- und Behandlungszentren für Krankheiten durch hochpathogene Erreger am Robert Koch-Institut. - Hinweise zu Erkennung, Diagnostik und Therapie von Patienten mit COVID-19.
- [2] Susanne Greiner. COVID-19-Patient: Ein Arzt aus dem Klinikum Landsberg berichtet. (<https://www.kreisbote.de/lokales/landsberg/covid-19-patient-arzt-klinikum-landsberg-berichtet-13713274.html>)

Intelligent Gait Analysis using Marker Based Motion Capturing System

Danny Möbius^a, Marika Kaden^b, Daniel Staps^b, and Thomas Villmann^b

^a Institut für Mechatronik, Chemnitz, Germany

^b SICIM, Hochschule Mittweida, Mittweida, Germany

Marker-based systems can digitally record human movements in detail. Using the digital biomechanical human model Dynamicus, which was developed by the Institut für Mechatronik, it is possible to model joint angles and their velocities such accurately that it can be used to improve motion analysis in competitive sports or for ergonomic evaluation of motion sequences. In this paper, we use interpretable machine learning techniques to analyze the gait. Here, the focus is on the classification between foot touchdown and drop-off during normal walking. The motion data for training the model is labeled using force plates. We analyze how we could apply our machine learning models directly on new motion data recorded in a different scenario compared to the initial training, more precise on a treadmill. We use the properties of the interpretable model to detect drift and to transfer our model if necessary.

1. Introduction

In the *Institut für Mechatronik* (IfM) a digital biomechanical human model names *Alaska/Dynamicus* was developed [1,2]. It uses the data of a marker-based system, which record the human movements. By means of the digital model it is possible to obtain the joint angles and their velocities very precise. Marker-based systems have a broad application not only on competition sports, also in ergonomic evaluation and movie industry. Even marker-less systems get better and better in the recent years, especially under the use of machine learning techniques, it cannot achieve such precise results in predicting the joint angles, yet [3]. The model *Dynamicus* is successfully applied in several areas like detection/simulation of the movement of the subject and the environment when entering a passenger car or truck, automated evaluation of work processes with the EAWS method or the acquisition/simulation of a jump on the force plate.

In our current gait study, the goal is to precisely detect the time when a foot touchdown and drop-off during normal walking using only the data of the motion capturing system. In a first step we divide the normal walking into two phases: the swinging and the standing phase of each foot. Using artificial intelligence, more precise machine learning (ML) we could train a model to predict these phases for the left and right foot very precise. The training of the model is taken place with labeled data, i.e., beside the different joint angles of the body provided by *Dynamicus* the time of the phases is given. These labels were obtained by using additionally force plates. This measuring system provides the exact timing for the phases. As we can see later, the model show good practical performance.

Yet, the force plates cannot easily apply in every scenario to train a scenario specific model, e. g. during walking on a treadmill. The idea is to train a ML model on the

data, where the labels are given, i.e., in the scenario, where the force plates can be used, and then apply this model to other scenarios. But as we can imaging, standard walking on the floor and walking on the treadmill have different influences on the movement. The questions are: Can we apply our model to other scenario or is a transformation necessary? If the model should be transferred, how we could adapt our model? For easier communication we will name the data recorded on the floor with available labels *train data* and the data of the treadmill without knowledge of the ground truth label *test data*.

Detection of data-drift is still an ongoing research topic [5]. We want to use prototype-based machine learning models to tackle the problem. These models have several advantages: Due to their competitive strategy they are quite intuitive and interpretable [7]. During minimizing the corresponding cost function the margin is maximized, which results in a robust decision making [margin]. Further the complexity of the model can be directly chosen by the applicant, i.e., the complexity of the model is not given, but can be adapted to boundary conditions. Another advantage is that we can use our model to detect drift.

2. Prototype Based Models

We do not want to overload the reader with formulas; thus, we describe the ideas only in rough illustrating manner. In the Generalized Learning Vector Quantization model (GLVQ, [11]) we have, beside the data with respective class assignments called labels, prototypes also equipped with labels. During learning the prototypes are adapted following geometrically an attraction and repulsing scheme, i.e., if the label of a training data sample and the corresponding nearest prototype agrees the prototype is attracted by the data point and it is pushed away otherwise. This very intuitive learning scheme is the result of applying stochastic gradient

descent on the GLVQ-cost function. This cost function approximates the classification error to be optimized during training and, additionally, describes the (hypothesis) margin margin . In the recall phase, a new data point gets the label of the nearest prototype, also known as Winner-Takes-All rule.

The dissimilarity measure used in the competition is in general the squared Euclidean distance, but any other measure can be applied like kernel distances or divergences [9,10] under mild mathematical conditions. A powerful extension is the Generalized Matrix LVQ (GMLVQ, [12]). Here, the distance is replaced by a parameterized version:

$$d_{\Omega}(\mathbf{x}, \mathbf{w}) = (\Omega(\mathbf{x} - \mathbf{w}))^2$$

where $\mathbf{x} \in \mathbb{R}^n$ is a given data sample and $\mathbf{w} \in \mathbb{R}^n$ represents a prototype and $\Omega \in \mathbb{R}^{m \times n}$ is called the mapping matrix, which maps the data from an n -dimensional to a m -dimensional vector space. This means that the above matrix Ω has the goal to map the data in such a way that they become better separable. For m we obtain a mapping into the two-dimensional visualization space while maximizing the classification performance.

In the specific gait analysis task to be investigated here, pure classification accuracy tells nothing about the applicability of the model. Yet, *IfM* provides a tool, which shows the single phases of the left and right foot. An illustration of the output of the learned GLVQ model is depicted in Figure 4. By means of this tool it is possible to evaluate the classification result visually. The obtained results are promising and show that it is possible to obtain a model to detect these phases only using the joint angles provided by *Dynamicus*.

Another possibility to evaluate the performance is the use of the time-difference between ground truth and the prediction for detecting the starting point of each phase and the detection of ending a phase. The mean time-difference for starting is 0.2 seconds with a standardization of ± 0.23 and for ending 0.31 ± 0.42 seconds. Thus, our ML-model works in praxis.

3. Drift Detection

Another challenge in our scenario is that we only have labeled data for the normal working mode on the floor. We do not have any label information for the data obtained from the treadmill due to technical difficulties. One possibility to overcome the lack of information is to visualize the high dimensional data using a maybe non-linear visualization method like t-distributed Stochastic Neighbor Embedding (t-SNE) [6]. In Figure 1 we illustrate the result of t-SNE on a subset of the data. On a first glance the data of the treadmill seems to be similarly distributed like the normal labeled data. But if we take a closer look, we recognize a slightly different shape of the manifold. However, we cannot decide with

certainty out of this visualization whether the treadmill data are drifted compared to normal mode.

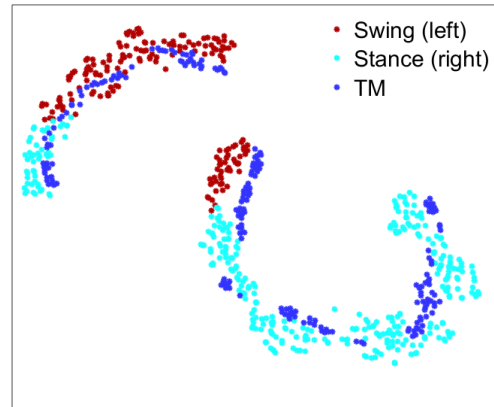


Figure 1: t-SNE visualization of the data (red - data with the label swing left, cyan- data with the label stance left and blue - data recorded on the treadmill)

Therefore, we use the GMLVQ with the interpretable recall phase to detect drift. In particular, we analyze the distribution of the distances of the data points to their nearest prototypes. The assumption for this is that if the new data are similar distributed in the vector space like the training data, the distance distributions should be similar, too. In contrast, if drift has taken place, the distribution of distance values should differ. The big advantage on looking on the distribution of the distances is that these distributions are only one-dimensional and thus easier to handle than comparing distributions in a high-dimensional data space. In Figure 2 the distance distributions are illustrated by respective histogram plots. We can detect a slightly different distribution of the distances referring to a drift. Especially for distances in the range of [10,40], i.e., data samples with larger distances to the winning prototype, the distribution is changed.

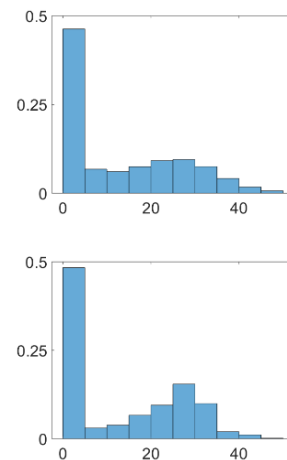


Figure 2: Histogram plots of the distances between data and closest prototype (top: data recorded by normal walking, bottom: data recorded on the treadmill)

In the 2D data visualization by means of the t-SNE the label information is not considered. Hence, if a drift in the data would be detected inspecting the visualization, the consequences of this drift for class discrimination is neither clear nor obvious. Here, the GMLVQ provides a possibility for visualization by taking the perspective of the later discrimination model. Setting the mapping dimension m in GMLVQ, we can visualize the data and optimize the model to be as best class discriminating as possible. In Figure 3 we depict the training and test data mapped using the trained Ω matrix of the GMLVQ classifier. It can be observed that the test data seems to be slightly drifted orthogonal to the decision hyperplane. The difference is not huge, but directly influences the decision. So, it might result in a higher misclassification rate for the *stance* class.

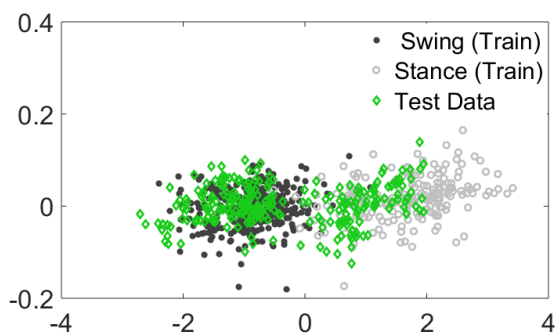


Figure 3: Visualization of the mapped data (training data - grey, test data - green)

In consequence of this observation, a direct application of the learned model on the data recorded during working on the floor to data captured in another scenario (treadmill) is not recommended. Thereby, it should be pointed out again that we do not have any labels for the second data set, such a direct evaluation of the classification results is not possible

4. Conclusion and Future Work

We apply matrix variant of the Generalized Learning Vector Quantization for a classification task in the field of motion detection using a digital bio-mechanical human model. Moreover, we figured out possibilities for drift detection in the data using the interpretability of this GMLVQ approach. To apply a model on novel data, we must check in advance whether the data have drifted, especially if the data are measured using new or slightly different positioned sensors, or the environment of the measurement process has been changed. The application of GMLVQ enables to detect those drift types, which have a high impact to the class separation. This detection is based on the utilization of the learned GMLVQ mapping matrix for class discriminating data visualization.

So far, the analyze of the GMLVQ-based visualization is done manually to detect class drift. A future work is to automatize these detections. Moreover, the mapping matrix could be also learned in a way that the model

decision is not influenced by the drift, which is related to robust adversarial learning [13, 14]. Yet, for this task data labels are required also for the new (drifted) data. Hence, we must modify these ideas accordingly, to apply it successfully for this task.

Acknowledgements

This work is co-financed by tax funds on the basis of the budget adopted by the Saxon Landtag's budget.



Diese Maßnahme wird mitfinanziert durch Steuermittel auf der Grundlage des vom Sächsischen Landtag beschlossenen Haushaltes.

References

- [1] www.ifm-chemnitz.de
- [2] H. Hermsdorf, N. Hofmann, and A. Keil: Chapter 16 - Alaska/dynamicus - human movements in interplay with the environment. Academic Press, DHM and Posturography, p. 187-198, 2019
- [3] E. Ceseracciu, Z. Sawacha, C. Cobelli. Comparison of marker less and marker-based motion capture technologies through simultaneous data collection during gait: proof of concept. PLoS One.2014
- [4] N. Hofmann, J. Alhakeem, H. Hermsdorf, D. Möbius, S. Öhmichen, H. Schulz, M. Kaden, and T. Villmann: Analysemethoden zur Abschätzung von Belastungsintensitäten bei Kniebeugen unter Verwendung interpretierbare Modelle der Künstlichen Intelligenz: Machine Learning Report 01, MLR-01-2020, 2020
- [5] A. S. Iwashita and J. P. Papa: An Overview on Concept Drift Learning. IEEE Access, vol. 7, pp. 1532-1547, 2019 L. J. P. van der Maaten, G. E. Hinton: Visualizing Data Using t-SNE. Journal of Machine Learning Research. 9: 2579- 2605, 2008
- [6] C. Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead, Nature Machine Intelligence 1 (5), 206-21, 2019
- [7] K. Crammer, R. Gilad-Bachrach, A. Navot and N. Tishby. Margin Analysis of the LVQ Algorithm. Advances in Neural Information Processing Systems, 2003
- [8] T. Villmann, S. Haase, and M. Kaden. Kernelized vector quantization in gradient-descent learning. Neurocomputing 147, 83-95, 2015.
- [9] T. Villmann and S. Haase. Divergence based vector quantization. Neural Computation, 2011
- [10] A. Sato and K. Yamada. Generalized learning vector quantization. Advances in neural information processing systems, 1996.
- [11] P. Schneider, M. Biehl, and B. Hammer. Adaptive relevance matrices in learning vector quantization. Neural Computation, vol. 21, pp. 3532-3561, 2009
- [12] S. Saralajew, and T. Villmann. Transfer learning in classification based on manifold models and its relation to tangent metric learning. International Joint Conference on Neural Networks (IJCNN). IEEE, 2017

- [13] B. Paassen, A. Schulz and B. Hammer. Supervised Linear Transfer Learning for GMLVQ. Proceeding on Workshop New Challenges in Neural Computation, MLR-03-2016, 14-19, 2016



Figure 4: Classification result integrated in a tool provided by IfM

Quantum Computing for Efficient Learning in Prototype-based Vector Quantization

Thomas Villmann¹, Alexander Engelsberger

Univ. of App. Sciences Mittweida, Saxon Institute for Computational Intelligence and Machine Learning,
Technikumplatz 17, 09648 Mittweida, Germany

Prototype-based Vector Quantization is one of the key methods in data processing like data compression or interpretable classification learning. Prototype vectors serve as references for data and data classes. The data are given as vectors representing objects by numerical features. Famous approaches are the Neural Gas Vector Quantizer (NGVQ) for data compression and Learning Vector Quantizers (LVQ) for classification tasks. Frequently, training of those models is time consuming. In the contribution we discuss modifications of these algorithms adopting ideas from quantum computing. The aim for this is a least twofold: First quantum computing provides ideas for enormous speedup making use of quantum mechanical systems and inherent parallelization. Second, considering data and prototype vectors in terms of quantum systems, implicit data processing is performed, which frequently results in better data separation. We will highlight respective ideas and difficulties when equipping vector quantizers with quantum computing features.

1. Introduction

The combination of methods from physics and machine learning frequently is stimulating the development of new paradigms in machine learning [2]. Starting with the pioneering work and vision by R.P. Feynman [4,5], quantum computing became one of the most challenging approaches in efficient data processing while offering a great perspective to accelerate data processing [10]. This expected behavior makes them attractive for machine learning, which usually requires huge computational power [1]. Otherwise, machine learning may help to improve quantum approaches. Regardingly, theoretical concepts known from successful machine learning paradigms can be transferred to quantum computing [16].

Thereby, an important aspect to be kept is to achieve computational machine learning approaches, which allow easy interpretation [24]. This is different from explaining the result of underlying neural network – it is more to obtain interpretability by network design [12]. The family of prototype based neural networks in supervised and unsupervised provides a variety of smart networks which are easy to interpret and which have low computational complexity [3,8]. This makes them to be an alternative for deep networks if time complexity and memory restrictions do not allow complex neural networks as well as if interpretability is required as it is frequently demanded, for example in technical and medical systems.

Otherwise, both, quantum and machine learning systems are powerful due to inherent parallelism (Fig. 1) and, hence, are predestined for combination to benefit from their advantages twice [19].

Hence, combining both areas, prototype-based vector quantization and quantum computing, provide a promising concept to benefit from the best properties of quantum and neural computing.

2. Prototype-based Vector Quantization and Quantum Computing

2.1. Ingredients

Prototype-based vector quantization distributes reference vectors $W = \{w_k\}$ in the data space \mathbb{R}^n such that the data distribution $P(x)$ is approximated. In case of classification learning, this approximation is class dependent. Thus, the prototype vectors are taken as reference for the original data or classes. They are frequently denoted as prototype vectors or simply prototypes.

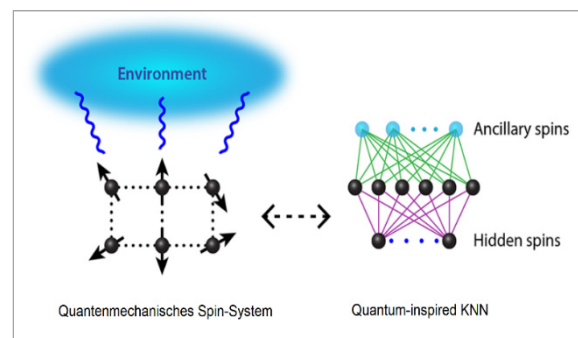


Fig. 1: Conceptual similarity between quantum mechanical systems and quantum-inspired artificial neural network.²

Depending on the task we can distinguish unsupervised and supervised vector quantization. Unsupervised models deal with clustering and data representation whereas supervised models are for regression and classification. Thus, conditional probabilities $P(x|w_k)$ are estimated in

¹ corresponding author: Thomas.villmann@hs-mittweida.de

² adaptiert von: <https://physics.aps.org/articles/v12/74>, M. Schuld, Xanadu, Toronto, Canada

unsupervised learning. In supervised learning the conditional class probabilities $P(c(x)|w_k)$ are subject of learning, where $c(x)$ is either a class label or a regression value to be predicted by the model after training.

In contrast to deep networks which usually require a huge memory and, hence, take long training time [6], prototype-based networks are interpretable, have low complexity and are robust against disturbed/corrupted data. The reference vectors correspond to perceptron weights, but the geometry interpretation by distance-based vector quantization is kept. Learning takes place as Hebbian competition learning and vector shift operations for the prototypes, if the Euclidean distance is used. Particularly, the prototypes are 'attracted' by a presented data sample, i.e. shifted toward or are repelled (repulsing).

After training, the network response is calculated for a given data sample based on the prototype with minimum distance (winner-take-all rule). It turns out that beside the interpretability these neural networks are robust regarding disturbed data and give mathematical guarantees for classification certainty [13].

2.2. Quantum Computing

Quantum computing is the realization of numerical calculation based on quantum mechanical systems [10,18]. Currently, first quantum computers are available like systems by IBM, Google or Xanadu. Thus, quantum algorithms can be tested on real hardware systems and first application can be realized.

Yet, one restriction for those quantum algorithms is that all computations have to follow the rules of underlying quantum physics [15]. This makes it difficult to simply adapt existing algorithms to quantum computing requirements.

Mathematically speaking, the description of those systems has to be in agreement with the postulates of quantum mechanics. In consequence, all training of machine learning systems have to be described by unitary operators and unitary transformation while the data live in a Hilbert space [17,20]. In this context, the data are assumed to be states of quantum systems, which can be described as respective amplitude vectors of the corresponding quantum mechanical system or, equivalently, in terms of density matrices. Depending on the task, the choice has to be made for that model, which gives more mathematical convenience.

From a computer science perspective, the concept of bits as information storing units taking states (values) zero and one has to be replaced by quantum bits (qubits), which can take any intermediate state if not observed/measured (see Fig. 2). This property, usually denoted as superposition, leads to much greater flexibility

and inherent parallelism providing the computational power. Further, those superpositions are also denoted as mixed states which closely relate to qubit entanglement [15,18].

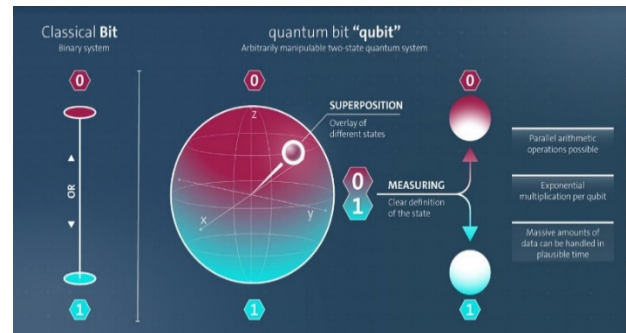


Fig. 2: Comparison of a quantum bit and a classical bit.³

Frequently, quantum algorithms are described by means of quantum circuits based on quantum gates in analogy to usual digital circuits and gates, an example is depicted in Fig. 3.

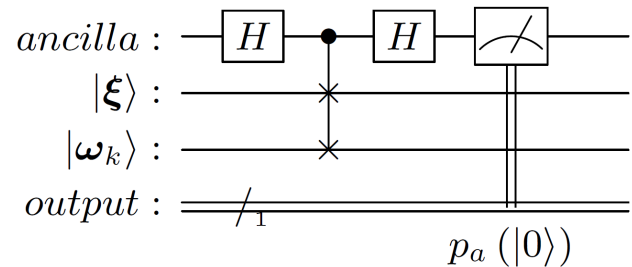


Fig.3 Quantum circuit of a cSWAP-gate realizing a swap of the quantum registers ξ and ω_k . The measurement is used to estimate the probability of the zero state. This probability relates to the inner product of the states in mathematical terms.

2.3. Quantum Computing for Vector Quantization

In quantum vector quantization the data as well as the prototypes are assumed to represent quantum registers of qubits [22]. The learning has to be in accordance with mathematics of quantum systems but always to realizing an nearest neighbor (prototype) scheme based on the winner-takes-all principle [25]. The most prominent examples are the quantum-k-means and quantum support vector machines for unsupervised and supervised learning, respectively [7,11]. Both methods were developed by means of the mathematical framework for Hilbert space representations.

Alternatively, classical clustering can be performed in terms of minimization of a cost function, which is related to the Schrödinger equation in quantum computing context.

³ Taken from <https://iot-automotive.news/volkswagen-optimizes-traffic-flow-with-quantum-computers/>

3. Neural Vector Quantization and Quantum Computing

Neural vector quantization frequently outperforms standard vector quantizers like the widely applied k-means algorithm and shows generally a more robust behavior. Further, these learning algorithms are insensitive to initialization. This behavior is obtained adopting parallelization and cooperativeness of real neural systems as it is observed in the cortical areas. This cooperativeness is incorporated in respective attraction and repulsing schemes of the prototypes to achieve faster learning as well as the demanded robust and insensitive behavior regarding disturbed data.

The most prominent approaches are the neural gas vector quantizer and the self-organizing map for unsupervised learning [9,8]. For classification learning, the family of learning vector quantizers became famous for interpretable and smart classification learning [21]. They have in common that the attraction learning is realized in all those networks [3].

According to the attraction scheme for prototype updates, the density matrix representation of quantum states is beneficial if these neural vector quantizers should be applied in quantum machine learning [22]: The distance between data (states) and prototypes can be calculated by means of the Frobenius-norm regarding the density matrix difference. The attraction scheme for a prototype update can be realized by a convex sum of the participating density matrices. The resulting adapted prototype density matrix represent in general a mixed quantum state. For a detailed description of the approach we refer to [22].

The respective mathematical (numerical) operations can be (theoretically) realized by quantum hardware systems. Currently, we implement them in a simulation software environment, such that we can study the numerical behavior and stability.

4. Conclusions

In this contribution we have highlighted aspects of vector quantization for unsupervised and supervised learning in context of quantum computing, which currently is close to the step of becoming working on real hardware systems.

We have explained that neural vector quantizers are interpretable and smart artificial neural networks, which can be efficiently combined with quantum computing concepts and algorithms.

Further, it turns out that machine learning quantum vector quantizers show a close relation to kernel approaches [16,20]. In particular, the pre-processing of usual data in such a way that the data can be seen as quantum state representations, generally is a non-linear mapping. By means of the Hilbert space description we are able to manipulate the learning quantum system in

that Hilbert space in an explicit way. In contrast, the frequently successful kernel trick in classical vector quantization allows to work with the original data while the manipulations in the kernel feature space are done implicitly using the kernel property [14,23]. Yet, this scheme reduces the interpretability of the network behavior.

However, a mathematical understanding of Hilbert space representation in quantum vector quantization is given considering kernel methods and beyond in context of vector quantization and studying the topological as well as geometrical analogies to the quantum approach.

In summery, the authors emphasize to make effort to understand, analyze and finally realize neural vector quantization approaches in the context of quantum computing systems to benefit from both, an interpretable neural network approach and the quantum parallelization for the speed-up of machine learning. For that reason, research for respective mathematical concepts is demanded as well as realizations on quantum computing hardware.

Acknowledgement

A. Engelsberger is supported by a ESF PhD grant during the time period 2020 to 2022.

This work is co-financed by tax funds on the basis of the budget adopted by the Saxon Landtag's budget.



Diese Maßnahme wird mitfinanziert durch Steuermittel auf der Grundlage des vom Sächsischen Landtag beschlossenen Haushaltes.

Bibliography

- [1] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, S. Lloyd: Quantum machine learning. *Nature* 549(9), 193-202, 2017.
- [2] M. Biehl, N. Caticha, M. Opper, T. Villmann: Statistical physics of learning and inference. Proceedings of the 33rd European Symposium on Artificial Neural Networks and Machine Learning. P. 501-509, i6doc.com, 2019.
- [3] M. Biehl, B. Hammer, T. Villmann: Prototype-based models in machine learning. *Wiley Interdisciplinary Reviews: Cognitive Science* 7(2),92-111, 2016.
- [4] R.P. Feynman: Simulating physics with computers. *International Journal of Theoretical Physics* 21(6/7),467-488, 1982.
- [5] R.P. Feynman: Quantum Mechanical Computers. *Optic News* 11(2), 11-20, 1985.
- [6] I. Goodfellow, Y. Bengio, A. Courville: Deep Learning. MIT Press, 2016.
- [7] I. Kerenidis, J. Landman, A. Luongo, A. Prakash: q-means: A quantum algorithm for unsupervised machine learning. *Advances in Neural Information Processing Systems* 32 (Proc. NeurIPS 2019), 4134-4144, 2019.
- [8] T. Kohonen. *Self-Organizing Maps*. 2nd ed., Springer, 1997.

- [9] T. Martinetz, S.G. Berkovich, K.J. Schulten. 'Neural-gas' network for vector quantization and its application to time-series prediction. *IEEE Transactions on Neural Networks* 4(4), 558-569, 1992.
- [10] M.A. Nielsen, I.L. Chuang: Quantum Computation and Quantum Information. Cambridge University Press, 2016.
- [11] P. Rebentrost, M. Mohseni, S. Lloyd: Quantum support vector machines for big data classification. *Physical Review Letters*, 113(13050):1-5, 2014.
- [12] C. Rudin: Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence*, 1(5):206-215, 2019.
- [13] S. Saralajew, L. Holdijk, T. Villmann: Fast Adversarial Robustness Certification of Nearest Prototype Classifiers for Arbitrary Seminorms. Proceedings of the 34th Conference on Neural Information Processing Systems (NeurIPS 2020), MIT Press, 2020.
- [14] B. Schölkopf, A. Smola: Learning with Kernels. MIT Press, 2002.
- [15] W. Scherer: Mathematics of Quantum Computing. Springer, 2019.
- [16] M. Schuld: Quantum machine learning models are kernel methods. arXiv:2101.11020 [quant-ph], 2021.
- [17] M. Schuld, N. Killoran: Quantum machine learning in feature Hilbert space. *Physical Review Letters* 122(4), 1-6, 2019.
- [18] M. Schuld, F. Petruccione: Supervised Learning with Quantum Computers. Springer, 2018.
- [19] M. Schuld, I. Sinayskiy, F. Petruccione: An introduction to quantum machine learning. *Contemporary Physics*, DOI: 10.1080/00107514.2014.964942, 2014.
- [20] T. Villmann, A. Engelsberger: Quantum-Inspired Learning Vector Quantizers for Prototype-based Classification. *Neural Computing and Applications*, 2020, 1-10.
- [21] T. Villmann, A. Bohnsack, M. Kaden: Can Learning Vector Quantization be an Alternative to SVM and Deep Learning? *Journal of Artificial Intelligence and Soft Computing Research*, 7(1), 65-81, 2017.
- [22] T. Villmann, A. Engelsberger: Quantum-hybrid Neural Vector Quantization – A Mathematical Approach. Internat. Conference on Artificial Intelligence and Soft-Computing 2021, Zakopane (Poland), Lecture Notes in Computer Science, Springer, 2021.
- [23] T. Villmann, S. Haase, M. Kaden: Kernelized Vector Quantization in Gradient-Descent Learning. *Neuro-computing*, 147, 83-95, 2015.
- [24] T. Villmann, S. Saralajew, A. Villmann, M. Kaden: Learning Vector Quantization Methods for Interpretable Classification Learning and Multilayer Networks. Proceedings of the 10th International Joint Conference on Computational Intelligence (IJCCI), Sevilla, SCITEPRESS - Science and Technology Publications, Lda., p. 15-21, 2018.
- [25] N. Wiebe, A. Kapoor, K.M. Svore: Quantum algorithms for nearest-neighbor methods for supervised and unsupervised learning. *Quantum Information and Computation*, 15(3-4), 316-356, 2015.

An attempt to explain double descent in modern machine learning

Jochen Merker, Gregor Schuld
HTWK Leipzig, PF 30 11 66, D-04251 Leipzig

This article aims to explain mathematically, why the so called double descent observed by Belkin et al., Reconciling modern machine-learning practice and the classical bias-variance trade-off, PNAS 116(32) (2019), p. 15849-15854, occurs on the way from the classical approximation regime of machine learning to the modern interpolation regime. We argue that this phenomenon may be explained by a decomposition of mean squared error plus complexity into bias, variance and an unavoidable irreducible error inherent to the problem. Further, in case of normally distributed output errors, we apply this decomposition to explain, why LASSO provides reliable predictors avoiding overfitting.

1. Introduction

While standard statistical machine learning theory [1] predicts that bigger models should be more prone to overfitting, modern machine learning practitioners claim that bigger models are always better. Within the approximation regime of machine learning, the first point of view is supported by the classical property of bias-variance trade-off well-known in literature since the pioneering work [2]. However, in the modern interpolation regime of deep learning, where there are so many parameters that data points do not have to be approximated but may be interpolated, a decay of the bias towards an asymptote was observed by [3] in neural networks as the network width increases. Belkin et al. [4] provided strong evidence for the existence and ubiquity of such a double descent (see Figure 1) for a wide spectrum of data models including decision trees and simple neural networks.

A concise explanation of this double descent phenomenon still seems to be missing in literature, although there are recent attempts [5] using a fine-grained bias-variance decomposition of the mean-squared error (MSE). In this article, we follow a similar strategy and aim to provide an explanation for the occurrence of double descent in machine learning using complexity [6]. This corresponds to the fact that

- an underfitted model has a high MSE on training data, high bias, low variance, low complexity,
- an overfitted model has a low MSE on training data, low bias, high variance, high complexity,
- a complexity reduced interpolation model has zero MSE on training data, low bias, reduced variance, reduced complexity.

Further, we will offer an explanation why LASSO [7], i.e. L^1 -regularized approximation, occurs naturally in statistical learning and provides reliable predictors avoiding overfitting.

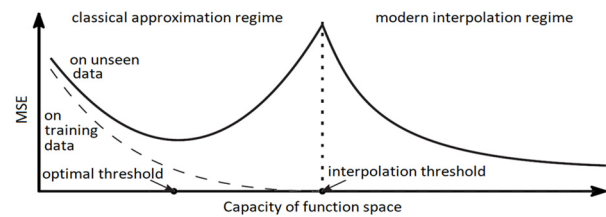


Figure 1: MSE on unseen data is a double descent curve [4]

2. Standard Bias-Variance Decomposition

Machine learning methods estimate an unknown deterministic input-output map from given inputs $u_i \in \mathcal{U}$ and corresponding outputs $y_i \in \mathcal{Y}$, $i = 1, \dots, N$, as training data. In the mathematical framework of statistical learning theory, it is assumed that

- the inputs u_i , $i = 1, \dots, N$, are drawn from a measurable space \mathcal{U} independently according to a fixed but unknown distribution of a random variable U ,
- for every input u_i an output y_i is drawn from a measurable space \mathcal{Y} according to a fixed but unknown conditional distribution of a random variable Y given $U = u_i$.

For simplicity, let us further assume that \mathcal{U} and \mathcal{Y} are finite-dimensional real Hilbert spaces, and that the distribution of the input resp. the conditional distribution of the output to an input are continuous, i.e. have densities $f_U(u)$ resp. $f_{Y|U}(y|u)$ w.r.t. Lebesgue measure. Then, the density of the joint distribution of (U, Y) is given by $f_{(U, Y)}(u, y) = f_{Y|U}(y|u)f_U(u)$, and the unknown deterministic input-output map $F: \mathcal{U} \rightarrow \mathcal{Y}$ is the expected output given an input, i.e.

$$Fu = E(Y|U = u) = \int_{\mathcal{Y}} y f_{Y|U}(y|u) dy.$$

Because a machine learning method selects on the basis of training data (u_i, y_i) , $i = 1, \dots, N$, a map $F_{(u_1, y_1), \dots, (u_N, y_N)}$ from a D -dimensional space \mathcal{F} of functions mapping \mathcal{U} to \mathcal{Y} as estimation of the unknown deterministic input-output map $F = E(Y|U)$, in statistical learning theory the predictor itself has to be considered as random vari-

able $F_T = F_{(U_1, Y_1), \dots, (U_N, Y_N)}$ induced by the machine learning method and by the random variables (U_i, Y_i) , $i = 1, \dots, N$, which are i.i.d. as (U, Y) .

Within this theory, the failure of overfitted models – low MSE on training data, but high variation w.r.t. training data and therefore bad generalization to unseen data – can be mathematically explained by the decomposition

$$E_{(U, Y), T}(|F_T U - Y|_Y^2) = E_{U, T}(|F_T U - E_T(F_T)U|_Y^2) + E_U(|E_T(F_T)U - E(Y|U)|_Y^2) + E_{(U, Y)}(|Y - E(Y|U)|_Y^2) \quad (1)$$

of the expected squared error of F_T w.r.t. training data T into three parts: The variance of F_T w.r.t. training data, the squared bias of the predictor $E_T(F_T)$ expected from training data, and the variance of Y as an unavoidable irreducible error inherent to the problem. Now, if you enhance the capacity of the function space \mathcal{F} by increasing the dimension D beyond an optimal threshold, then the squared bias of the predictor $E_T(F_T)$ often decreases, because $E_T(F_T)U$ can better resemble the output $FU = E(Y|U)$ expected at U . However, variance $E_{U, T}(|F_T U - E_T(F_T)U|_Y^2)$ w.r.t. training data T often increases more strongly, leading in total to a larger expected squared error of F than at the optimal threshold. This bias-variance trade-off explains the failure of overfitted models in the classical approximation regime of machine learning. Yet, there is no necessity that a machine learning method shows this behaviour, and in fact, a simultaneous decrease of variance and bias can be observed in the modern interpolation regime of machine learning [3,4].

3. Bias-Variance-Complexity Decomposition

In this section, we aim to explain double descent and simultaneous decrease of variance and bias in the modern interpolation regime of machine learning by a bias-variance decomposition of mean squared error plus complexity. Let us consider the mean squared error with random training data, i.e. the random variable

$$\frac{1}{N} \sum_{i=1}^N |F_T U_i - Y_i|_Y^2$$

If we try to calculate the expected value of this random variable w.r.t. training data and to obtain a similar decomposition as (1), it makes sense to notationally separate the i -th variable from the other variables by setting $F_{(u_i, y_i)} := E_{T_i}(F_{(u_i, y_i), T_i})$, where T_i does not contain (U_i, Y_i) . As a first partial result, we then obtain the decomposition

$$E_T(|F_T U_i - Y_i|_Y^2) = E_T(|F_T U_i - F_{(u_i, y_i)} U_i|_Y^2) + E_{(u_i, y_i)}(|F_{(u_i, y_i)} U_i - Y_i|_Y^2) \quad (2)$$

of the expected squared error at the i -th input into variance of F_T w.r.t. training data T and expected squared error of the predictor $F_{(u_i, y_i)} U_i$ with variable i -th input for training, because $E_T = E_{(u_i, y_i)} E_{T_i}$ and the expectation of the middle term in

$$|F_T U_i - Y_i|_Y^2 = |F_T U_i - F_{(u_i, y_i)} U_i|_Y^2$$

$$- 2 (F_T U_i - F_{(u_i, y_i)} U_i)(F_{(u_i, y_i)} U_i - Y_i) + |F_{(u_i, y_i)} U_i - Y_i|_Y^2$$

vanishes due to $E_T(F_T U_i - F_{(u_i, y_i)} U_i) = 0$. However, in the further decomposition of $E_{(u_i, y_i)}(|F_{(u_i, y_i)} U_i - Y_i|_Y^2)$ the middle term does not vanish: The middle term on the right of

$$\begin{aligned} & E_{(u_i, y_i)}(|F_{(u_i, y_i)} U_i - Y_i|_Y^2) \\ &= E_{(u_i, y_i)}(F_{(u_i, y_i)} U_i - E(Y_i|U_i))_Y^2 \\ &\quad - 2 E_{(u_i, y_i)}\left(\left(F_{(u_i, y_i)} U_i - E(Y_i|U_i)\right)(Y_i - E(Y_i|U_i))\right) \\ &\quad + E_{(u_i, y_i)}(|Y_i - E(Y_i|U_i)|_Y^2) \end{aligned} \quad (3)$$

is given by

$$\begin{aligned} & E_{(u_i, y_i)}\left(\left(F_{(u_i, y_i)} U_i - E(Y_i|U_i)\right)(Y_i - E(Y_i|U_i))\right) \\ &= \int_U \left(\int_Y (F_{(u, y)} u - Fu)(y - Fu) f_{Y|U}(y|u) dy \right) f_U(u) du, \end{aligned}$$

and this term could be named expected complexity of F_T . Note that the scalar product inside the integral is positive, if the deviation of the predicted output $F_{(u, y)} u$ from the deterministic output $Fu = E(Y|U = u)$ at u points in the same direction as the deviation of the stochastic output y , i.e. the complexity is positive if $F_{(u_i, y_i)} U_i$ captures the behaviour of the noise in the output Y_i . Of course, this is an undesired property, thus a machine learning method should be so that complexity of the predictor is low. While complexity has some similarities with covariance and correlation, see e.g. [8], the term is not a covariance or correlation, as $E(Y_i|U_i)$ is not the expectation value of $F_{(u_i, y_i)} U_i$.

If moreover – like for most machine learning methods – the order of the training data is not important for the estimation of the input-output map, then from (2) and (3) we obtain the following Bias-Variance-Complexity decomposition.

3.1 Main Theorem

Under the assumptions of statistical learning theory (and assuming existence of the expectation values), with the above definitions the decomposition

$$\begin{aligned} & E_T\left(\frac{1}{N} \sum_{i=1}^N |F_T U_i - Y_i|_Y^2\right) \\ &+ 2 E_{(U, Y)}\left(\left(F_{(U, Y)} U - E(Y|U)\right)(Y - E(Y|U))\right) \\ &= E_T\left(\frac{1}{N} \sum_{i=1}^N |F_T U_i - F_{(u_i, y_i)} U_i|_Y^2\right) \\ &\quad + E_{(U, Y)}(F_{(U, Y)} U - E(Y|U))_Y^2 \\ &\quad + E_{(U, Y)}(|Y - E(Y|U)|_Y^2) \end{aligned} \quad (4)$$

of the expected mean squared error w.r.t. training data T plus twice the complexity into variance of F_T w.r.t. training data T , squared bias of the predictor $F_{(U, Y)} U$ with one input variable for training, and the variance of Y as an unavoidable irreducible error inherent to the problem holds.

Hence, if the machine learning method is such that on an enhancement of the capacity of the function space \mathcal{F} the mean squared error plus complexity decreases, then variance plus bias decreases simultaneously, too. This seems to be the case for many machine learning methods in the modern interpolation regime and may explain that after the interpolation threshold there occurs a second descent of the error on unseen data, which is related to a decrease of complexity $E_{(U,Y)} \left((F_{(U,Y)}U - E(Y|U)) (Y - E(Y|U)) \right)$.

4. Complexity and LASSO

Complexity is tightly related to LASSO, i.e. L^1 -regularized approximation. In fact, if $\mathcal{Y} = \mathbb{R}^k$ and $Y - E(Y|U = u) \sim N(0, \sigma_u^2 Id)$ is normally distributed, then the inner integral in the definition of complexity allows a partial integration

$$\begin{aligned} & \frac{1}{\sigma_u (2\pi)^{k/2}} \int_{\mathbb{R}^k} (F_{(u,y)}u - Fu)(y - Fu) e^{-\frac{(y-Fu)^2}{2\sigma_u^2}} dy \\ &= \frac{-\sigma_u}{(2\pi)^{k/2}} \int_{\mathbb{R}^k} (F_{(u,y)}u - Fu) \nabla_y \left(e^{-\frac{(y-Fu)^2}{2\sigma_u^2}} \right) dy \\ &= \frac{\sigma_u}{(2\pi)^{k/2}} \int_{\mathbb{R}^k} \text{div}_y (F_{(u,y)}u) e^{-\frac{(y-Fu)^2}{2\sigma_u^2}} dy . \end{aligned}$$

Thus, complexity is given in this case by

$$\begin{aligned} & E_{(U,Y)} \left((F_{(U,Y)}U - E(Y|U)) (Y - E(Y|U)) \right) \\ &= E_{(U,Y)} (\sigma_u^2 \text{div}_Y (F_{(U,Y)}U)) \end{aligned}$$

and the decomposition (4) reads as

$$\begin{aligned} & E_T \left(\frac{1}{N} \sum_{i=1}^N |F_T U_i - Y_i|_{\mathcal{Y}}^2 \right) \\ & \quad + 2 E_{(U,Y)} (\sigma_u^2 \text{div}_Y (F_{(U,Y)}U)) \\ &= E_T \left(\frac{1}{N} \sum_{i=1}^N |F_T U_i - F_{(u_i, y_i)} U_i|_{\mathcal{Y}}^2 \right) \\ & \quad + E_{(U,Y)} (F_{(U,Y)}U - E(Y|U))_{\mathcal{Y}}^2 \\ & \quad + E_{(U,Y)} (|Y - E(Y|U)|_{\mathcal{Y}}^2) \end{aligned} \quad (5)$$

Hence, under the assumption that the true input-output map $F = E(Y|U)$ has a derivative $\frac{\partial}{\partial u} Fu =: \gamma_u$ w.r.t. u and the estimated input-output map $F_{(\tilde{u}, y)}u$ has approximately the same derivative w.r.t. u at $\tilde{u} = u, y = Fu$, i.e. if there are $\sigma_i \approx \sigma_{u_i}^2 \gamma_{u_i}$ such that

$$\sigma_{u_i}^2 \text{div}_{y_i} (F_{(u_i, y_i)} u_i) = \sigma_i \nabla F_{(\tilde{u}, y)} u_i |_{\tilde{u}=u_i, y=y_i},$$

a machine learning method could estimate for training data $t = (u_1, y_1), \dots, (u_N, y_N)$ the input-output map by solving

$$\frac{1}{N} \sum_{i=1}^N |F u_i - y_i|_{\mathcal{Y}}^2 + \frac{1}{N} \sum_{i=1}^N \sigma_i |\nabla F u_i| = \min_{F \in \mathcal{F}} ! \quad (6)$$

to simultaneously decrease variance and bias when enhancing the capacity of the function space \mathcal{F} . Now assume that σ_i is bounded by a constant C , then the second term is dominated by

$$\frac{C}{N} \sum_{i=1}^N |\nabla F u_i| .$$

This is exactly the L^1 -regularizer in LASSO regression, and if this regularizer is used instead of the second term in (6), then the minimal value dominates the left and right hand side of (5). Hence, under the above assumptions LASSO regression provides reliable predictors in machine learning and leads to a simultaneously decrease of variance and bias when enhancing the capacity of the function space \mathcal{F} .

Further, even if no LASSO regression problem is solved, often machine learning methods – like e.g. stochastic gradient descent – generate complexity reduced interpolation models, and thus automatically lead to a descent of the MSE on unseen data after the interpolation threshold. Let us mention in 4.1 one such case, and in 4.2 a case where LASSO is explicitly used.

4.1 Autoencoder neural networks

In the example of an autoencoder neural network in [9] trained by data $y_i = u_i$, where gradient descent is used to minimize mean squared error for a deep neural network in the modern interpolation regime of machine learning, the weights do not converge to the identity but to the matrix W of lowest rank satisfying $W u_i = u_i$. Hence, MSE is minimal and complexity is reduced, so that variance and bias are simultaneously low and the predictor is nearly the identity on training data.

4.2 Sparse Support Vector Machines

To generate a sparse soft margin support vector machine [10], the constrained optimization problem

$$\begin{aligned} & \frac{1}{2} \sum_{i=1}^N |\xi_i|^2 + C \sum_{j=1}^D |\omega_j| = \min_{(\omega, \xi, \beta)} ! \\ & \text{s. t. } (\omega \cdot u_i + \beta) y_i \geq 1 - \xi_i \end{aligned}$$

may be solved for data points u_i with given classification $y_i = \pm 1$. The predictor then is the linear classifier $Fu = \omega \cdot u + \beta$, and $\nabla F = \omega$ holds so that above really LASSO is used, i.e. $\sum_{j=1}^D |\omega_j|$ is the L^1 -regularization term. Therefore, variance plus bias decreases simultaneously when enhancing the capacity of the function space \mathcal{F} by increasing the dimension D .

Acknowledgements

The first author gratefully acknowledges the financial support of the German Federal Ministry for Economic Affairs and Energy within the ZIM project 16KN072420. The second author gratefully acknowledges the financial support of the Saxonian Ministry of Science and Arts as well as the European Union within the research project SmartKMU (Project No. 100 38 21 59).

Bibliography

- [1] V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer, 1995.
- [2] S. Geman, E. Bienenstock, R. Doursat, *Neural computation* 4 (1992), 1-58.
- [3] S. Spigler, M. Geiger, S. d'Ascoli, L. Sagun, G. Biroli, M. Wyart, *Journal of Physics A: Mathematical and Theoretical* 52 (2019), 474001.
- [4] M. Belkin, D. Hsu, S. Ma, S. Mandal, *Proceedings of the National Academy of Sciences* 116 (2019), 15849-15854.
- [5] B. Adlam, J. Pennington, arXiv:2011.03321 (2020).
- [6] B. Ghogh, M. Crowley, arXiv:1905.12787 (2019).
- [7] T. Hastie, R. Tibshirani, M. Wainwright, *Statistical learning with sparsity: the lasso and generalizations*, CRC press, 2015.
- [8] J. Merker, G. Schuldt, *Proceedings of ICoMS 2020, ACM* (2020), DOI: 10.1145/3409915.3409920
- [9] A. Radhakrishnan, M. Belkin, C. Uhler, *Proceedings of the National Academy of Sciences* 117 (2020), 27162-27170.
- [10] J. Merker, *Journal of Advances in Applied Mathematics* 2 (2017), 109-114.

Sensors Data Fusion for Smart Decisions Making Using Interpretative Machine Learning Models

Feryel Zoghlami^a, Marika Kaden^b, Thomas Villmann^b, Germar Schneider^a and Harald Heinrich^a

^a Automation, Maintenance and Factory Integration Infineon Technologies Dresden GmbH & Co. KG, Dresden, Germany

^b SICIM, Hochschule Mittweida, Mittweida, Germany

Sensor fusion is an important and crucial topic in many industrial applications. One of the challenging problems is to find an appropriate sensor combination for the dedicated application or to weight their information adequately. In our contribution, we focus on the application of the sensor fusion concept together with the reference to the distance-based learning for object classification purposes. The developed machine learning model has a bi-functional architecture, which learns on the one side the discrimination of the data regarding their classes and, on the other side, the importance of the single signals, i.e., the contribution of each sensor to the decision. We show that the resulting bi-functional model is interpretative, sparse, and simple to integrate in many standard artificial neural networks

Sensor fusion gains much interest during the last years. There are lot of advantages to combine the signals of several sensors: systems can get better reliability, stability and confidentiality or sensor failures can be compensated. Yet, the fusion of several sensor signals leads to challenges, too. Several questions must be answered for successful fusion: Which sensors should be applied to solve a specific problem? Which sensors are more robust? Are all sensors necessary and should all have the same influence on the decision?

Beside technical expertise for the task and sensors, artificial intelligence (AI) can go for answering these questions. In several publications the performance benefit applying different sensor fusion approaches using AI is analyzed [1,2,3]. From that, we notice that researchers' efforts are mainly dedicated to try different combination formats of sensing data. Yet, the importance of the single sensors for the performance are still hard to detect. Obviously, the influence is sensor and task depending and generally difficult or impossible to answer in general. However, for a given specific task AI may help to capture this information analyzing sensor behavior and comparing this to expected behavior.

In our contribution we present a neural network architecture taking a multi-sensor data input [4]. Thereby the network is built by two main blocks: the feature generation block and the bi-functionality block including the classification and evaluation of the influence of multi sensor input (see illustration Fig.\ref{bi}). For the feature generation block, standard deep neural networks are chosen, which have been shown to be immensely powerful in many application areas, e.g., ResNet for image processing [7]. For the second block we use the prototype-based Generalized Learning Vector Quantization for the final classification task (GLVQ, [6]) integrated in a network structure [5]. This network ar-

chitecture provides a competition layer, i.e., the data in the feature space were evaluated by calculating a distance to learned class-dependent prototypes and comparing them to find the best matching prototype. This is also known as Winter-Takes-All rule. Thus, the input is assigned to the class of this nearest prototype. The used distance measures can be also parameterized, in this sense that the influence of each sensor signal (feature) is weighted according to the importance for the discrimination [8,9]. We use this feature adapting property of GLVQ to learn the contribution of each sensor for our classification task additionally. Thus, our network response not only a classification model, but it also provides us the sensor importance.

We applied our network with the bi-functional block to detect humans using image of different sensor signals. Here three sensor signals are used, depth images and the aligned amplitude images of a time-of-flight camera and a radar image. It turns out, that the highest contribution has the amplitude image. Nevertheless, the depth and radar sensors are not unimportant.

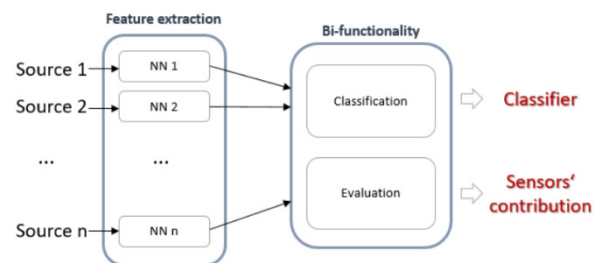


Figure 1 Illustration of a bi-functionality system for sample classification and sensors evaluation [4]

Acknowledgements

Project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 826452. The JU receives support from the European Union's Horizon 2020 research.



References

- [1] Dehzangi, O., Taherisadr, M., Changal-Vala, R. (2017). IMU-based gait recognition using convolutional neural networks and multi-sensor fusion. *Sensors*, 17(12), 2735.
- [2] E. Münzner, S., Schmidt, P., Reiss, A., Hanselmann, M., Stiefelhagen, R., Dürichen, R. (2017). CNN-based sensorfusion techniques for multimodal human activity recognition. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers*, pp. 158-165.
- [3] N Liu, J., Hu, Y., Wang, Y., Wu, B., Fan, J., Hu, Z. (2018). An integrated multisensory fusion-based deep feature learning approach for rotating machinery diagnosis. *Measurement Science and Technology*, 29(5).
- [4] Zoghalmi, F., Kaden, M., Villmann, T., Schneider, G. and Heinrich, H. (2021). Sensors data fusion for smart decisions making: A novel bi-functional system for the evaluation of sensors contribution in classification problems. *Advances in Digital Transformation of Manufacturing Environments - Integrated Development 4.0*, accepted
- [5] Saralajew, S., Rees, M., Holdijk, L., and Villmann, T. (2018). Prototype-based neural network layers: Incorporating
- [6] Sato, A. and Yamada, K. (1996). Generalized learning vector quantization. *Advances in neural information processing systems*.
- [7] ImageNet. <http://www.image-net.org>.
- [8] Hammer, B. and Villmann, T. (2002). Generalized relevance learning vector quantization. *Neural Networks*, 15, pp 1059-1068
- [9] Schneider, P., Biehl, M. and Hammer, B. (2009). Adaptive relevance matrices in learning vector quantization, *Neural Computation*, vol. 21, pp. 3532-3561

Recurrent Learning Vector Quantization

Jensun Ravihandran and Thomas Villmann

SICIM, Hochschule Mittweida, Mittweida, Germany

Learning Vector Quantization (LVQ) methods have been popular choices of classification models ever since its introduction by T. Kohonen in the 90s. These days, LVQ is combined with Deep Learning methods to provide powerful yet interpretable machine-learning solutions to some of the most challenging computational problems. However, techniques to model recurrent relationships in the data using prototype methods still remain quite unsophisticated. In particular, we are not aware of any modification of LVQ that allows the input data to have different lengths. Needless to say, such data is abundant in today's digital world and demands new processing techniques to extract useful information. In this paper, we propose the use of the Siamese architecture to not only model recurrent relationships within the prototypes but also the ability to handle prototypes of various dimensions simultaneously.

Learning Vector Quantization (LVQ) methods have been popular choices of classification models ever since its introduction by T. Kohonen in the 90s [1]. Since then, a plethora of improvements have been made over the years to the original formulation of the LVQ algorithm to handle several shortcomings. One essential extension was the Generalized LVQ (GLVQ) which introduced a cost function [6] by keeping the intention of the LVQ. These days, LVQ is combined with Deep Learning methods to provide powerful yet interpretable machine-learning solutions to some of the most challenging computational problems [4, 5].

However, techniques to model recurrent relationships in the data using prototype methods still remains quite unsophisticated. In particular, we are not aware of any modification of LVQ that allows the input data to have different lengths. This means that they are not particularly suited for applications in domains where the data are sequential, since requiring that all sequences in the dataset to be of the same length is severely limiting. Examples of such domains include but are not exclusive to DNA/RNA sequences, time-series prediction, and natural language processing. Needless to say, such data is abundant in today's digital world and demands new processing techniques to extract useful information. To this end, appropriate modifications to the LVQ algorithm are required to successfully employ them on sequential data.

In this contribution, we propose the use of the siamese architecture to model not only recurrent relationships within the data as well prototypes but also the ability to handle data samples of various lengths simultaneously. A siamese network, introduced first in 1993 [2], is a network type using a shared memory architecture. Here it was used to transform two images in high-dimensional spaces to simpler representations before comparing them. Since then, a plethora of publications have made use of the siamese architecture in very interesting ways. In our novel architecture that we denote as ReLVQ, we

employ a siamese architecture not to process two different inputs, but rather to process input data and trainable prototypes.

Interpretability is very often not the primary goal for machine learning methods. In fact, this is not a problem for many applications. However, recently there has been pushback against the use of black-box models for high-stakes decisions [3]. In the last years, deep learning has become so popular today in part because of its effectiveness for many practical problems. Further, a regarding network is easy to build-up by means of modern deep learning frameworks like Tensorflow and Py-Torch. However, we feel that the deep learning revolution has also turned machine learning into black magic. When certain deep architectures work very well for particular applications, we often do not understand why the decisions are made in this or that way. Even worse, when they do not work, we do not know why they failed. This means that choosing the right architecture frequently is a matter of trial and error. Needless to say, this is very time consuming, and without any guarantees of success. Models like GLVQ and matrix GLVQ (GMLVQ, [7]) are fully interpretable in that all layers in those networks lend themselves naturally to human interpretation. Particularly, this network type interprets data geometrically and adapts so-called prototypes in the data space acting their like sensors. Adaptation of them is realized by attraction towards a presented sample or repulsing, depending on correct classification or failed decision, respectively. However, these networks cannot directly handle sequential data as effectively as recurrent network models. This is in part because raw sequence data of different lengths cannot be processed without artificial manipulation like padding or clipping. Manual and labor-intensive feature engineering is therefore necessary such that pre-processed data can be feed into the network. ReLVQ circumnavigates this problem by incorporating recurrent structures to do the heavy-lifting, while still keeping the well-studied attraction-repulsion scheme from LVQ. Yet, ReLVQ is not longer fully interpretable

in the same fashion like GLVQ or GMLVQ. Here, the prototypes have to be interpreted as processing units to reflect the recurrent data information. There is no free lunch, but it may be a fair price to pay in certain applications

Acknowledgements

Project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 826452. The JU receives support from the European Union's Horizon 2020 research

This work is co-financed by tax funds on the basis of the budget adopted by the Saxon Landtag's budget.



Europa fördert Sachsen.
ESF
Europäischer Sozialfonds



Diese Maßnahme wird mitfinanziert durch Steuermittel auf der Grundlage des vom Sächsischen Landtag beschlossenen Haushaltes.

References

- [1] Dehzangi, O., Taherisadr, M., Changal-Vala, R. (2017). IMU-based gait recognition using convolutional neural networks and multi-sensor fusion. *Sensors*, 17(12), 2735.
- [2] Teuvo Kohonen. *Self-Organizing Maps*. Springer, Berlin, 1997.
- [3] Jane Bromley, James W. Bentz, Leon Bottou, Isabelle Guyon, Yann Lecun, Cliff Moore, Eduard Säckinger, and Roopak Shah. Signature verification using a "siamese" time delay neural network. *International Journal of Pattern Recognition and Artificial Intelligence*, 07, 1993.
- [4] Cynthia Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead, *Nature Machine Intelligence* 1 (5), 206-21, 2019.
- [5] Sascha Saralajew, Maike Rees, Lars Holdijk, and Thomas Villmann. Prototype-based neural network layers: Incorporating vector quantization, arXiv:1812.01214. 2018.
- [6] Jensun Ravichandran, Marika Kaden, Sascha Saralajew, and Thomas Villmann. Variants of drop connect in learning vector quantization networks for evaluation of classification stability. *Neurocomputing*, 403, 2020.
- [7] Atsushi Sato and Keiji Yamada. Generalized learning vector quantization. *Advances in neural information processing systems*, 1996.
- [8] Atsushi Sato and Keiji Yamada. Generalized learning vector quantization. *Advances in neural information processing systems*, 1996.

Machine Learning Algorithms for Classifying Component Defects for Predictive Maintenance

Christoph Kammerer, Micha Küstner, Michael Gaust, Pascal Starke

Roman Radtke, Alexander Jesser
University of Applied Sciences Heilbronn
Max-Planck-Str. 39, 74081 Heilbronn, Germany
CeraCon GmbH
Talstraße 2, 97990 Weikersheim, Germany

Reducing costs is an important part in today's business. Therefore manufacturers try to reduce unnecessary work processes and storage costs. Machine maintenance is a big, complex, regular process. In addition, the spare parts required for this must be kept in stock until a machine fails. In order to avoid a production breakdown in the event of an unexpected failure, more and more manufacturers rely on predictive maintenance for their machines. This enables more precise planning of necessary maintenance and repair work, as well as a precise ordering of the spare parts required for this. A large amount of past as well as current information is required to create such a predictive forecast about machines. With the classification of motors based on vibration, this paper deals with the implementation of predictive maintenance for thermal systems. There is an overview of suitable sensors and data processing methods, as well as various classification algorithms. In the end, the best sensor-algorithm combinations are shown.

1. Introduction

The topic of predictive maintenance (PMA) is becoming more and more important for industrial plants and is the key topic in mechanical engineering from the Industry 4.0 aspect [1]. PMA is defined as condition-based maintenance which is carried out on the basis of a wear or service life forecast [2]. PMA uses methods that allow for individual maintenance intervals of an industrial plant to be determined and the maintenance process to be initiated automatically. As part of a R&D cooperation project between CeraCon GmbH and the Heilbronn University of Applied Sciences, a thermal system is to be set up under automation and a PMA strategy is to be implemented, which should then be adaptable to other industrial plants¹. Due to the complexity of industrial plants, an intelligent solution is required in order to be able to offer individual maintenance strategies depending on the state of the plant. For this reason, the project uses machine learning (ML) methods. The essential steps of an intelligent PMA strategy are the digital acquisition of (sensor) data, their evaluation, the analysis of the acquired data and the prediction of probable events. First, possible component defect combinations (CDC) of the industrial plant were analyzed using standard technical risk analysis methods (FMEA, risk graph, fault tree analysis) [3]. CDC is the assignment of a wear component of the industrial system to a potentially occurring defect. Depending on the number of possible defects, a component can therefore have several CDCs. Each CDC was assigned an potential detection measure, e.g. physical vibration measurement or electrical current measurement. Suitable sensors were selected for the analyzed

detection measures and analyzed with regard to the PMA strategy. CDC's with the same detection methods were combined and measurement data recorded with the respective sensors. The core of this work is the evaluation of a combination of detection measures for data processing methods and ML algorithms. The optimal combination of these is a prerequisite for an efficient PMA strategy that can be used for the respective industrial plant.

2. State of the Art

A study by Bearingpoint [4] shows that PMA implementations capture 76% of the relevant data using suitable sensors, although only 59% of the process, measurement and machine data are evaluated in a targeted manner. There are three basic approaches to implementing a PMA strategy [5]. A basic approach is to use the already implemented sensors of the plant for process monitoring. This passive method is particularly suitable for systems that are already in operation. Another passive approach is to introduce dedicated sensors into the system. The additional sensors are introduced to monitor defined wear components and to detect potential defects. In the third approach, a test signal is actively fed into the system. The degree of wear of the components to be monitored can be deduced from the feedback. An example of this is Time Domain Reflectometry (TDR) [5].

3. Sensors and Test Set-Up

The requirement for a condition-based PMA is a structured data collection of sensor values. The following sensors were used to obtain status data:

¹ Sponsored by ZIM ZF4644801

- Three-axis acceleration sensors (Accelerometer):
 - LIS 3DH, MMA 8451, ADXL 343, ADXL 345
- Three-axis acceleration sensors with three-axis yaw rate sensor (gyroscope):
 - MPU 60.50
- Three-axis magnetic field sensor (magnetometer):
 - MLX 90393
- Multi sensors with three-axis acceleration, three-axis yaw rate and three-axis magnetic field measurement:
 - MPU 92.65, BNO 055, GY 250, GY 521

These recorded the acceleration, the rotation rate and the surrounding magnetic field of the fan motor R3G180-AJ11-XF from ebm-papst Mulfingen GmbH & Co. KG used in the thermal system. The fan motor was operated at fixed speeds, which were divided into 7 classes. This classification was based on the specific values 0%, 50%, 60%, 70%, 80%, 90% and 100% of the maximum engine speed. During the operation of the fan motor the vibration of the crankcase was sensed and recorded by the sensors. More than 980,000 structured sensor data sets per measurement series and sensor type were recorded. A total of more than 2.6 million data sets have thus been recorded for all sensor types [6]. An example of a recorded data set is shown in Figure 1. It shows the measured acceleration from the housing vibration in the spatial x- and z- orientation. The individual classes are highlighted in color to make a distinction possible. Due to the highest spatial coverage, it can be seen that the measurement results for class 80% can be assigned to the resonance range of the fan motor, since the acceleration values in the x- and z- alignment are at their maximum values here.

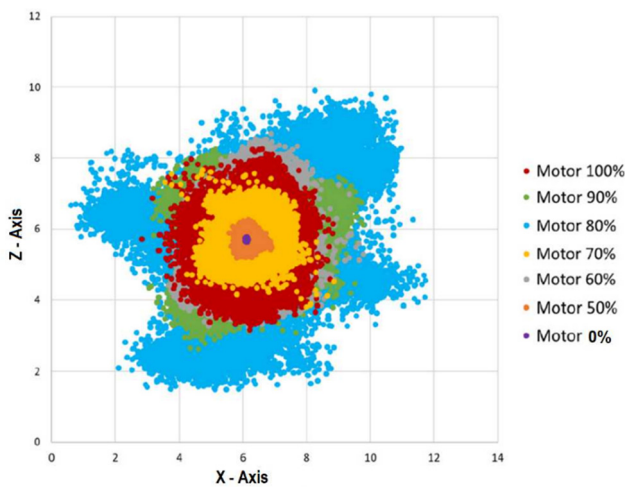


Figure 1: Measurement results from the acceleration sensor MMA 8451.

4. Data Conditioning

In order to be able to better differentiate the individual classes, it is in some cases advantageous if the data records are processed before classification. The methods used for data conditioning are presented here: One possibility to process the data sets consists of the differencing and absolute value formation of subsequent values according to (1)

$$X_{dif} = |X_i - X_{i+1}| \quad (1)$$

X_i and X_{i+1} are the successive sensor values. Another processing method is the integration of the data according to (2). Here the area under two successive values X_i and X_{i+1} is calculated.

$$X_{int} = \begin{cases} X_i + 0.5 * (X_{i+1} - X_i) & \text{if } X_i < X_{i+1} \\ X_i - 0.5 * (X_i - X_{i+1}) & \text{if } X_i > X_{i+1} \\ X_i & \text{if } X_i = X_{i+1} \end{cases} \quad (2)$$

In both the processing methods, an additional smoothing can be carried out by calculating the moving average according to equation (3).

$$X_i = \frac{1}{G} * \sum_{i-g}^{i+g} X_i \quad (3)$$

The parameter G specifies the degree of smoothing. The parameter g is the difference between the indices between the instantaneous value X_i and the maximum value $X_{g\pm i}$ specified by the degree of smoothing. Thus g depends on the degree of smoothing G and can be determined according to (4).

$$g = \frac{G-1}{2} \quad (4)$$

With the degree of smoothing G, first optimizations regarding the classification of the measured sensor data can be carried out. With the degree of smoothing G, first optimizations regarding the classification of the measured sensor data can be carried out [7]. Figure 2 shows the effect of processing by means of differencing compared to the unprocessed raw data. The coloring in the pictures illustrates the different class assignments. Figure 2 (a) shows the acquired raw data of an acceleration sensor in the x-orientation. It can be seen that a delimitation regarding the classes is not clear. For example, the acceleration in the direction of the x-axis at -8 m/s^2 is not unique and can in principle be assigned to any class. In Figure 2 (b) the classes are more delimited after the differencing and smoothing and thus a class assignment is clearer. For example, the value 0.7 can be clearly assigned to the class shown in gray.

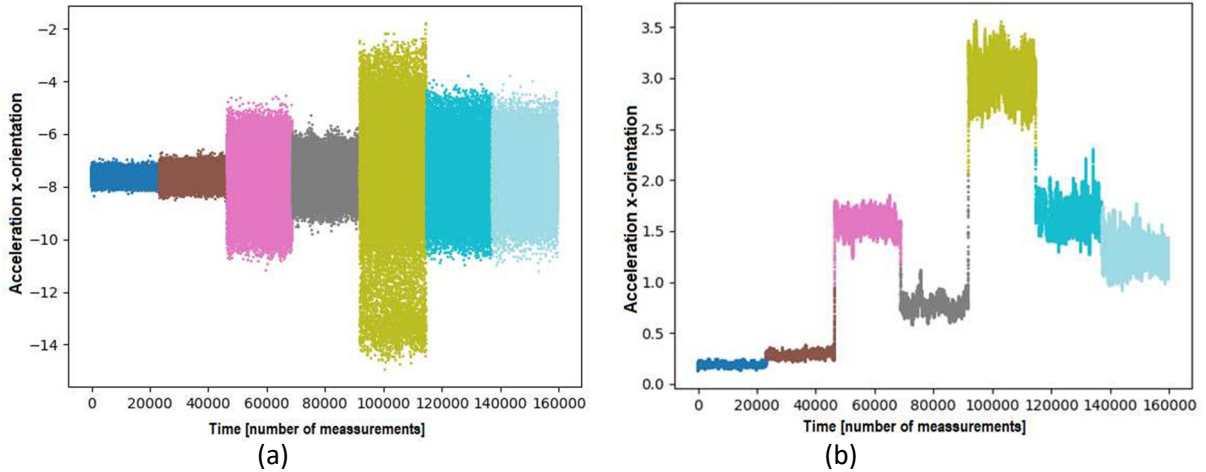


Figure 2: Comparison of raw data (a) and data prepared by differencing (b).

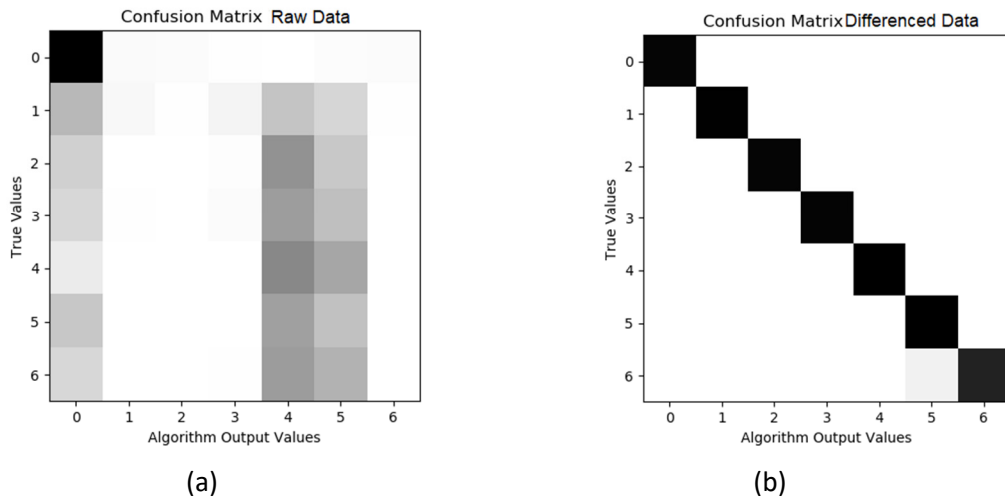


Figure 3: Comparison of the prediction results of the focus cluster algorithm with raw data (a) and data prepared by differencing (b).

5. Evaluation

The evaluation of the ML algorithms with regard to the respective sensors and the data processing was divided into a training and a test phase. In the training phase, the data records were divided evenly by feeding every tenth data value of the respective training method to the ML algorithm. As a result, the respective ML algorithm was trained with 10% of the data. The complete data set was then evaluated in the test phase. Several ML algorithms were considered for the recorded data sets. These included decision trees [8] [9], the gradient boost method [10] [8] a focus cluster algorithm [11] and artificial neural networks (ANN) [12]. The investigations revealed that ANNs are less suitable for data sets with low attribute numbers due to the long duration in the training phase. Therefore, only the decision trees, the gradient boost method and the focus cluster algorithm were used for the further experiments [7]. Figure shows two confusion matrices [13] for the focus cluster algorithm, which show the distribution between the actual class and the class determined by the algorithm. The numbers

on the axes correspond to the seven defined classes in which the data records have been categorized. The darker an area, the more often the ML algorithm has assigned data records to a class. A correct assignment is obtained if the assigned class corresponds to the actual class. Ideally, you would get a black diagonal from top left to bottom right. Figure (b) shows the result for the data sets prepared after differencing and smoothing. It can be seen that the majority of the data records were assigned to the actual classes, the hit rate here was over 98%. On the other hand, it can be seen in Figure (a) that a significantly lower hit rate has been achieved for the unprepared data sets.

6. Results

To compare the results of the ML algorithms for the respective sensors, a matrix with the relevant properties was created for each combination of ML algorithm and sensor:

- Classification accuracy (performance) of the algorithms

- Computing time for the training and testing phase of the algorithms
- Smoothing factor G

Table 1 shows the comparison of the processing methods examined using the combination of the multifunction sensor GY 521 and the gradient boost method.

Table 1: Comparison of the data preparation methods with the combination of the multifunction sensor GY521 and the gradient boost method.

| GY521 with Gradient Boost method | Performance (%) | | Computing Time Testing (s) | | Computing Time Testing (s) | |
|----------------------------------|-----------------|-------|----------------------------|------|----------------------------|------|
| | G=0 | G=99 | G=0 | G=99 | G=0 | G=99 |
| Raw Data | 98.51 | 83.97 | 4.85 | 6.88 | 1.11 | 0.90 |
| Integration | 82.84 | 85.15 | 5.45 | 3.82 | 1.07 | 0.66 |
| Differencing | 79.76 | 85.1 | 5.04 | 3.83 | 1.50 | 0.61 |

It can be seen that the highest performance is achieved when using the raw data. In the training and test phases the integrated and the differenced data are slightly faster. A comparison was made for each sensor and ML algorithm combination. The best performing data processing method was then selected for each combination. The comparison tables of the sensors which have achieved the best results of the sensor types examined are listed in Tables 2 to 4. These were the ADXL 345 (accelerometer), the MPU 60.50 (gyroscope) and the GY 521 (accelerometer, gyroscope and magnetic field). The processing method with the highest performance for the respective algorithm is shown for each of the sensors.

Table 2: Comparison of the ML algorithms on the combination ADXL 345 (accelerometer) and the best performing data processing method.

| ADXL345 | Decision Tree | Gradient Boost Method | Focus Cluster Algorithm |
|-------------------------|-------------------|-----------------------|-------------------------|
| Conditioning method | Differencing G=99 | Differencing G=99 | Differencing G=99 |
| Performance | 93.79% | 96.6% | 98.98% |
| Computing time training | 1.78s | 5.88s | 0.32s |
| Computing time testing | 31.99s | 1.44s | 34.88s |

The result of the examinations according to Table 2 is that all acceleration sensors achieved the greatest performance with smoothing (G = 99) and data prepared by differencing. The focus cluster algorithm achieved the highest performance.

Table 3: Comparison of the ML algorithms on the combination MPU 60.50 (gyroskope) and the best performing data processing method.

| MPU60.50 | Decision Tree | Gradient Boost Method | Focus Cluster Algorithm |
|---------------------|---------------|-----------------------|-------------------------|
| Conditioning method | Raw data, G=0 | Raw data, G=0 | Differencing G=99 |

| | | | |
|-------------------------|--------|--------|--------|
| Performance | 91.46% | 95.24% | 90.89% |
| Computing time training | 0.09s | 5.54s | 0.91s |
| Computing time testing | 0.28s | 1.46s | 31.99s |

The result according to Table 3 is that the highest performance was achieved with unprocessed and unsmoothed data with the gyroscopes. The gradient boost process achieved the highest performance.

Table 4: Comparison of the ML algorithms on the combination GY 521 and the best performing data processing method.

| GY521 | Decision Tree | Gradient Boost Method | Focus Cluster Algorithm |
|-------------------------|---------------|-----------------------|-------------------------|
| Conditioning method | Raw data, G=0 | Raw data, G=0 | Differencing G=99 |
| Performance | 95.7% | 98.51% | 99.89% |
| Computing time training | 0.09s | 4.85s | 1.27s |
| Computing time testing | 0.28s | 1.11s | 37.63s |

In the case of the multifunction sensors with acceleration, magnetic field sensors and gyroscope, it can be seen from Table 4 **Fehler! Verweisquelle konnte nicht gefunden werden.** that the highest performance was achieved with smoothing (G = 99) and differenced data using the focus cluster algorithm.

7. Conclusion

It has been found that the processing of the raw data in the form of smoothing and differencing in combination with the focus cluster algorithm gave the best results for acceleration sensors. The gyroscopes examined showed that the unprocessed raw data without smoothing in combination with the gradient boost method achieved the highest classifiability. The multisensors examined gave the best results when using the focus cluster algorithm in combination with smoothed and differenced data.

8. Outlook

Based on these results, the combination of detection measure, data processing method and ML algorithm can in the next step be used for a PMA strategy. For a complete PMA, further detection measures have to be examined. For that purpose, this procedure is continued with further sensor types in order to find an optimal combination for all necessary detection measures. In the future, a prediction model is to be developed on the basis of these results, with which predictions can be made about the degree of wear of system components of a thermal system under automation. Formal aging and error models of the respective system components must also be created in order to map the aging process of components. These models can then be used to make probabilistic statements about the failure probabilities

of the individual assemblies. Such models could be based on Dynamic Bayesian Networks (DBN) [11] auto regression and moving average (ARMA) [14] or, as the focus cluster algorithm has yielded such a high performance, a multi dimensional focus trajectory.

References

- [1] O. H. H. R. a. P.-M. S. S. Feldmann, Predictive Maintenance - Service der Zukunft, Roland Berger GmbH, 2017.
- [2] „DIN EN 13306: Instandhaltung – Begriffe der Instandhaltung; Dreisprachige Fassung EN 13306:2017 (Deutsche Norm),“ Beuth Verlag, Berlin, 2018.
- [3] B. Ebert, Prozessoptimierung bei Industrie 4.0 durch Risikoanalysen, Berlin, Heidelberg: Springer-Verlag, 2018.
- [4] R. B. S. G. Frank Duscheck, „<https://www.bearingpoint.com/>,“ BearingPoint GmbH, 18 01 2018. [Online]. Available: https://www.bearingpoint.com/files/BearingPoint_Studie_Maintenance_.pdf. [Zugriff am 07 12 2019].
- [5] H. M. Hashemian, „State-of-the-Art Predictive Maintenance Techniques,“ in IEEE Transactions On Instrumentation And Measurement Vol. 60, 2011.
- [6] M. Küstner, „Auswahl und Bewertung eines Sensorkonzepts für die Implementierung einer Predictive Maintenance zur Temperaturbehandlung unter Automation,“ Weikersheim, Künzelsau, 2019.
- [7] M. Gaust, „Entwurf und Implementierung von maschinellen Lernalgorithmen zur Klassifikation von Maschinendaten für eine vorausschauende Wartung von Industrie-Thermosystemen,“ Weikersheim, Künzelsau, 2019.
- [8] Pedregosa, „JMLR 12 Scikit-learn: Machine Learning in Python,“ Journal of Machine Learning Research, 2011.
- [9] P. V. A. E. S. Thomas T., Applications of Decision Trees. In: Machine Learning Approaches in Cyber Security Analytics, Singapore: Springer, 2020.
- [10] N. D. a. S. P. A. Binding, „Machine Learning Predictive Maintenance on Data in the Wild,“ in IEEE 5th World Forum on Internet of Things (WF-IoT), Limerick, Ireland, IEEE, 2019, pp. 507-512.
- [11] A. Géron, Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques for Building Intelligent Systems, Boston: O'Reilly UK Ltd., 2017.
- [12] T. Rashid, Neuronale Netze selbst programmieren: Ein verständlicher Einstieg mit Python, Boston: O'Reilly, 2017.
- [13] M. Füllsack, „Systems sciences at ISIS,“ Institute for Systems Sciences, Innovation and Sustainability Research at the Karl-Franzens University Graz, Austria, 2013. [Online]. Available: <http://systems-sciences.uni-graz.at/etextbook/bigdata/confusion-matrix.html>. [Zugriff am 28 11 2019].
- [14] Y. L. Y. Z. R. X. a. F. Z. Jinjiang Wang, „An integrated fault diagnosis and prognosis approach for predictive maintenance of wind turbine bearing with limited samples,“ in Renewable Energy Vol. 145, 2020, pp. 642 - 650.

Edge detection in Images with multiplicative Noise by using Ant Colony System

Sergio Baltierra⁽¹⁾, Marco Mora⁽²⁾

(1)Facultad de Ingeniería, Universidad Autónoma de Chile, 346 0000

(2)Department of Computer Science, Universidad Católica del Maule, 346 6706

Abstract

We propose a method for edge detection in images with multiplicative noise based on Ant Colony System (ACS). To adapt the Ant Colony System algorithm to multiplicative noise, global pheromone matrix is computed by the Coefficient of Variation. We carried out a performance comparison of the edge detection Ant Colony System algorithm among several techniques, the best results were found in the gradient and the coefficient of variation.

1. Introduction

The extraction of features in images with noise is one problem difficult to solve, above all if the noise present is the noise multiplicative, one of the most aggressive [1] and [6].

The algorithms that currently used as the Coefficient of Variation (CV) [2] and [7] do not give good solutions, so we must pay attention to other algorithms.

One of these algorithms is the algorithms bioinspired, which are adaptations of the behavior, reasoning, and solution of problems of systems biological present in nature. A bioinspired algorithm is Ant Colony System (ACS) [3].

In this paper, an adaptation of the ACS algorithm is proposed to detect edges in images with multiplicative noise at various levels, also some results of experiments on a synthetic image are also presented. In the end, it is analysed qualitatively, with visual observation, and quantitatively using a performance function [5].

2. The Ant Colony System

The ACS was proposed in one principle as a solution to the traveling salesman problem (TSP) [3]. The instructions of ACS can be summarized below:

- Determine the number of ants "m" in the colony, and the number of iterations of the algorithm "n".
- For each iteration and for each one of the ants of the colony, is made following sub instructions:
 - Apply the State Transition Rule.
 - Store the visited city in a list called a Taboo List.
 - Apply the Pheromone Local Update Rule.
- Apply the Pheromone Global Update Rule.
- At end of the last iteration will you get the route more cuts.

3. Adaptation of Ant Colony Algorithm for Edge Detection

To be able to detect contours using ACS first, we must define the type of neighborhood, which for our experiments is the neighborhood of Von Neumann as is shown below.

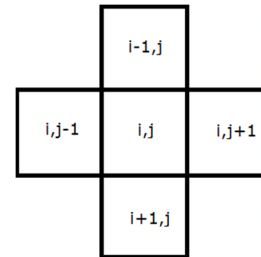


Figure 1: Neighborhood Von Neumann.

Also, must make some modifications to be able to adapt ACS for the detection of contours, these changes are:

- The edge detection in images is preserved the States Transition Rule, but it modifies the function of visibility.
- It preserves the Pheromone Local Update Rule.
- In the Pheromone Global Update Rule, only it replaces part of the expression with one constant [4].

In the end, the result to be obtained in the processing of an image using ACS is the path of the ant colony that is stored in a matrix called the Pheromone Matrix, which when thresholding we obtain an estimate of the image edge.

Two types of pheromone are defined in ACS: the first is the Global Pheromone (GF) and the second is the Local Pheromone (LF). FG is defined by one equation, the second is calculated to through of the path of the ants.

For our solution, we will use more than one FG, since we want to obtain the best contours of an image with multiplicative noise.

The first GF used is the gradient [4], which is:

$$|G(I)| = \sqrt{\frac{\delta I^2}{\delta x} + \frac{\delta I^2}{\delta y}}$$

Where $\frac{\delta I}{\delta x}$ is the partial derivate of the image I with respect to x and $\frac{\delta I}{\delta y}$ is the partial derivate of the image I with respect to y .

It should be mentioned that to obtain the result more optimal, they are used four types of gradient denominated: Gradient A (GA), Gradient B (GB), Gradient C (GC), and Gradient D (GD).

The second GF to use is the CV, which is expressed below:

$$CV(I) = \frac{\sigma}{|\bar{x}|}$$

Where σ is the standard deviation, and $|\bar{x}|$ is the absolute value of the average. Both are computed on the neighborhood defined in figure 1 of the image I to be processed.

4. Experiments and Results

In order to carry out the experiments, we will use a synthetic image and its terrain of truth, where the first is an image built from a tool, free of noise and the second is the true contours of the image. The image synthetic is contaminated to with noise multiplicative in three levels: level low of noise, level medium noise, and a level high of noise, as is shown below:

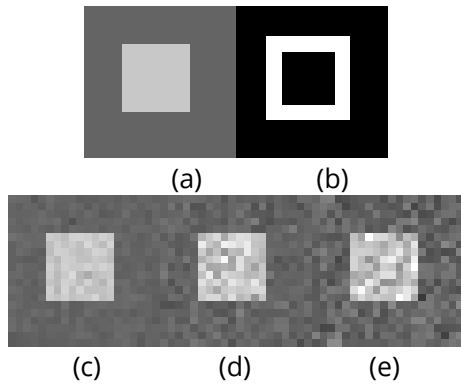


Figure 2: Simple Synthetic Image: (a) Image Original, (b) Ground of Truth, (c) Level low of noise ($\sigma = 0,05$), (d) Level medium of noise ($\sigma = 0,1$) and (e) Level high of noise ($\sigma = 0,15$).

In the experiments, the ACS parameters proposed in [3] are declared. The detail of these parameters is the following: the number of ants “ m ” is considered one percent of the area of the image, and the number of iterations “ n ” is the maximum between the width and length of the image. The values of the parameters remaining are $\tau_{ini} = 1, \alpha = 1, \beta = 1, \rho = 0.3, \gamma = 0.5$.

In Figures 3, 4, 5, 6, and 7 are shown the contours detected, where: Figure 3 GF of the GA (GF-GA), Figure 4 GF GB (GF-GB), Figure 5 GF of the GC (GF- GC), figure 6 GF of the GD (GF-GD) and figure 7 GF of the CV (GF-CV).

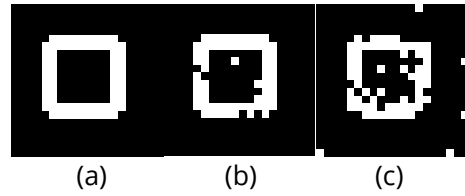


Figure 3: Edge detected in one image synthetic: (a) Level low of noise GF-GA, (b) Level medium of noise GF-GA, (c) level high of noise GF-GA.

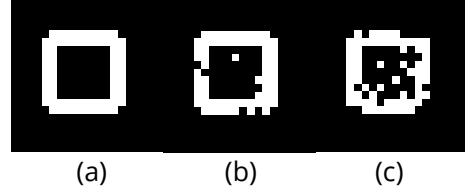


Figure 4: Edge detected in one image synthetic: (a) Level low of noise GF-GB, (b) Level medium of noise GF-GB, (c) level high of noise GF-GB.

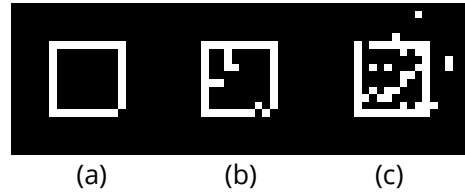


Figure 5: Edge detected in one image synthetic: (a) Level low of noise GF-GC, (b) Level medium of noise GF-GC, (c) level high of noise GF-GC.

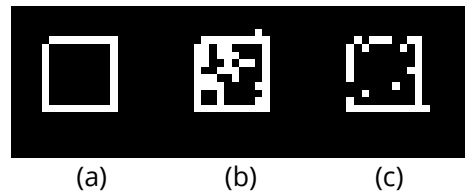


Figure 6: Edge detected in one image Synthetic: (a) Level low of noise GF-GD, (b) Level medium of noise GF-GD, (c) level high of noise GF-GD.

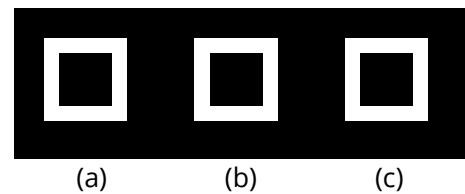


Figure 7: Edge detected in one image synthetic: (a) Level low of noise GF-CV, (b) Level medium of noise GF-CV, (c) level high of noise GF-CV.

As can be seen, as noise increases there are more errors in contour detection. Notwithstanding the foregoing, in GF-CV fewer errors are visually obtained.

To quantify this analysis, we will use a performance function defined in [5] whose results are shown in the following tables:

| Level low of noise | | | | | |
|---------------------|-------|-------|-------|-------|-------|
| | GF-GA | GF-GB | GF-GC | GF-GD | GF-CV |
| Optimal performance | 0.95 | 0.95 | 0.48 | 0.48 | 1 |
| Optimal threshold | 0.53 | 0.53 | 0.42 | 0.46 | 0.56 |

Table 1: Optimal values of the performance function for level low of noise.

| Level medium of noise | | | | | |
|-----------------------|-------|-------|-------|-------|-------|
| | GF-GA | GF-GB | GF-GC | GF-GD | GF-CV |
| Optimal performance | 0.83 | 0.85 | 0.44 | 0.47 | 1 |
| Optimal threshold | 0.52 | 0.54 | 0.58 | 0.44 | 0.64 |

Table 2: Optimal values of the performance function for the level medium of noise.

| Level high of noise | | | | | |
|---------------------|-------|-------|-------|-------|-------|
| | GF-GA | GF-GB | GF-GC | GF-GD | GF-CV |
| Optimal performance | 0.66 | 0.7 | 0.45 | 0.44 | 1 |
| Optimal threshold | 0.51 | 0.5 | 0.41 | 0.56 | 0.68 |

Table 3: Optimal values of the performance function for the level high of noise.

The results of the tables confirm the seen in qualitative analysis, that the GF-CV has better results in comparison to GF with the four types of gradients.

5. Conclusion and Future Works

It has proposed one method of edge detection in images with noise multiplicative using the algorithm ACS.

In this proposal a comparison of the Global Pheromone is considered based on gradient of the image and based on CV of the image.

Image synthetic allow compare easily such the detection of contours using gradient and CV. The analysis of the results allows to conclude that the ACS algorithm based on GF-CV adapts and detects contours in a better way in an image with multiplicative noise.

It is proposed as future works:

- Perform tests on real images with ground of truth.
- Propose other detector of contours as GF and make the comparisons between is a new GF and already analysed.
- Extend the algorithm proposed for other types of noises like for example the noise Salt and Pepper.

Acknowledgements

Special thanks to Facultad de Ingeniería of the Universidad Autónoma de Chile by depositing in me the confidence to present my area of research.

References

- [1] R. Touzi, A review of speckle filtering in the context of estimation theory. *IEEE Transactions on Geoscience and Remote Sensing*, 40 (2002), 2392–2404.
- [2] D. Kuan, A. Sawchuk, T. Strand, P. Chavel, Adaptive restoration of images with speckle, *IEEE Transactions on Acoustics Speech and Signal Processing*, 35 (1987), 3, 373–383.
- [3] M. Dorigo, L. M. Gambardella, Ant colony system: A cooperative learning approach to the traveling salesman problem, *IEEE Transactions on Evolutionary Computation*, 1 (1997), 1, 53–66.
- [4] S. A. Etemad, T. White, An ant-inspired algorithm for detection of image edge features, *Applied Soft Computing*, 11 (2011), 8, 4883–4893.
- [5] C. Grigorescu, N. Petkov, M. Westenberg, Contour detection based on nonclassical receptive field inhibition, *IEEE Transactions on Image Processing*, 12 (2003), 7, 729–739.
- [6] V. Bhateja, A. Tripathi, A. Gupta, A. Lay-Ekuakille, Speckle suppression in SAR images employing modified anisotropic diffusion filtering in wavelet domain for environment monitoring, *Measurement: Journal of the International Measurement Confederation* 74 (2015), 246–254.
- [7] I. F. Ince, O. F. Ince, F. Bulut, MID filter: An orientation-based nonlinear filter for reducing multiplicative noise. *Electronics* 8 (2019), 9.

The challenge of using artificial intelligence in autonomous driving

Michael Sachse, André Müller, Klaus Krumbiegel
 Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

Durch die steigende Leistungsfähigkeit von Prozessoren und Datenübertragungstechniken hat die Entwicklung und Anwendung von künstlicher Intelligenz, exemplarisch das maschinelle Lernen (engl. Machine Learning – ML) und die Methode des Deep Learning, in den letzten Jahren stark an Bedeutung gewonnen. Hierbei stellt sich die Frage, wie diese Technologien in einem weiteren zukunftssträchtigen Entwicklungsfeld, zum Beispiel bei der Entwicklung moderner Mobilitätskonzepte und hochautomatisierter/autonomer Fahrzeuge, eingesetzt werden können.

Potentielle Möglichkeiten der Anwendung von AI im Entwicklungsprozess eines hochautomatisierten Fahrzeugs werden vorgestellt, aber auch die entscheidenden Herausforderungen diskutiert. Darüber hinaus wird der Unterschied zwischen verschiedenen Ansätzen ausgeführt. Dazu werden sowohl Randbedingungen als auch Herausforderungen mit Hilfe eines einfachen Beispiels aus dem täglichen Verkehrsgeschehen veranschaulicht.

1. Einleitung

Der Einsatz von Künstlicher Intelligenz ist einer der aktuellen Trends und Forschungsschwerpunkte verschiedenster Domänen in Wirtschaft und Forschung. Dabei wird der Begriff oft sehr weit gefasst – von eher klassischen intelligent erscheinenden Algorithmen und statistischen Methoden, bis hin zu selbstlernenden Systemen, etwa auf Basis von künstlichen neuronalen Netzen.

Die Verwendung von künstlichen Intelligenzen stellt die Automotive Industrie nicht nur technologisch, sondern auch methodisch vor Herausforderungen. Aktuelle Entwicklungsprozesse, z.B. basierend auf Automotive Spice, benötigen Anforderungen als Grundlage für die Entwicklung von Software. Der Einsatz von Neuronalen Netzen mit sichtbaren Schichten steht diesem entgegen. Diese Herausforderung wird auch im Rahmen des SOTIF-Normierungsvorhabens diskutiert [1]. Resultierend sollte ein besonderes Augenmerk auf die Spezifikation der gewünschten Funktionalität in Kombination mit Anwendungsfällen und Szenarien gelegt werden. Die relevanten Szenarien für die Anwendungsfälle und die Funktion werden in Kategorien eingeteilt, welche in Abbildung 1 dargestellt sind. Das Ziel ist es unbekannte und gefährdende Szenarien zu minimieren.

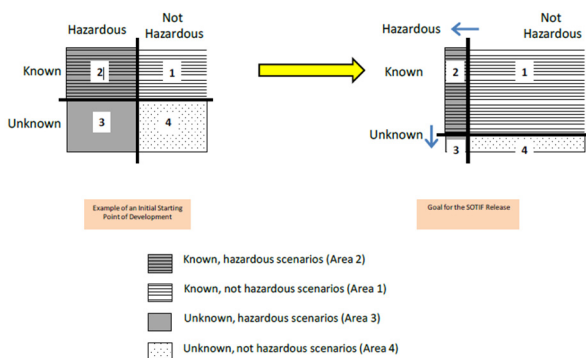


Abbildung 1: Scenario Classification and SOTIF-Goal [1]

2. Entscheidungskette an einem Beispiel

Der Informationsfluss im Fahrzeug kann angelehnt an die Informationsverarbeitung des Menschen in verschiedene Phasen unterteilt werden. Die Wahrnehmung der Umgebung erfolgt mit Hilfe von Sensoren oder aber auch a priori Wissen (z.B. Karten). Die daraus resultierenden Daten müssen anschließend verarbeitet und mögliche Situationen oder Szenarien antizipiert werden. Basierend auf diesen Informationen kommt es zu einer Entscheidung der Fahrstrategie und folgender Umsetzung.

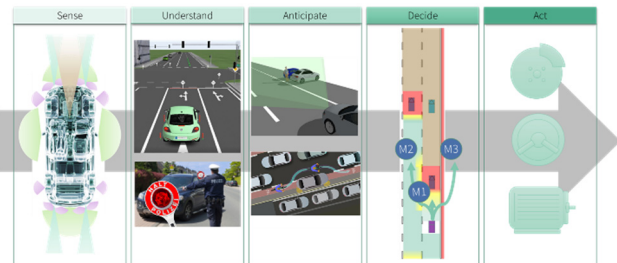


Abbildung 2: Verarbeitungskette im Fahrzeug

In der Sensorik werden schon seit einigen Jahren Mechanismen der künstlichen Intelligenz angewendet. Insbesondere im Bereich der kamerabasierten Umfelderkennung erfolgt mit Hilfe von Machine Learning die Detektion und Klassifikation von Objekten wie Fahrzeugen, LKWs und Verkehrszeichen. In der Analyse und Interpretation der aktuellen Situation ist die Nutzung von Machine Learning ebenfalls möglich oder sogar in spezifischen Problemen erforderlich.

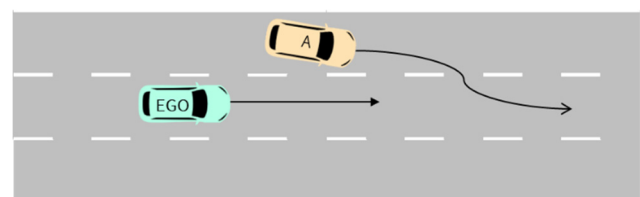


Abbildung 3: Einscherserszenario

Eine nahezu alltägliche Situation stellt das Einscheren eines anderen Fahrzeugs im Straßenverkehr dar. Für den heutigen Stand der Technik beinhaltet dieses Szenario eine Vielzahl von Herausforderungen. So lässt sich das genannte Szenario mit über 20 Parametern abbilden, welche untereinander Abhängigkeiten haben und deren Werte lediglich einen konkreten Zeitpunkt beschreiben und nur bedingt zeitliche Abhängigkeiten repräsentieren. Darüber hinaus ist die weitere Entwicklung des Szenarios nicht eindeutig, selbst wenn physikalische Grenzen betrachtet werden. Tabelle 1 zeigt eine Auswahl an möglichen Parametern.

| | |
|------------|--|
| x_A | Abstand in X-Richtung Ego zu Fahrzeug A |
| y_A | Abstand in Y-Richtung Ego zu Fahrzeug A |
| vx_A | Geschwindigkeit in X-Richtung von Fahrzeug A |
| vy_A | Geschwindigkeit in Y-Richtung von Fahrzeug A |
| ax_A | Beschleunigung in X-Richtung von Fahrzeug A |
| ay_A | Beschleunigung in Y-Richtung von Fahrzeug A |
| α_A | Kurswinkel von Fahrzeug A |
| L_t | Typ der Ego Fahrspur |
| L_w | Breite der Ego Fahrspur |
| | Etc... |

Tabelle 1: Mögliche Parameter zur Szenariobeschreibung

Bei der Prozessierung des Szenarios sind in der Regel mehrere Softwaremodule beteiligt. Die Funktionalität soll dafür über klare deterministische Regeln und Modelle als Anforderungen beschrieben und implementiert werden. Dafür wird eine Vielzahl von Parametern eingeführt, welche voneinander abhängen und die Funktionalität beeinflussen. Klassische Beispiele für solche Parameter sind Schwellwerte für Entscheidungen, die Trajektorienberechnung und -auswahl oder Regelungsparameter. In einem autonomen Fahrzeug kann nach heutigem Stand der Technik von mehreren tausend bis zehntausenden Parametern ausgegangen werden.

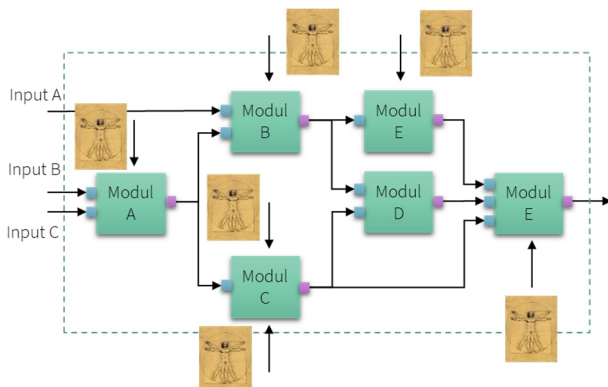


Abbildung 4: SW-Architekturschema einer Entscheidungskette

Die manuelle Kalibrierung dieser Anzahl an Parametern, insbesondere in Abhängigkeit von den Szenario-Parametern, ist zeitlich sehr aufwendig und somit nicht praktikabel. Hinzu kommt, dass in der Regel die Lösungen und Funktionen eine Abstraktion der Realität darstellen und so unter anderem, nicht lineare Zusammenhänge in meist linearisierten Regeln abgebildet werden. Die manuelle Einstellung eines Parameters bezüglich der Verbesserung eines Szenarios kann zu einer Verschlechterung eines anderen Szenarios führen, was für den menschlichen Entwickler auf Grund der Komplexität nicht vorhersehbar ist.

Eine weitere Herausforderung stellen die zumeist gegensätzlichen Ziele aus den Bereichen Komfort und Sicherheit dar. So kann eine Fahrfunktion mit einer gewählten Parametrierung sicher ausgelegt sein, widerspricht aber gewissen Anforderungen an den Fahrkomfort. Weiterhin muss die Gesamtheit von auftretenden oder zu beherrschenden Fahrsituationen betrachtet werden, denn eine sichere Auslegung einer Fahrfunktion in Bezug auf eine Einzelsituation kann das Risiko in einer anderen Fahrsituation erhöhen.

3. Kalibrierung der Entscheidungskette

Wie bereits erwähnt, ist die Kalibrierung von mehreren Parametern auf Grund der gegenseitigen Beeinflussung über eine Entscheidungskette manuell nur schwer möglich. Aus diesem Grund kann ein Parameteroptimierungsablauf bezüglich der Entscheidungskette entwickelt werden. Dieser Workflow ist grob in Abbildung 2 dargestellt und unterstützt den Ansatz der szenarienbasierten Funktionsentwicklung.

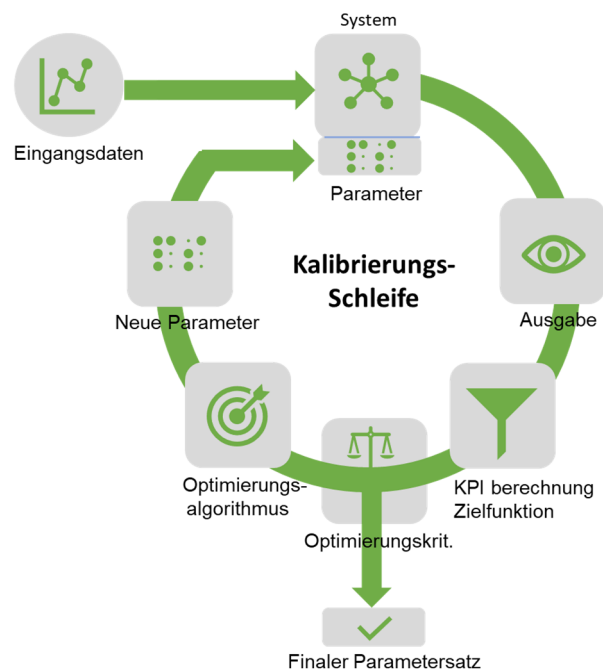


Abbildung 5: Optimierungsworkflow

Die Szenarien mit ihren vielfältigen Parametern stellen die Eingangsdaten dar. Die Parameter der im System beteiligten SW-Module sollen nun mit Hilfe von Optimierungsschleifen kalibriert werden. Die verschiedenen Kriterien aus den Bereichen Komfort und Safety können in eine Zielfunktion oder Nebenbedingung für die Optimierung einfließen. Bezogen auf das Einschereerbeispiel können dies zum Beispiel eine Kritikalitätsmetrik, die Verletzung eines definierten Sicherheitsabstands und die Verzögerung des Egofahrzeugs sein.

Die Wahl der Szenarioparameter als Input für die Kalibrierungsschleife ist von entscheidender Bedeutung, um die Bereiche Komfort und Sicherheit angemessen abzudecken. Eine Optimierung ausschließlich basierend auf kritischen Szenarien ist demzufolge nicht zielführend. Wiederum können sich durch eine vollfaktorielle Zerlegung des Parameterraumes von mehr als 20 Parametern wie im oben beschriebenen Einschereerszenario schnell mehrere Millionen Szenarien ergeben.

Um dem Ressourcen- und Zeitaufwand schließlich gerecht zu werden, müssen Samplingmethoden des Design of Experiment (DoE) zur effizienten Abdeckung des Szenariotesttraums mit dem Parameteroptimierungsworkflow verknüpft werden.

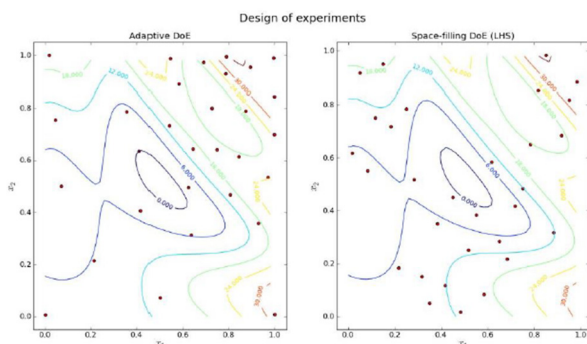


Abbildung 6: 2D-Beispielgrafik DoE-Sampling Methoden

Eine weitere wesentliche Herausforderung ist die Wahl des Optimierungsverfahrens für die Parameter. Auf Grund der Komplexität und nichtlinearer Zusammenhänge kann der Einsatz klassischer Algorithmen an seine Grenzen stoßen und möglicherweise zu keinem verwertbaren Ergebnis führen. Hier ergibt sich die Möglichkeit KI-basierte Methoden einzusetzen, auch wenn das System selbst noch ein regelbasiertes ist, da es auf Anforderungen beruht. So könnte z.B. ein nichtlineares ML Modell wie beispielsweise eine Support Vector Machine (SVM) eingesetzt werden, um einen funktionalen Zusammenhang zwischen den Stellparametern und der Zielfunktion herzustellen. Die SVM transformiert den Merkmalsraum mittels Kernel-Funktion in einen dualen Raum, in dem mittels Hyperebenen die Datenpunkte getrennt werden. Durch Rücktransformation wird eine nichtlineare Trennung im ursprünglichen Raum erzeugt. Für den Optimierungsschritt wird die resultierende Funktion schließlich optimiert, z.B. mittels eines genetischen Algorithmus, der als Werkzeug im Rahmen von KI-

Anwendungen bereits erfolgsversprechende Resultate erzielt hat (vgl. [2]). Weitere Optimierungsverfahren, welche an dieser Stelle zum Einsatz kommen können, sind das Nelder-Mead-Verfahren, Partikelschwarmoptimierung und diverse Greedy Algorithmen.

4. Validierung der Entscheidungskette mit KI

Wie bereits beschrieben, können KI-basierte Verfahren an unterschiedlichen Stellen der Verarbeitungskette autonomer Fahrfunktionen zum Einsatz kommen. Neben der Verwendung in der Sensorik, wurde bereits die Verwendung im Bereich der Softwareentwicklung, z.B. bei der Kalibrierung der SW-Module erwähnt. Wieviel Potenzial sich hinter dem Thema tatsächlich verbirgt, wird deutlich, wenn man sich einem weiteren wesentlichen Bestandteil der Entwicklungskette autonomer Fahrfunktionalitäten widmet, dem komplexen Thema der Validierung von Funktionskomponenten.

Tatsächlich birgt der Bereich Validierung große Herausforderungen, da immer mehr autonome Fahrfunktionalitäten in der Simulation abgetestet werden müssen. In diesem Zusammenhang stellt sich die Frage, welche Szenarien konkret getestet werden müssen. Neben den Szenarien, die für die Kalibrierung verwendet wurden, sollten weitere kritische Szenarien identifiziert und getestet werden. Das SOTIF-Normierungsvorhaben sieht vor, bei der Entwicklung von autonomen Fahrfunktionalitäten die Anzahl der unbekannt und gefährdenden Szenarien zu minimieren, jedoch wird sowohl die Auswahl mittels eines vollfaktoriellen Ansatzes, als auch die Auswahl durch geeignete DoE Methoden nicht ausreichen, diese vollständig zu identifizieren.

Neben den Samplingmethoden des DoE, die eine möglichst effiziente Abdeckung des Parameterraumes anstreben und so auch kritische Szenarien identifizieren können, ist ein weiterer Ansatz für die Erkennung der relevanten Szenarien zielführend: Reinforcement Learning gesteuertes Testen der autonomen Fahrfunktionalitäten.



Abbildung 7: Validierungsworkflow mit KI-basiertem Agenten zur Szenariowahl

Im Rahmen dieser Methodik entscheidet ein sogenannter Agent über die konkreten Szenarien, die nacheinander getestet werden. Das Ziel des Agenten ist dabei, die einzelnen Parameter für das Szenario (also z.B. Geschwindigkeit in X-Richtung des Fahrzeugs A) so zu wählen, dass die implementierte Funktionalität herausgefordert wird und, wenn möglich, Grenzen für sicherheitsrelevante Kenngrößen verletzt werden. Das Verhalten des Agenten wird dabei über eine Belohnungsfunktion erzwungen, welche der Zielfunktion aus der Kalibrierung entsprechen kann, wobei hier das umgekehrte Optimierungsproblem gelöst wird. Je höher der Wert, desto sicherheitskritischer ist die Situation und desto größer die Belohnung für den Agenten. Der Agent wählt unter Berücksichtigung einer definierten Explore-Rate eine zufällige Kombination von Parameterwerten. Mit einer festgelegten Schrittweite für eine Parameteranpassung werden die für die Fahrfunktionalität kritischsten Szenarien im Laufe der Simulationen gezielt identifiziert und getestet ohne sämtliche Parameterkombinationen testen zu müssen. Das könnte den Simulationsaufwand erheblich reduzieren.

5. Realisierung der Entscheidungskette durch KI

Neben der Unterstützung der Entwicklung von autonomen Fahrfunktionen stellt sich die Frage, inwieweit KI auch Bestandteil der eigentlichen Funktion sein kann. D.h. die regelbasierten SW-Module mit einer großen Anzahl an Kalibrierungsparametern werden durch KI-basierte Module ersetzt. Wie bereits erwähnt, ist dies im Falle der kamerabasierten Umfelderkennung bereits State of the Art. Hier sind die einfachen klassischen Bildverarbeitungsalgorithmen hinsichtlich Prognosegüte weit unterlegen, so dass sich die Frage nach der Zulässigkeit der Verwendung der KI mittlerweile erübrigt.

Zusätzlich zum Bereich der Sensorik sind sämtliche Prozesse der Planung für den Einsatz von KI aus technischer Sicht potenziell interessant. Ein Beispiel findet sich in der Pfadplanung in Form der Berechnung der Trajektorie auf Basis von Umgebungsinformationen. Die Trajektorie gibt vor, wo sich das Fahrzeug während der folgenden Zeitschritte jeweils befinden sollte. Ein Ansatz, eine solche Funktionalität zu erlernen, könnte beispielsweise über Prinzipien des Reinforcement Learning erfolgen, die schon in Abschnitt 4 vorgestellt wurden (siehe auch [3]). Für den konkreten Fall übernimmt der Agent die Funktion, wird im Rahmen von ausgewählten Szenarien getestet und erhält Feedback in Form einer Belohnungsfunktion, die unterschiedliche KPIs (Key Performance Indicator z.B. Kollision ja/nein oder Eintauchtiefe in den Sicherheitsabstand) berücksichtigt. Auf Basis dieses Feedbacks passt der Agent wiederum die Funktion an, also ändert z.B. die longitudinale Beschleunigung in Abhängigkeit der Szenarioparameter.

Aber auch die Verwendung von KI in der Funktion und den damit verbundenen Entwicklungs- und Validierungsprozessen bringt einige Herausforderungen mit sich. Laut [1] sind dies u.a.

- Unvollständige Trainingsdaten und somit nicht aussagekräftige Modelle.
- Ungünstige Wahl von Trainingsdaten und somit Erlernen falscher Korrelationen.
- Ungünstige Wahl der Testdaten und somit ungenügende Verifizierung der gelernten Modellparameter.
- Keine unmittelbare oder unzureichende Beeinflussung des Trainingsprozesses durch den Menschen.

Bezogen auf die Entwicklung autonomer Fahrfunktionen bedeutet dies insbesondere, dass auf einer repräsentativen Menge von Szenarien gelernt werden sollte, wobei sichergestellt werden muss, dass auch die relevante Menge der kritischen Situationen erfasst und vom Modell verstanden wird. Die geeignete Kombination und Menge von Realdaten und simulativen Daten ist entscheidend für die Qualität.

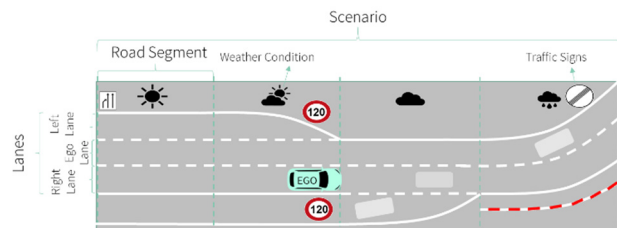


Abbildung 8: Visualisierung der Szenarienvielfalt

Allein durch die große Menge an Parametern bei der Definition eines Szenarios, Positionen, Geschwindigkeiten, Beschleunigungen, Wiedereinflüsse, Fahrbahnverlauf, etc. führt eine vollfaktorielle Zerlegung bei Berücksichtigung sämtlicher möglicher Parameterkombinationen schnell zu einer Menge an Szenarien, die von keinem Rechner dieser Welt in überschaubarer Zeit simuliert werden können. Ein simples Rechenbeispiel mit 200 Parametern und 10 Variationen führt beispielsweise zu 10^{200} Szenarien. Zum Vergleich kommt die Schätzung der Anzahl der Atome im Universum auf gerade mal 10^{96} . Diese Herausforderung löst auch die KI nicht.

Literaturverzeichnis

- [1] ISO/DIS 21448 - Road vehicles — Safety of the intended functionality
- [2] Jia, Weikuan & Zhao, Dean & Ding, Ling. (2016). An optimized RBF neural network algorithm based on partial least squares and genetic algorithm for classification of small sample. Applied Soft Computing. 48. 10.1016/j.asoc.2016.07.037.
- [3] Rose, Dominic & Mair, Jamie & Garrahan, Juan. (2021). A reinforcement learning approach to rare trajectory sampling. New J. Phys. 23 013013.

Transformation der Mobilität

–

***ökologisch, autonom und
vernetzt***

Autonomes Fahren – Traum der Ingenieure, Alptraum der Juristen?

Daniel Graewe

NORDAKADEMIE Hochschule der Wirtschaft, Köllner Chaussee 11, 25337 Elmshorn

Der Technologiesprung zum autonomen Fahren steht kurz bevor – und damit auch die Umsetzung im öffentlichen Straßenverkehr. Was technisch bereits in greifbare Nähe gerückt ist, stellt juristisch aber eine große Herausforderung dar. Wie soll autonomes Fahren reglementiert werden? Haftungs- und Strafbarkeitsfragen müssen vollkommen neu gedacht werden – und sind vermutlich nicht rechtssicher lösbar. Gleichzeitig gerät der deutsche Gesetzgeber unter internationalen Umsetzungsdruck.

The technological leap to autonomous driving is imminent - and so is its implementation in public road traffic. What is already technically within reach, however, presents a major legal challenge. How should autonomous driving be regulated? Questions of civil and criminal liability must be completely rethought - and presumably cannot be solved with legal certainty. At the same time, German lawmakers are coming under international implementation pressure.

1. Einleitung

Der Megatrend der Digitalisierung beschleunigt die Welt immer mehr und stellt alle Lebensbereiche des Menschen vor neue Herausforderungen. Er erfasst selbstverständlich auch die Automobilindustrie mit allen damit einhergehenden Vorzügen, wie technischen Innovationen, aber auch Nachteilen, wie z.B. Arbeitsplatzverlusten. Der technische Fortschritt, von je her der Traum der Ingenieure, hat das Automobil von der Entwicklung des Verbrennungsmotors, über die Serienfertigung bis hin zum (teil-) automatischen Fahren weiterentwickelt. Die Kehrseite der Medaille fortschreitender Technologisierung und Automatisierung bekommen hingegen auch die Juristen zu spüren, die versuchen müssen, damit Schritt zu halten, die Rechtsordnung und Technologisierung fortwährend aufeinander abzustimmen. Von Zeit zu Zeit stößt dies auf größere Probleme, wie etwa der rechtliche Umgang mit dem autonomen Fahren, also der juristischen Regelung der Bewegung des Autos ohne menschliche Kontrolle.

2. Autonomiestufen

Die Klassifizierung eines Fahrzeugs hinsichtlich seiner Automatisierung wird ganz überwiegend in sechs Abstufungen (sog. SAE-Level) vorgenommen. Auf der untersten Stufe („Stufe 0“) gilt das „driver only-Prinzip“: die Kontrolle der Fahrzeugsteuerung (Längs- und Querbewegungen) obliegt allein dem Fahrer. In den letzten einhundert Jahren hat sich dabei das Autonomielevel stetig erhöht, und auch den Begriff des „Führens“ eines Fahrzeugs einer Veränderung unterzogen. Die eigentliche Steuerungsaufgabe hat sich gewandelt von einer jederzeitigen aktiven Einwirkung auf das Fahrzeug (Beschleunigen, Lenken, Verzögern etc.), hin zu einer Überwachungsaufgabe der Fahrzeugfunktionen. Moderne Fahrzeuge können heute in der Regel den Autonomiestufen 2 und 3 zugerechnet werden. Die Autonomiestufe 4 würde ein dauerhaftes automatisches Führen des Fahrzeugs bedeuten, wobei der Fahrer aufgefordert werden

kann, in besonderen Situationen die Kontrolle wieder zu übernehmen. Die Autonomiestufe 5 würde ein menschliches Eingreifen nur noch in Notfallsituationen ermöglichen („Not-Stopp“).

Das Ziel der Automatisierung des Verkehrs ist dabei neben einem Komfortgewinn für den Fahrer auch die sog. „Vision Zero“ – die Reduktion der Anzahl der Unfälle und Geschädigten im Straßenverkehr auf null (wobei im Jahr 2020 nach Angaben des ADAC in Deutschland bei etwa 2,7 Millionen Verkehrsunfällen rund 2.800 Menschen getötet wurden; aufgrund der COVID19-Pandemie vergleichsweise geringere Fallzahlen).

| SAE Level | SAE Name | SAE Narrative Definition | Execution of Steering/ Acceleration/ Deceleration | Monitoring of Driving Environment | Fallback Performance of Dynamic Driving Task |
|--|------------------------|--|---|-----------------------------------|--|
| Human Driver monitors the driving environment | | | | | |
| 0 | No Automation | the full-time performance by the human driver of all aspects of the dynamic driving task | Human Driver | Human Driver | Human Driver |
| 1 | Driver Assistance | the driving mode-specific execution by a driver assistance system of either steering or acceleration/deceleration | Human Driver and Systems | Human Driver | Human Driver |
| 2 | Partial Automation | Part-time or driving mode-dependent execution by one or more driver assistance systems of both steering and acceleration/deceleration. Human driver performs all other aspects of the dynamic driving task | System | Human Driver | Human Driver |
| Automated driving system ("system") monitors the driving environment | | | | | |
| 3 | Conditional Automation | driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task - human driver does respond appropriately to a request to intervene | System | System | Human Driver |
| 4 | High Automation | driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task - human driver does not respond appropriately to a request to intervene | System | System | System |
| 5 | Full Automation | full-time performance by an automated driving system of all aspects of the dynamic driving task under all roadway and environmental conditions that can be managed by a human driver | System | System | System |

Abbildung 1: Automatisierungslevel (CEDR 2014, S. 7)

3. „Führungsrolle“ Deutschlands

„Deutschland soll eine Führungsrolle beim autonomen Fahren einnehmen“, so das Bundesministerium für Verkehr und digitale Infrastruktur (BMVI). Daher trat bereits am 21. Juni 2017 ein Gesetz zum automatisierten Fahren in Kraft. Hierbei wurden im Wesentlichen die Rechte und Pflichten des Fahrzeugführers während der automatisierten Fahrphase geregelt, wonach die Fahraufgabe unter bestimmten Voraussetzungen durch das Fahrzeug übernommen werden darf (SAE-Level 3); ein – (rechtlich)

verantwortlicher – Fahrer ist dabei aber weiterhin jederzeit zwingend notwendig. Eine rechtlich große, umfassende Regelung des autonomen Fahrens war damit freilich nicht verbunden. Es handelte sich eher um eine dringend notwendige Minimalanpassung des Rechts in eher unkritischen Bereichen. Mit dem neuen Regierungsentwurf zu einem weiteren Gesetz zum autonomen Fahren vom 8. Februar 2021 möchte die Bundesregierung darauf aufbauend nun den Rechtsrahmen schaffen, damit automatische Kraftfahrzeuge (SAE-Level 4) in festgelegten Betriebsbereichen im Regelbetrieb fahren können. Unter „festgelegten Betriebsbereichen“ sollten dabei etwa Shuttle-Verkehre, automatisches Parken und Hub2Hub-Verkehre zu verstehen sein. Ein (teil-) autonomes Fahren im allgemeinen öffentlichen Straßenverkehr ist, entgegen den Ankündigungen des BMVI, hierunter jedoch nicht zu verstehen und wird auch nicht gesetzlich geregelt werden. Hierzu sind die sich stellenden rechtlichen Fragestellungen (noch) zu zahlreich und zu komplex. Die schon heute verfügbare Technologie hat insofern die gesellschaftliche und rechtliche Verarbeitung deutlich abgehängt.

4. Rechtsprobleme und Anpassungsbedarf

Mit dem autonomen Fahren gehen zahlreiche Rechtsprobleme einher, die unter der bestehenden Rechtsordnung nur schwer lösbar sind. Aufgrund der Vielzahl und Komplexität der Herausforderungen ist allerdings anzunehmen, dass die Fortentwicklung des Rechts an autonome Systeme in erster Linie nicht von großen Reformen bestimmt sein wird, sondern sich einzelfallbezogen viele neue Regeln durch Gesetzgeber und Rechtsprechung entwickeln werden (Borges, NJW 2018, 977). Im Zentrum steht dabei die Frage der Haftung. Hierbei stellt sich insbesondere die Frage nach der haftenden Person und der Art der Haftung.

Die Frage nach der Person, die haften soll, muss verschiedene Rollen unterscheiden: auf der einen Seite die Hersteller und Lieferanten und auf der anderen Seite die Eigentümer und Betreiber (Zech, S. 177 ff.). Im Falle autonomer Systeme wird es sich darüber hinaus anbieten, noch eine weitere Gruppe zu differenzieren: die Bediener. Dieses Konzept, das auch unterschiedliche Haftungsmodelle für die unterschiedlichen Rollen definiert, findet im Straßenverkehrsrecht bereits grundsätzlich Anwendung, wo Hersteller, Halter und Fahrer eines Fahrzeugs unterschieden werden, und müsste für autonome Systeme „nur“ weiterentwickelt werden. Interessant ist in diesem Zusammenhang eine Drucksache des EU-Parlaments, die neben den natürlichen Personen (Menschen) und juristischen Personen die Einführung einer dritten Gruppe von Rechtssubjekten erwähnt, die „elektronischen Personen“ (Europäisches Parlament, Drs. 2019/0081, Nr. 114).

Gänzlich neue Ansätze müssen hingegen im Bereich der Art der Haftung diskutiert werden. Auch hier kennt das Zivilrecht grundsätzlich bereits verschiedene Modelle.

Der klassischen „Kausalhaftung“ (Rechtsgutsverletzung – Rechtswidrigkeit – Verschulden – Schaden – Kausalität) steht die sog. „Gefährdungshaftung“ gegenüber. Hier kommt es schon durch die von einer bloßen Inbetriebnahme oder Vornahme einer potentiell gefährlichen Handlung ausgehenden typischen Gefahren (Betriebsgefahr) zu einer Haftung des Halters bzw. Handelnden für Schäden auch ohne dessen Verschulden. Insofern bestehen inzwischen Tendenzen, die Haftung für autonome Fahrzeuge, analog der Haftung für Tiere und Kinder, nach einer (abgestuften) Gefährdungshaftung zu beurteilen.

5. Größte Herausforderung: Dilemmasituationen

„Der Knackpunkt für das vollautomatische Fahren ist ein bislang ungelöstes ethisches Dilemma“, so Holger Steltzner in der Frankfurter Allgemeinen Zeitung (FAZ). Damit spielt er auf eine Situation an, in der autonome Fahrzeuge bereits im öffentlichen Straßenverkehr fahren, eine (korrekt funktionierende) Software das jeweilige Fahrzeug selbsttätig steuert und die auftretenden Verkehrssituationen auch richtig erfasst, es aber dennoch ein (unlösbares) ethisches und juristisches Dilemma auftritt; die größte juristische Herausforderung des autonomen Fahrens:

Ein autonom gelenktes Fahrzeug fährt im Gebirge vorschriftsmäßig um eine langgezogene Rechtskurve. In der Kurve liegt ein gestürzter Radfahrer mitten auf der Straße. Ein rechtzeitiges Abbremsen oder eine Kontrollübernahme durch den Fahrer sind aus Zeitgründen nicht mehr möglich. Es existieren damit nur zwei Alternativen: (i) die Spur halten, den Radfahrer überfahren und tödlich verletzen oder (ii) nach links einlenken und über die Leitplanke den Abhang hinunterstürzen, wobei der Fahrer getötet wird (Fall nach Weber, NZV 2016, 249). Welche Alternative soll der Algorithmus nun wählen?

5.1. Strafrechtliche Konsequenzen im herkömmlichen Straßenverkehr

Strafrechtlich wäre ein solcher Fall vergleichsweise einfach zu lösen, wenn ein menschlicher Fahrer das Fahrzeug gesteuert hätte. Wenn dieser sich nicht dazu entschließt, den Abhang hinunter zu stürzen und selbst ums Leben zu kommen, kommt durch das Töten des Radfahrers eine Strafbarkeit wegen Totschlags gem. § 212 Abs. 1 Strafgesetzbuch (StGB) in Betracht. Der Fahrer befindet sich allerdings in einer akuten Gefahr für sein eigenes Leben, die nur abgewendet werden kann, wenn er den Radfahrer tötet; ein anderer Ausweg existiert nicht. Zudem muss er, ohne lange mögliche Alternativen abwägen zu können, in Sekundenschnelle reagieren. Da der Straßenverkehr ein allgemein anerkanntes Lebensrisiko darstellt, stellt die (vorschriftsmäßige) Teilnahme daran auch keine erhöhte strafrechtliche Verantwortlichkeit des Fahrers dar. Eine strafrechtliche Verurteilung käme daher beim Überfahren und Töten des Radfahrers

nicht in Betracht; der Fahrer wäre entschuldigt gem. § 35 StGB.

5.2. Probleme bei autonomen Fahrzeugen

Anders als bei einem herkömmlich gelenkten Fahrzeug hängt die Auswahl der Alternative bei einem autonom agierenden Fahrzeug davon ab, ob der Programmierer (unter Umständen bereits Jahre zuvor und in gründlicher Abwägung) in einer solchen Situation dem Leben des Fahrzeuginsassen oder des Radfahrers den Vorrang eingeräumt hat. Da sich Fahrzeuge, die im Zweifel das Leben der Insassen opfern, nur schwer bis gar nicht verkaufen lassen würden, muss davon ausgegangen werden, dass die Software des Fahrzeugs so programmiert wurde, dass im Zweifel der Radfahrer getötet wird. Diese feste Lösung wäre allerdings im Ergebnis genauso rechtswidrig wie eine flexible, je nach Unfallsituation angepasste Lösung.

Unter einer solchen flexiblen Lösung werden Programmieransätze verstanden, die etwa in Unfallsituationen Strafpunkte für Sach- und Personenschäden vergeben (Sachschaden = 1 Strafpunkt, verletzte Person = 2 Strafpunkte, getötete Person = 5 Strafpunkte etc.) und verschiedene Handlungsalternativen im Hinblick auf die Anzahl an möglichen Strafpunkten durchrechnen, um dann den „besten“ Unfallverlauf (im Sinne der wenigsten Strafpunkte) zu wählen. Andere Ansätze bevorzugen etwa den Einsatz eines Zufallsgenerators, der in einer Unfallsituation die zu wählende Alternative für das Fahrzeug zufällig vorgibt. Alle flexiblen Ansätze haben dabei das Problem, dass zum einen der Unfallverlauf nicht sicher bis ins letzte Details vorausberechnet werden kann und dass in jeder Lösung der Mensch auf eine bloße Zählgröße herabgewürdigt werden würde. Insofern hat die BMVI-Ethik-Kommission richtigerweise schon im Jahr 2017 festgestellt, dass zwar Sachschäden zu akzeptieren sind, um Personenschäden zu vermeiden, aber gleichzeitig bei ausschließlichem Vorliegen von möglichen Personenschäden keine Klassifizierung von Personen erfolgen darf (BMVI, Ethik-Kommission, S. 11), etwa nach Anzahl, Geschlecht, Krankheitserscheinungen oder Alter. Diese Lehre aus den massiven Verstößen gegen die Menschenwürde (heute Art. 1 Abs. 1 Grundgesetz) aus der Zeit des Dritten Reichs hatte seinerzeit auch das Bundesverfassungsgericht (BVerfG) aufgegriffen, als es die Abschussermächtigung des Verteidigungsministers für entführte Passagierflugzeuge nach § 14 Abs. 3 Luftsicherheitsgesetz im Jahr 2006 für verfassungswidrig erklärte (BVerfG, NJW 2006, 751).

Aus dem gleichen Grund ist auch der Einsatz selbstlernender Algorithmen im Straßenverkehr grundsätzlich nur mit größten rechtlichen Unsicherheiten verbunden, da in diesen Fällen nicht einmal mehr die Programmierer wissen, in welche Richtung sich die Software entwickelt und welche Verhaltensweisen sie im Straßenverkehr „erlernt“. Abgesehen von diesen Unwägbarkeiten

kann und darf es nach Ansicht der Ethik-Kommission niemals eine irreversible Unterwerfung des Menschen unter technische Systeme geben.

6. Lösungsmöglichkeit

Ist einmal ein Status erreicht, in dem flächendeckend autonomes Fahren erfolgt und kritische (Unfall-) Situationen gar nicht erst entstehen (können), ist das autonome Fahren auch rechtlich zu bewältigen. Auf dem Weg dorthin wird es aber zwangsläufig zu einer Vielzahl an Dilemmasituationen kommen, die rechtlich nicht eindeutig normierbar und auch nicht ethisch zweifelsfrei programmierbar sind (Lange, NZV 2017, 345). Es stellt sich daher die Frage, wie der Weg bis dahin rechtssicher gestaltet werden kann.

Der von der Bundesregierung durch das neue Gesetz zum autonomen Fahren im Februar 2021 offenbar eingeschlagene Weg deutet auf eine „Trippelschrittlösung“ hin. In vielen kleinen, abgrenzbaren Bereichen, in denen jeweils kritische Unfallsituationen mit Menschen weitgehend vermieden werden können, sollten jeweils autonome (und untereinander vernetzte) Fahrzeuge zugelassen werden. Irgendwann sollen dann diese einzelnen Inseln zusammengelegt und zu einem flächendeckenden System verbunden werden. Solange muss allerdings die technische Entwicklung auf die gesellschaftlichen und damit auch die rechtlichen Verarbeitungsprozesse warten. Fraglich ist insoweit nur, ob andere Länder mit anderen Rechtssystemen die gleiche Strategie verfolgen, oder mit einem höheren Risiko schneller bei der Umsetzung des autonomen Fahrens agieren werden.

Als Beispiel sind hier die USA zu nennen, wo das Straßenverkehrsrecht nicht, wie in Deutschland, eine Bundesangelegenheit ist, sondern der Jurisdiktion der einzelnen Bundesstaaten unterliegt. Dies führte unlängst zu einem regelrechten Wettbewerb der Bundesstaaten um die liberalsten Gesetze zum autonomen Fahren, um Unternehmen wie Google und Tesla anzulocken. Derzeit bieten Kalifornien, Arizona und Nevada die Spitze der Liberalisierungsbewegung, wo die Autohersteller ihre Systeme auch selbst für den Straßenverkehr zertifizieren können und in weiten Teilen bei guten Wetterbedingungen nur eine geringe Bevölkerungsdichte herrscht; optimale Bedingungen für den Betrieb von Testflotten.

7. Fazit

Autonome Systeme – und insbesondere der autonome Straßenverkehr – sind ein Kernelement der gegenwärtig entstehenden digitalen Gesellschaft. Das autonome Fahren stellt dabei hohe Anforderungen an den technischen Fortschritt. Die Entwicklung der letzten Jahre verbunden mit dem aktuellen Forschungsstand lässt allerdings vermuten, dass die Hürde zum autonom agierenden Fahrzeug in nicht mehr allzu ferner Zukunft, jedenfalls aus technischer Sicht, erfolgreich genommen werden wird.

Hierdurch wird allerdings die Rechtsordnung vor nur sehr schwer lösbare Probleme gestellt, denn sie ist auf autonome Systeme (noch) nicht ausreichend vorbereitet; und auch ihre Fortentwicklung ist mit ethischen Fragen verbunden, die sich – im Gegensatz zu technischen Herausforderungen – vielleicht niemals lösen lassen. Nur wenn es gelingt, die durch den technologischen Fortschritt geschaffenen neuen Risiken im Rahmen neuer rechtlicher Haftungssysteme auch akzeptabel zuzuweisen, können autonome Systeme in die menschliche Gesellschaft voll integriert werden.

Die Frage ist nur, ob eine solche Integration auch wünschenswert ist.

Literaturverzeichnis

- [1] Bundesministerium für Verkehr und digitale Infrastruktur: Ethik-Kommission Automatisiertes und vernetztes Fahren, Juni 2017.
- [2] Borges, Georg: Rechtliche Rahmenbedingungen für autonome Systeme, NJW 2018, S. 977-982.
- [3] Bundesverfassungsgericht: Nichtigkeit der Abschussermächtigung im Luftsicherheitsgesetz, Urteil vom 15.02.2006 - 1 BvR 357/05, NJW 2006, 751.
- [4] Conference of European Directors of Roads (CEDR): Transnational Road Research Programme Call 2014, http://www.bast.de/DE/BASSt/Forschung/Forschungsfoerderung/Downloads/cedr_call_2014_2.pdf?__blob=publication-file&v=2 (letzter Abruf am 15.02.2021).
- [5] Lange, Ulrich: Automatisiertes und autonomes Fahren – eine verkehrs-, wirtschafts- und rechtspolitische Einordnung, NZV 2017, S. 345-352.
- [6] Weber, Philipp: Dilemmasituationen beim autonomen Fahren, NZV 2016, S. 249-254.
- [7] Zech, Herbert: Zivilrechtliche Haftung für den Einsatz von Robotern – Zuweisung von Automatisierungs- und Autonomierisiken Intelligente Agenten und das Recht, in: Gless, S. / Seelmann, K., Intelligente Agenten und das Recht, 2016, S. 163-204.

Betrieb Sächsischer Testfelder für das automatisierte und vernetzte Fahren

Christian Andrä

Sächsische Energieagentur – SAENA GmbH, Pirnaische Straße 9, 01069 Dresden

Automatisierte und vernetzte Fahrfunktionen sind bei vielen Herstellern / OEMs bereits verbaut. Dabei ändert sich der Charakter dieser Funktionen von der reinen Anzeige über unterstützende Eingriffe hin zur automatisierten Bewältigung von Fahraufgaben. Wir zeigen wie diese Fahraufgaben durch die Vernetzung der Fahrzeuge mit der Infrastruktur, aber auch mit anderen Verkehrsteilnehmern zur Steigerung der Effizienz und Verkehrssicherheit beitragen kann. Dafür werden bestehende und in der Entwicklung befindliche Nachrichtentypen untersucht und deren Beiträge zur Mobilität der Zukunft herausgearbeitet. Abschließend werden die Testkorridore in verschiedenen sächsischen Städten betrachtet, die speziellen Arbeitsinhalte dargestellt und ein Ausblick auf weitere Entwicklungen und den Betrieb solcher Testfelder am Beispiel Sachsen vorgenommen.

Automated and connected driving functions are entering the market already. Carmakers introduce such functions in their premium-segment models. The character of the functionality alters from an additive information, rather displaying approach to a more and more supportive one in the sense of intervention, even up to take over for particular situation. We will demonstrate that these functions are helping the traffic as such to be more efficient and that they are supporting road safety. Therefore, the currently being used messages but also messages in the development phase will be investigated and their contribution to future traffic systems will be highlighted. Concluding with an overview of existing corridors and the work that is been carried out, we will lay out how such test corridors can be operated based on the example in Saxony.

1. Einleitung

Der Aufbau von Testinfrastruktur für die Erprobung von automatisierten und vernetzten Fahrfunktionen ist seit spätestens 2016 mit der Auslobung der Urbanen Testfelder des Bundesministeriums für Verkehr und digitale Infrastruktur (BMVI) im öffentlichen Raum angekommen. In Dresden sind unter dem Projektschirm der „Synchronen Mobilität 2023“ verschiedene Projekte durchgeführt worden, die neben den Fahrfunktionen auch die Vernetzung und die Absicherung in den Blick nahmen. Die dabei entstandene und aufgebaute Infrastruktur soll nach Ablauf der Projektlaufzeiten in den Produktiv-Betrieb überführt werden und weitere Erprobungen seriennaher Funktionalitäten der kooperativen und vernetzten Mobilität ermöglichen, sowie den Fahrbetrieb von modernen Fahrzeugen unterstützen. Darüber hinaus werden weitere Korridore ausgestattet. Es sollen über die ansässigen Unternehmen hinaus auch deutschland- und europaweit Partner angesprochen werden, diese Möglichkeiten zu nutzen. Dafür wird der Betrieb organisatorisch und technisch sichergestellt.

Im Weiteren wird auf die verschiedenen Funktionen eingegangen, die das vernetzte und automatisierte Fahren unterstützen, wie diese auf den sächsischen Testfeldern bei der Entwicklung von Fahrfunktionen gewirkt haben und auch weiterhin wirken und neue Nachrichteninhalte für zukünftige Fahrfunktionen entstanden sind. Abschließend wird die Testfeldstruktur und Testfeldorganisation dargestellt, die die im Aufbau befindlichen und bestehenden Testinfrastrukturelemente koordiniert,

durch Öffentlichkeitsarbeit und Wirken in überregionalen Gremien die Außendarstellung unterstützt, sowie den regelmäßigen Austausch zwischen den Testfeldern, aber auch mit neuen und externen Partnern anregt und organisiert.

2. Anzeige und Fahrfunktionen

Über die Fahrassistentenfunktionen sind teillautomatisierte Systeme in den Funktionsumfang heutiger Oberklassefahrzeuge in die Breite migriert. Die aktuelle Entwicklung fokussiert jedoch bereits darauf, die Fahraufgabe in bestimmten Situationen oder für gegebene Umstände vollständig dem System zu übertragen (SAE Level 3 und 4 [1]). Entlang des Entwicklungsprozesses werden die Komponenten und Systeme auf verschiedene Weise getestet, z.B. SIL (Software in the Loop) oder HIL (Hardware in the Loop), bevor sie dann im Gesamtfahrzeug von den Entwicklern appliziert und getestet werden. Die Komponenten, Sub-Systeme und Systeme werden dafür viele Stunden bzw. hunderte Kilometer in virtuellen Umgebungen und Szenarien betrieben und ihr Systemverhalten analysiert. Dabei nimmt die Abstraktion und Virtualisierung immer weiter ab und die Umgebung, in der die Funktionen arbeiten, ähnelt immer mehr dem Zielfahrzeug.

Die heutigen Fahrfunktionen basieren noch weitestgehend auf den im Fahrzeug verbauten Sensoren und Karten. Komfortfunktionen, wie zum Beispiel ParkNow [2] nutzen die Datenverbindung des verbauten Mobilfunk-

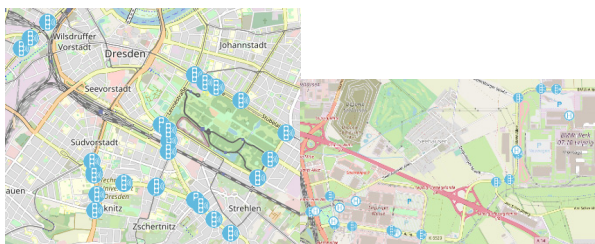


Abbildung 1: Standorte der Roadside Units auf dem Digitalen Testfeld Dresden, Stand: Februar 2021

datenzugangs. Vernetzte Fahrfunktionen werden in immer mehr Modellen der Hersteller verbaut und unterstützen so den Fahrer in verschiedenen Situationen.

C-ITS Services

Mithilfe der C-ITS (Cooperative Intelligent Transport Systems) Services, die das automatisierte und vernetzte Fahren sicherer und effizienter gestalten, können vernetzte Fahrfunktionen dem Fahrer verfügbar gemacht werden. C-ITS Services basieren auf der vernetzten Datenkommunikation zwischen einzelnen Fahrzeugen (V2V) oder zwischen Fahrzeugen und der Infrastruktur (V2I), allgemein als Vehicle to Everything (V2X) zusammengefasst. Dabei werden unter C-ITS Services eine Reihe an verschiedenen Diensten für vernetzte Fahrfunktionen gesammelt, die sich in ihrem Nachrichtentyp, der Informationen die sie enthalten und der daraus resultierenden Möglichkeiten für das vernetzte und automatisierte Fahren voneinander abheben. Dafür werden Day 1, Day 2, Day 3+ Services unterschieden: Während sich Day 1 Services auf den Austausch von Informationen, die das vorausschauende Fahren fördern, konzentrieren, verbessern Day 2 Services hingegen die Servicequalität und können Informationen zur Wahrnehmung und zum Situations-Bewusstsein enthalten. Darüber hinaus ebnet Day 3+ Services den weiteren Weg zum kooperativen, unfallfreien automatisierten Fahren mittels hochentwickelter Dienste, wie z. B. das Teilen von Absichten oder die Kooperationsunterstützung. [3]

Ein Beispiel für einen Day 1 Anwendungsfall ist der GLOSA-Dienst (Green Light Optimised Speed Advisory): Hierbei wird mittels der an Lichtsignalanlagen (LSA) installierten Kommunikationstechnik, der Roadside Unit (RSU), der Austausch der Restgrünzeit ermöglicht. Diese Information wird dem Fahrer im Fahrzeug als Empfehlung bereitgestellt. Er kann daraufhin entscheiden, ob eine Grünphase noch zu schaffen ist oder er bereits ausrollen kann, um Energie zu sparen (vor allem im Hinblick auf den Einsatz von E-Fahrzeugen relevant). Andersherum kann eine Restrot-Information dazu genutzt werden, das Heranfahren an die Ampel etwas zu verlangsamen und einen Stopp zu vermeiden um auch so wiederum Energie zu sparen. Bei einer zentralen Ampelsteuerung sowie untereinander vernetzter LSAs bietet solch

¹ Als eines der vom Bundesministerium für Verkehr und digitale Infrastruktur (BMVI) benannten digitalen Testfelder in Deutschland, werden auf dem Digitalen Testfeld Dresden Vorhaben für

ein System sogar weitere Vorteile: Durch den gegenseitigen Informationsaustausch zwischen den LSAs können grüne Wellen geschaffen bzw. aufrechterhalten werden. Dies ist selbst bei sich ändernden Verkehrssituationen möglich.

Die Entwicklung weiterer Funktionalitäten, wie zum Beispiel die Nachrichtenformate für die Day 2 Services „Emergency Vehicle Approaching“ (EVA) oder „Vulnerable Road User“ (VRU) erfordern eine RSU, die frei programmierbar ist. So kann auf der einen Seite sichergestellt werden, dass die bereits standardisierten Nachrichten weiterhin versendet werden und auf der anderen Seite darüber hinaus zusätzliche Nachrichtenformate für Erprobungszwecke auf dem Funkkanal bereitgestellt werden können. Diese sogenannten Forschungs-RSUs sind derzeit an verschiedenen Stellen auf dem Digitalen Testfeld Dresden¹ verbaut und werden von der Forschungsinstitution betrieben. Die RSUs für Produktivzwecke werden zumeist von den örtlichen Behörden oder Ämtern betrieben.

Die Erprobungsinhalte sehen neben den Informationen von der Infrastruktur an die Fahrzeuge auch noch weitere Inhalte vor. Unter anderem wird ein Nachrichtenaustausch ermöglicht, der kooperative Fahrmanöver unterstützt. Dafür tauschen die Fahrzeuge Nachrichten über ihre Intention aus, die Fahrspur zu wechseln. Weitere vernetzte und automatisierte Fahrzeuge können dann im fließenden Verkehr eine Lücke zum Einfädeln aufmachen und so einen steten Verkehrsfluss ermöglichen. Anwendung kann ein solches Manöver beim Einfädeln von ÖPNV Fahrzeugen in den Fließverkehr finden.

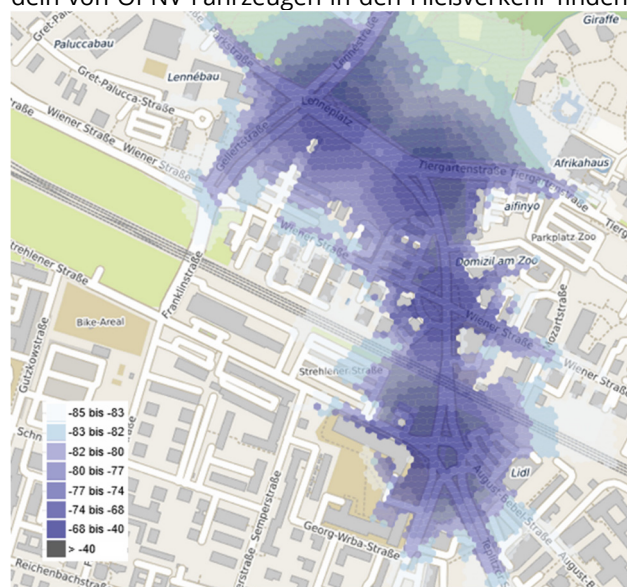


Abbildung 2: Roadside Unit zur Steigerung der Effizienz durch Bereitstellung von Kommunikationsmöglichkeiten

Im Oktober 2019 wurde ein solches Manöver auf dem ausgestatteten Korridor am Dresdner Flughafen als Pro-

das automatisierte und vernetzte Fahren durch das BMVI unterstützt.

jektabschluss verschiedener Projekte der „Synchronen Mobilität 2023“ demonstriert.

Entwicklung der Nachrichten und Korridore

Die sächsischen IVS-Testfelder umfassen derzeit verschiedene Straßenkategorien auf einer Gesamtlänge von

ca. 32 km. Das Testfeld im Leipziger Norden beinhaltet eine Strecke, die auf einer Länge von ca. 7 km über eine Ausstattung mit Produktiv-RSUs verfügt. Weitere Ausstattungsmerkmale sind die zum Einsatz kommenden Landmarken, die die Präzision der Lokalisierung erhöhen. Die Korridore in Dresden befinden sich auf einer Gesamtlänge

von ca. 25 km überwiegend im innerstädtischen Bereich. Das Testfeld besteht zum Teil aus zusammenhängenden Abschnitten über mehrere Kreuzungen und RSUs hinweg (siehe Standorte der Roadside Units auf dem Digitalen Testfeld Dresden, Stand: Februar 2021). Die Korridore sind sowohl mit Produktiv-RSUs als auch teilweise mit Forschungs-RSUs ausgestattet. Ein weiterer Abschnitt ist der zum Teil außerorts verlaufende Korridor zwischen Dresden und Bannewitz (B170). Dieser wird zukünftig an die urbanen Korridore signaltechnisch angeschlossen und befindet sich derzeit in der Ausstattungsphase. Die unter dem Digitalen Testfeld Dresden zusammengefassten Korridore befinden sich entlang verschiedener Straßenkategorien (u.a. Haupt-, Neben-, Zufahrtsstraßen), bieten Interaktionen mit weiteren Verkehrsteilnehmern (u.a. Straßenbahnen, Fußgänger, Radfahrer) und weisen weitere spezielle Straßeninfrastrukturen (u.a. Kreuzungen, straßenbündiger Bahnkörper) auf. Durch diese Testfeldcharakteristika sowie durch die fortgeführte Infrastrukturertüchtigung im urbanen als auch perspektivisch im ruralen Raum, ergeben sich eine Vielzahl an Erprobungs- und Testszenarien für die Vernetzung und Automatisierung des Straßenverkehrs.

Im aktuellen Projekt C-ROADS Germany - Urban Nodes² findet mithilfe der Forschungsinfrastruktur des Digitalen Testfelds Dresden die Weiterentwicklung von Nachrichtenformaten und C-ITS Services statt. Neue Inhalte werden auf den frei programmierbaren Forschungs-RSUs implementiert und erprobt. Dafür arbeiten die Dresdner Forscher zusammen mit Partnern in Kassel und Hamburg.

3. Zusammenwirken und Angebot

Die Sächsische Energieagentur – SAENA GmbH hat für den 24/7 Betrieb der verbauten Anlagen die verschiedenen technischen Betreiber zusammen gebracht und gemeinsam ein Koordinatoren-Konzept erarbeitet. Hierbei ist eine klare Aufteilung der Zuständigkeiten vorgenommen worden, wobei die SAENA die Rolle des organisatorischen Koordinators ausübt. Im Rahmen dessen steht

die SAENA interessierten Partnern und externen Akteuren als zentraler Ansprechpartner für die Infrastrukturelemente und -ausstattung der IVS-Testfelder in Sachsen zur Verfügung. Darüber hinaus unterstützt die SAENA die Außendarstellung und das Vermarktungskonzept der auf den sächsischen Testfeldern bestehenden Inhalte und Angebote. Im Zuge dessen bietet die SAENA ein kartenbasiertes Onlinetool unter <https://www.energieportal-sachsen.de/> an, das einen Überblick über die Ausstattungsmerkmale der jeweiligen Testfelder gibt. Darin können Informationen zu den Standorten der an den LSA installierten RSUs, zum bereitgestellten Nachrichtenumfang und zu den Projekten, in dessen Zuge die Ausstattung vorgenommen wurde, abgerufen werden.

Ausblick

Das Kartentool zur IVS-Infrastrukturausstattung in Sachsen wird in einem nächsten Schritt um die Darstellung der empfangenen Signalstärke von verschiedenen RSUs in deren Umfeld (siehe Roadside Unit zur Steigerung der Effizienz durch Bereitstellung von Kommunikationsmöglichkeiten) erweitert.

Perspektivisch wird die Kartendarstellung um weitere Inhalte ergänzt, die über die Verzeichnung der RSUs hinausgehen. Es werden weitere Infrastrukturelemente mit aufgenommen, die das automatisierte und vernetzte Fahren immer effizienter und sicherer gestalten. Dazu zählen z. B. Fahrbahnmarkierungen, wie sie bereits auf der B170 zwischen Bannewitz und Dresden aufgebracht wurden. Diese Markierung ist gerade bei widrigen Bedingungen, z. B. bei einem Wasserfilm auf der Straße, mit verbesserten Reflexionseigenschaften für eine erhöhte Sichtbarkeit ausgestattet, die es dem menschlichen Auge aber auch optischen Systemen zum automatisierten Fahren vereinfacht, diese besser zu erkennen. Darüber hinaus kommt entlang des erwähnten Streckenabschnitts auf der B170 eine neue Antennentechnologie zum Einsatz, die hier im Realbetrieb erprobt wird. Der auf der Technik des Leckwellenleiters basierende Antennentyp kann aufgrund seiner Eigenschaften nicht nur entlang der Strecke eine homogene Signalabstrahlung ermöglichen, sondern bietet darüber hinaus die Möglichkeit neben den Mobilfunkstandards wie 4G und 5G, Automotive WLAN 11p, DAB+ auch noch weitere Kommunikationstechnologien bis zu 6 GHz abzustrahlen. Besonders interessant wird diese Technologie dadurch, dass neben der Bereitstellung verschiedener Simplex und Duplex Kommunikationsmöglichkeiten auch der Verkehrsfluss erfasst werden kann. So entsteht für die angeschlossene Verkehrsleitzentrale ein präzises Bild von der aktuellen Verkehrssituation entlang des Korridors. Diese Informationen können wiederum dazu genutzt werden, dynamische Routen zu berechnen und In-

² C-ROADS Germany - Urban Nodes wird von der Europäischen Union als „Connected Europe Facility“ gefördert.

dividualfahrzeugen somit eine optimierte und damit energiesparsamere Route vorzuschlagen.

Literaturverzeichnis

- [1] SAE International, „Levels of Driving Automation,“ 2021. [Online]. Available: <https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic>. [Zugriff am 15. Januar 2021].
- [2] Park Now GmbH, „Parken integriert im Navigationssystem,“ 2021. [Online]. Available: <https://de.park-now.com/fahrzeugintegration/>. [Zugriff am 15. Januar 2021].
- [3] Car 2 Car Communication Consortium, „About C-ITS: C-ITS Services,“ [Online]. Available: <https://www.car-2-car.org/about-c-its/>. [Zugriff am 20. Januar 2021].

Smarte Mobilitätsketten als Zukunftsbild für den ländlichen Raum

Marco Rehme, Udo Wehner, Julia Döring, Steve Rother
Technische Universität Chemnitz, 09107 Chemnitz

In ländlichen Räumen bestehen große sowie wachsende Herausforderungen bei der Versorgung mit öffentlichen Verkehrsangeboten. Zugleich besitzt ländliche Mobilität große Potenziale in allen drei Nachhaltigkeitsdimensionen. Im Projekt „Smarte Mobilitätsketten im ländlichen Raum“ entwickelt ein interdisziplinäres Team aus Wissenschafts- und Praxispartnern bis Ende 2021 komplementäre Lösungen für den ÖPNV im Erzgebirge auf Basis vernetzter, multimodaler Angebote. Die Betrachtung reicht über neue Service-, Wege- und Wertschöpfungsketten bis hin zur Bewertung der Umsetzbarkeit und Vorteilhaftigkeit. Im Beitrag werden das Projektdesign, der angewendete Methodenmix und die Schnittstellen zwischen Gesellschaft/Nutzern, Technologie, Markt sowie Recht aufgezeigt und erste Zwischenergebnisse zu identifizierten Innovationsbarrieren, Verbesserungsmöglichkeiten sowie zu neuen Denkansätzen und Modulbausteinen für ländliche Mobilität präsentiert.

Rural areas are facing growing challenges in the provision of public transport services. At the same time, rural mobility has great potential in all three dimensions of sustainability. In the project "Smart Mobility Chains in Rural Areas", an interdisciplinary team of scientific and practical partners is developing complementary solutions for public transport in the Ore Mountains on the basis of interconnected, multimodal mobility offers. The analysis ranges from new service chains, trip chains and supply chains to the evaluation of their feasibility and benefits. In the paper the project design, the mix of applied methods and the interfaces between society/users, technology, market and law are pointed out. In addition, interim results on identified innovation barriers, as well as on optimization measures and new approaches and modules for rural mobility are presented.

1. Projekthintergrund

Das Projekt „Smarte Mobilitätsketten im ländlichen Raum“ ordnet sich in das BMBF-geförderte WIRI-Vorhaben „Smart Rail Connectivity Campus“ (SRCC), ein Bündnis unter Federführung der TU Chemnitz, der Stadt Annaberg-Buchholz und der DB RegioNetz, ein. Der SRCC versteht sich strategisch als – für die Region Erzgebirge-Chemnitz völlig neuartige – institutionelle und infrastrukturelle Plattform für den regionalen Wandel. Seine Forschungs- und Erprobungsinfrastruktur sowie seine institutionelle Neutralität und Offenheit bilden ideale Voraussetzungen für kooperative Innovationsaktivitäten. Der SRCC soll als ein dauerhaftes, agiles Innovationsnetzwerk etabliert werden und damit die Region insgesamt zum Treiber des Technologiefortschritts im weit gefassten Feld Smart Rail entwickeln.

Zu diesem Feld gehören auch ganzheitliche Mobilitätskonzepte unter Einbeziehung des Schienenpersonennahverkehrs. Im ländlichen Raum ist Mobilität heute weitgehend mit motorisiertem Individualverkehr (MIV) gleichzusetzen, was mit negativen Klima- und Umweltwirkungen und einer geringen Standortattraktivität für mobilitätseingeschränkte Personen einhergeht. Attraktive Angebote des öffentlichen Personennahverkehrs (ÖPNV) sind in der räumlichen Peripherie derzeit kaum wirtschaftlich darstellbar. Gleichzeitig eröffnen sich im Zuge der Digitalisierung neue Möglichkeiten, bestehende und alternative Mobilitätsangebote miteinander zu koordinieren und einfacher zugänglich zu machen, um so Nachhaltigkeitspotenziale zu erschließen.

Darauf abzielende Ansätze versuchen smarte Mobilitätsketten in ländlichen Räumen zu konzipieren und zu ermöglichen. Solche umfassen – bezogen auf den Personenverkehr – grundsätzlich technische Systeme und Geschäftsmodelle, die es außerhalb der Städte reisenden Verkehrsteilnehmern ermöglichen, ihre Zieldestinationen unter Inanspruchnahme unterschiedlicher Verkehrsmittel (multimodal) und IT-gestützter Mobilitätsdienste zu erreichen. Zu den Systembestandteilen integrierter Mobilitätsketten gehören verschiedene Kern- und Ergänzungsleistungen, wie die (zeitweise) Bereitstellung von Fahrzeugen, das Erbringen von Personenbeförderungsleistungen, Informations-, Integrations- und Vermittlungsleistungen für vorgenannte, die Bereitstellung benötigter Kommunikations- und Verkehrsinfrastrukturen und von Mobilitätsdaten sowie im Hintergrund ablaufende Basis-IT-Dienste.

2. Projektdesign und Methodenmix

Im Rahmen des Projektes soll ein smartes, multimodales und vernetztes Mobilitätskonzept für einen ausgewählten Mobilitätsraum innerhalb der Region Chemnitz-Erzgebirge entwickelt werden. Das Konzept soll den ÖPNV in der genannten Region stärken, dabei lokal vorhandene Infrastrukturen bestmöglich nutzen und, wo nötig, sinnvoll ergänzen sowie durch den umfassenderen Rückgriff auf öffentliche und kollaborative Mobilitätsleistungen gleichzeitig die ökologische Nachhaltigkeit des dortigen Verkehrs verbessern. Bei der Entwicklung von Lösungen sollen gezielt Barrieren berücksichtigt werden, welche einer Abkehr von etablierten, MIV-dominanten Verhaltensweisen entgegenstehen. Darüber hinaus

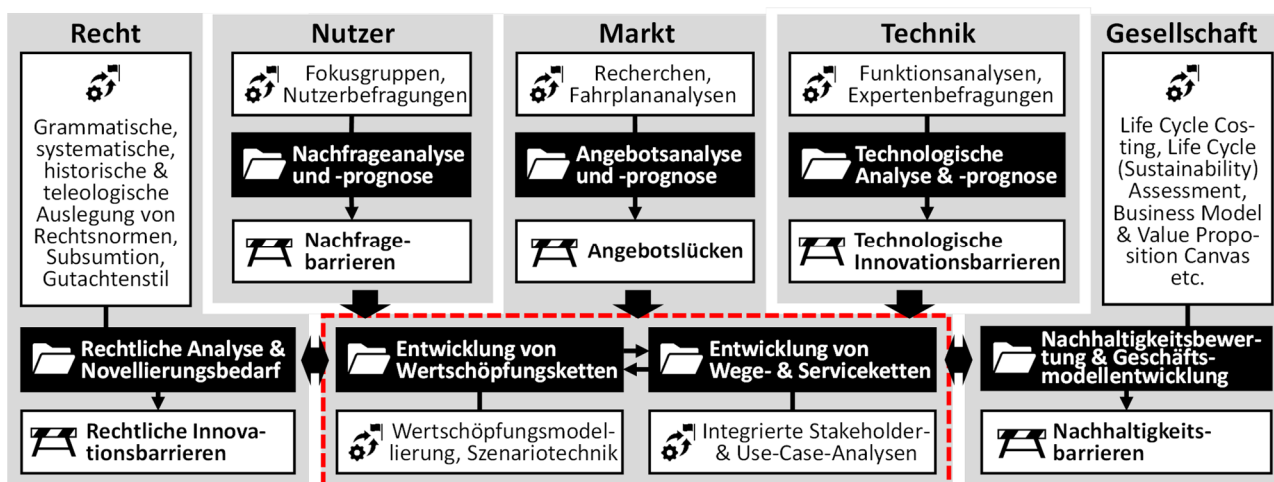


Abbildung 1: Projektdesign und Methoden

besteht der Anspruch, auch aus ökonomischer Sicht ein für ländliche Räume dauerhaft tragfähiges Gesamtkonzept zu entwickeln. In diesem Zusammenhang wird angestrebt, dass die für den Referenzraum zu entwickelnden Konzeptbausteine auch skalierbar und auf andere ländliche Räume übertragbar sind.

Das Projekt verfolgt einen interdisziplinären Ansatz. So werden die Kompetenzen von Juristen, Wirtschaftswissenschaftlern, Psychologen und Informatikern gekoppelt und ergänzt um die praktischen Erfahrungen regionaler ÖPNV-Anbieter und -Besteller sowie Unternehmen, die mit ihrem Geschäftsfeld an der Schnittstelle zum Nachfrager öffentlicher Mobilitätsleistungen aktiv sind. Die interdisziplinäre Ausrichtung des Vorhabens spiegelt sich einerseits in der Methodenwahl, andererseits in dem organisatorischen Projektaufbau wider (vgl. dazu Abb. 1). So werden umfassende empirische Untersuchungen verknüpft mit theoretisch-konzeptionellen Forschungsansätzen. Diese gegenläufige Ausrichtung der Forschungsarbeit bildet das Fundament zur Generierung valider Ergebnisse.

Das Projekt ist in sieben Arbeitspakete gegliedert (in Abb. 1 schwarz hervorgehoben), die sich in einen grundlegenden, einen flankierenden und einen konzeptionellen Teil differenzieren lassen. Im grundlegenden Teil werden zunächst Barrieren und Lücken im öffentlichen Mobilitätssystem der Region aus der Perspektive der Nachfrager, der Angebote und der eingesetzten Technologien identifiziert. Im flankierenden Teil erfolgt einerseits eine rechtliche Analyse, andererseits eine Nachhaltigkeitsbetrachtung aus ökonomischer und ökologischer Perspektive. Auch hier sollen insbesondere Barrieren, die aus rechtlicher und aus nachhaltigkeitsbezogener Perspektive bestehen, identifiziert und analysiert werden. Im Gegensatz zum grundlegenden Teil des Projekts, in dem lediglich der Ausgangszustand des Mobilitätssystems der Region untersucht wird, umfassen die im flankierenden Teil angestellten Analysen jedoch auch die im konzeptionellen Teil des Projekts entstehenden Ideen und Ansätze smarterer Mobilitätsketten in der Region.

Der Arbeit im konzeptionellen Teil des Projekts (in Abb. 1 rot umrahmt) liegt ein spezifisches Verständnis von Mobilitätsketten zu Grunde, das als Rahmen und Orientierung dient und in dessen Strukturen auch die Projektergebnisse ausgearbeitet und dargestellt werden sollen. Demnach lassen sich bezogen auf Mobilitätsketten drei Ebenen differenzieren: Die naheliegende erste Ebene bildet die Wegekette, die die eigentliche Mobilitätsleistung und damit das Überwinden räumlicher Distanzen von einem Ausgang- zu einem Zielpunkt unter Rückgriff auf verschiedene Verkehrsmittel umfasst. Diese wird flankiert von einer Servicekette, auf der u. a. das Informieren und Buchen zu sowie das Bezahlen von Mobilitätsleistungen erfolgen. Die dritte Ebene bildet schließlich die der Wertschöpfungskette, auf der die für die Erbringung von Service- und Wegeleistungen notwendigen Wertschöpfungsschritte abgebildet werden, um auch die Frage zu beantworten, wer – also welcher Akteur – entsprechende Wege- und Serviceleistungen anbieten kann.

Für die Überlegungen im konzeptionellen Teil wird die Expertise von Partnern aus dem Mobilitätssektor eingebunden. Darüber hinaus stehen die Projektverantwortlichen in Austausch mit sehr ähnlich ausgerichteten innovativen Mobilitätsvorhaben im Umkreis (in Zwönitz, Stollberg und Mittweida), um Synergien zu nutzen und abgestimmte Vorschläge zur Verbesserung ländlicher Mobilität zu unterbreiten.

3. Spannungsfeld der Mobilitätssicherung in ländlichen Räumen

Das verfolgte interdisziplinäre Projektdesign spiegelt letztlich auch das Spannungsfeld wider, in welchem sich das Bestreben nach einer nachhaltig(er)en Mobilitätssicherung bewegt. Wirksame Veränderungen des Status Quo, die helfen, ökonomischen, ökologischen und sozialen Ansprüchen besser gerecht zu werden als das etablierte Verkehrssystem, können nur dann gelingen, wenn alle relevanten Einflussbereiche – vom Nutzer über die Gesellschaft, den Markt und die Technik bis zum Recht –

mit ihren Faktoren und Schnittstellen adäquat berücksichtigt werden, und zwar beginnend bei der Konzeptionierung bis hin zur Umsetzung. Dies gilt insbesondere für ländliche Räume, welche mit Bezug auf die Mobilität strukturelle Besonderheiten und Herausforderungen, wie sinkende Bevölkerungsdichten, einen Anstieg des Durchschnittsalters, längere Mobilitätswege sowie zunehmend ausgedünnte und dennoch defizitäre ÖPNV-Angebote unter limitierenden rechtlich-organisatorischen Randbedingungen, aufweisen (vgl. [1], S. 13 ff.).

Abb. 2 stellt ohne Anspruch auf Vollständigkeit einige der zentralen Faktoren des Spannungsfeldes in den fünf genannten Einflussbereichen und die Schnittstellen zwischen diesen dar. Die geometrischen Figuren deuten die Vielzahl und Vielgestalt der Ausprägungen und Varianten einzelner Elemente an.

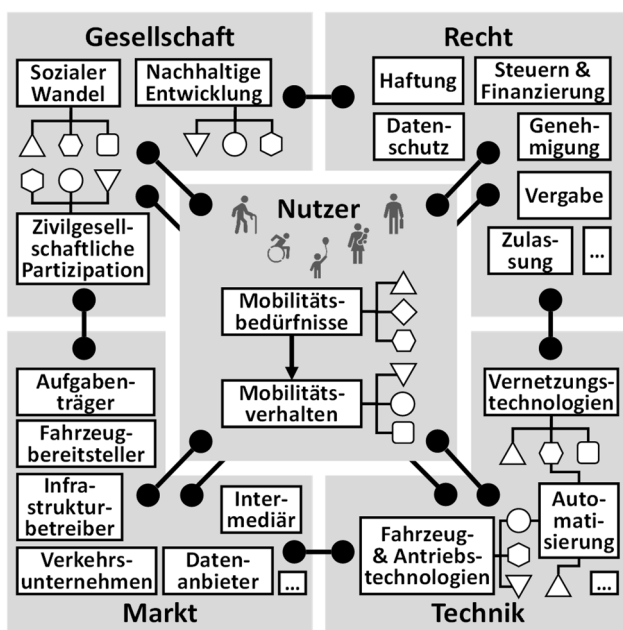


Abbildung 2: Spannungsfeld der Mobilitätssicherung

Im Mittelpunkt stehen die Nutzer, deren Bedürfnisse und Verhaltensweisen in Abhängigkeit von ihrer Gruppenzugehörigkeit vielfältige Erscheinungsformen annehmen können. Nachfolgend sind beispielhaft Schnittstellen der Nutzer zu den anderen vier Bereichen benannt:

- Zur Gesellschaft: der Einfluss soziokultureller, -ökonomischer und -demographischer Trends
- Zum Recht: die Vorschriften fahrerlaubnis-, steuer- und verkehrsstrafrechtlicher Regelungen
- Zur Technik: Technologieakzeptanz und -adoption
- Zum Markt: Kundenbeziehungen zu den unterschiedlichen Mobilitätsanbietern.

Im Markt agieren verschiedenste Dienstleister (z. B. Taxi, Bike-Sharing, Pannenhilfe), Hersteller (z. B. Pkw), Infrastrukturbetreiber (z. B. Ladesäule) und Intermediäre (welche Angebote vermitteln oder gebündelt vermark-

ten) mit den Nutzern sowie untereinander; zudem planen und bestellen öffentliche Aufgabenträger Verkehrsleistungen (vgl. [2]). Relevante Schnittstellen sind hier unter anderem:

- Zur Gesellschaft: ehrenamtlich erbrachte Fahrdienste
- Zum Recht: Organisation des ÖPNV, Beihilferecht, Ausschreibung/Direktvergabe von Verkehrsverträgen, Genehmigung der Personenbeförderung
- Zur Technik: Plattformtechnologien für die Realisierung multi- und intermodaler Mobilitätsoptionen.

Im Bereich Technik stechen an der Schnittstelle zur Gesellschaft bei der Mobilität (wie auch in anderen Anwendungsdomänen) die Trends der Digitalisierung und der sozial-ökologischen Transformation hervor, welche v. a. mit der Suche nach alternativen Fahrzeug- und Antriebstechnologien sowie fortschrittlichen Vernetzungs- und Automatisierungstechnologien (inklusive Künstliche Intelligenz) adressiert werden. Schnittstellen zum Recht bestehen hier beispielsweise in der Zulassung und Subventionierung technischer Innovationen.

Der Bereich Mobilitätsrecht selbst zerfällt in eine große Zahl an öffentlich- und privatrechtlichen Teilbereichen in der Normenhierarchie (Völker-, EU-, Bundes-, Landes-, Kommunalrecht), darunter viele verkehrsträgerspezifische Vorschriften (z. B. Straßen- und Schienenverkehrsrecht). Neben den schon genannten und in Abb. 2 dargestellten Bereichen gehören hierzu u. a. auch das Verkehrswege-, Bau- und Planungsrecht. Die zentrale Schnittstelle des Rechts zur Gesellschaft stellt die Politik dar, über die eine gemeinschaftliche Willensbildung (z. B. zu den Erfordernissen eines stärkeren Schutzes personenbezogener Daten sowie einer Mobilitätswende weg vom klassischen MIV) und die Transformation in gültiges Recht erfolgen. Im Bereich Gesellschaft spielt für den Mobilitätssektor neben der bereits erwähnten Nachhaltigkeitsorientierung und den Facetten sozialen Wandels die Bedeutungszunahme des sog. Dritten Sektors durch zivilgesellschaftliche Initiativen eine Rolle (vgl. [2]).

Die Bezeichnung Spannungsfeld macht deutlich, dass in und zwischen den fünf Bereichen auch gegensätzliche Kräfte aufeinander einwirken und Konflikte hervorrufen. Letztere gilt es, frühzeitig zu erkennen und bei der Suche nach intelligenten Mobilitätslösungen zu bedenken.

4. Barrieren für Alternativen zum MIV

Im Rahmen der grundlegenden Arbeitspakete wurde der Status Quo des öffentlichen Mobilitätssystems in der Region aus der Perspektive der Nutzer, des Angebots und der Technologien analysiert. Darauf aufbauend wurden Barrieren identifiziert und systematisiert, welche sowohl in den Wege- und Serviceketten als auch auf einer übergeordneten Ebene zu finden sind (vgl. dazu Abb. 3).

| | | | | |
|-----------------------------|---|---|--|---|
| akteursintern | ressourcen-/prozessbezogen | finanziell <ul style="list-style-type: none"> * Ausrichtung auf Schülerbeförderung, eingeschränktes ÖPNV-Angebot insb. an Tagesrandzeiten, Wochenenden, Ferien: z.B. fehlende Bedienung von Haltestellen, zu geringer u. ungleichmäßiger Takt ▪ Finanzielle Hürden für technologische Neuanschaffungen | materiell <ul style="list-style-type: none"> ○ Fahrzeuge: z.B. eingeschränkte Mitnahme (von Fahrrädern), Barrierefreiheit, fehlende Ausstattung für Echtzeitinformationen ○ Haltestellen: z. B. fehlende digitale Anzeigen, Fahrkartensystem, Unterstell-/Sitzmöglichkeit, Barrierefreiheit | organisational/personell <ul style="list-style-type: none"> ○ unzureichende Informationsbereitstellung: <ul style="list-style-type: none"> – schwer verständlich (Usability) – wenig zielgruppenorientiert (inkl. Erreichen der Nicht-Nutzer) – unvollständig (z.B. bei Umstiegen, Ticketerwerb, für Servicepunkte) ○ Lücken im Serviceumfang: z. B. Zugang zum Fahrplan, eingeschränkter Ticketerwerb |
| | output-bezogen | erfolgswirtschaftlich <ul style="list-style-type: none"> * Unwirtschaftlichkeit des ÖPNV (Daseinsvorsorge): s. finanzielle Barrieren (Ausrichtung auf Schülerbeförderung etc.) ○ Unwirtschaftlichkeit für gewerbliche Anbieter: keine alternativen Bedienformen ÖV/IV (wg. zu geringer Nachfrage) | | |
| akteursextern/-übergreifend | global | rechtlich-politisch <ul style="list-style-type: none"> * „historisch gewachsenes“ und kaum an veränderte Bedarfskenngrößen angepasstes Liniennetz * Keine Genehmigung „paralleler“ Angebote (Bus/Zug) ▪ PBefG (z. B. Hürde für neue Mobilitätslösungen) ▪ Konzessionsvergabe (z. B. Hürde für Integration v. Mikrolösungen) ▪ Komplexe Finanzierungsstruktur ▪ Auseinanderfallen von Zuständigkeiten ▪ rechtliche Hürden bei neuen Technologien (wie Shuttles) | gesellschaftlich <ul style="list-style-type: none"> ▪ Gesellschaftliche Einstellung zum IV/eigenen PKW ▪ Image des ÖPNV * Persönlicher Komfort: <ul style="list-style-type: none"> – erste/letzte Meile u. Übergangsstrecken – deutlich längere Reisezeit des ÖPNV ggü. IV – Flexibilität u. Anzahl der Umstiege * Persönliche Wirtschaftlichkeit: Preis für ÖPNV-Leistungen ▪ Akzeptanz der neuen Technologie in der Bevölkerung (z. B. moderne Bezahlssysteme, Shuttle) | technologisch <ul style="list-style-type: none"> ▪ Fehlende flächendeckende Internetkonnektivität mit hohen Übertragungsraten * Verfügbarkeit autonomer Shuttles bzw. anderer neuer Beförderungsmittel * Straßenlayout (z.B. Topologie, Kurvenradien) u. Straßenbeschaffenheit (z. B. Material, Zustand) ○ unvollständige Informationsbereitstellung: kaum Online-Auskunft über P+R o. ä. Parkmöglichkeiten ○ keine einheitliche u. einfache Technologie beim Ticketing |
| | | anbieter-spezifisch | Markt/Wettbewerbssituation/Kooperationsbeziehungen <ul style="list-style-type: none"> ▪ rechtliche u. wirtschaftliche Eigeninteressen beteiligter Stakeholder führen zu opportunistischem Verhalten beim Aufbau u. der Sichtbarkeit von Mobilitätsangeboten ▪ fehlende Abstimmung der Akteure | |
| | Legende: * Barrieren in den Wegeketten ○ Barrieren in den Serviceketten ▪ übergeordnete/-greifende Barrieren | | | |

Abbildung 3: Auszug der identifizierten Barrieren

Um im konzeptionellen Teil des Projektes gezielte Maßnahmen und Ansätze zur Überwindung der identifizierten Barrieren definieren zu können, ist eine tiefere ursachenbezogene Systematisierung der Barrieren notwendig (vgl. [3], S. 65 ff.). Hierfür wurde auf eine Systematisierung von Innovationsbarrieren in Anlehnung an [3] zurückgegriffen. Es erfolgten Anpassungen an den Sachverhalt des ÖPNV sowie die relevanten Akteure, wie z. B. Kommunen, Verkehrsunternehmen, Zweckverbände und andere Träger öffentlicher Belange.

Die erfassten Barrieren wurden in die jeweiligen Kategorien eingeordnet. Mit Blick auf die zugrundeliegende Systematisierung ist festzuhalten, dass nicht für jede Unterkategorie Barrieren im Rahmen des Projektes vorzufinden waren und solche deshalb in Abb. 3 ausgespart wurden (z. B. globale, ökonomische oder prozessbezogene, zeitliche Barrieren).

Eine durchweg überschneidungsfreie Abgrenzung der Barrieren ist nicht möglich; zum Teil sind sie voneinander abhängig. Letztlich lassen sich die meisten Barrieren auf die folgenden Hauptursachen aggregieren:

- die begrenzte, finanzielle Ausstattung und damit eng verbundene erfolgswirtschaftliche Ursachen,
- technologische Voraussetzungen, insbesondere die Aktualität der eingesetzten Technologien und
- die Wettbewerbssituation bzw. die unzureichenden Kooperationsbeziehungen der Anbieter.

Diese Barrieren, deren Überwindung notwendig ist für eine wesentliche Verbesserung im öffentlichen Mobilitätssystem und somit das Anbieten einer attraktiven Alternative zum klassischen MIV, sind prinzipiell als überwindbar einzustufen. Ihre Bewältigung ist Grundvoraussetzung für eine Veränderung der gesellschaftlichen Einstellung und des Verhaltens bezüglich der eigenen Mobilität. Ansätze hierzu werden in einem weiteren Projekt untersucht.

5. Voraussetzungen und Bausteine Smarter Mobilitätsketten auf dem Land

Im Projekt werden zukunftsweisende Mobilitätssysteme für den ländlichen Raum als modulare Rekombinationen einzelner Bausteine, basierend auf einigen grundlegenden Voraussetzungen, verstanden. Die Bausteine dienen der Überwindung einzelner oder mehrerer der in Abb. 3 aufgeführten Barrieren und können sowohl die Wege- und Serviceketten als auch die Wertschöpfungsketten betreffen.

Für die Gestaltung smarter Mobilitätsketten im ländlichen Raum stehen auf unterschiedlichen Evolutionsstufen insbesondere die nachfolgend beispielhaft aufgeführten Bausteine zur Verfügung:

- E-Car-, E-Scooter-Sharing,
- Steuerung des Parkraumangebots,
- Ausbau von Echtzeit-Informationen,
- funktional und modal umfassende (geographisch weitreichende) Mobilitätsplattform,
- On-Demand-Ridepooling,
- Letzte-Meile-Shuttles (autonom und nicht-autonom),

- virtuelle Haltestellen,
- nutzerorientiert-dynamisierter Fahrplan.

Insbesondere die Bausteine zur Erweiterung des Mobilitätsangebots scheitern im ländlichen Raum an finanziellen Barrieren (Finanzierung des ÖPNV, Wirtschaftlichkeit von Mobilitätsleistungen). Vor diesem Hintergrund werden - als weiterer flankierender Baustein - zivilgesellschaftliche oder selbstorganisierende Gemeinschaftsverkehre in den Mittelpunkt rücken müssen.

Alle genannten Bausteine erfordern die Erfüllung von Voraussetzungen. Als eine der wichtigsten ist die flächendeckende Konnektivität zu nennen, ohne die Echtzeit-Informationen, Plattformansätze und On-Demand-Services nicht realisierbar sind. Geographisch weitreichende Mobilitätsplattformen erfordern einen anbieterübergreifenden Austausch von Daten über offene standardisierte Schnittstellen bzw. Systemarchitekturen. Um die Erschließungsdefizite des eingeschränkten ÖPNV-Angebotes und des historisch gewachsenen Liniennetzes, welches sich kaum an veränderte Verkehrsströme und Bedarfskenngößen im Ergebnis sozialen und raumstrukturellen Wandels angepasst hat, zu überwinden, muss eine Abstimmung auf die Nutzerbedürfnisse erfolgen. Somit ist die Erfassung der Nutzerdaten eine zwingende Voraussetzung für bestmöglich platzierte virtuelle Haltestellen und einen nutzerorientierten Fahrplan.

Im Fortgang des Projektes wird an einer weiteren Operationalisierung und Analyse dieser Voraussetzungen in den verschiedenen Einflussbereichen des Spannungsfeldes sowie an einer zielorientierten Synthese der identifizierten Bausteine zu funktionsfähigen Wertschöpfungsketten und nachhaltigen Geschäftsmodellen gearbeitet.

Danksagung

Das Projekt wird vom BMBF im Rahmen des Programms „WIR! – Wandel durch Innovation in der Region“ gefördert. Das unterstützende Vorhaben „Transferassistent Alternative Antriebe und Digitalisierung Bahnverkehr“ wird mit Mitteln des Europäischen Sozialfonds und des Freistaates Sachsen mitfinanziert. Die Autoren danken den Fördermittelgebern sowie allen Projektbeteiligten.

Literaturverzeichnis

- [1] Richter, S.; Rehme, M.: Auto FipS. Automatisiertes Fahren in peripheren Siedlungsstrukturen, Studie im Auftrag des BMVI, Oberlungwitz, August 2018, verfügbar unter: <https://www.tib.eu/de/suchen/id/TIBKAT%3A1035298368/>
- [2] Rehme, M.; Wehner, U.; Rother, S.; Götze, U.: Management von Wertschöpfungsnetzwerken multimodaler Mobilität, in: Proff, H. (Hrsg.): Making Connected Mobility Work, Wiesbaden 2021 (39-59).
- [3] Rehme, M.; Lindner, R.; Götze, U.: Barrieren bei Geschäftsmodell-Innovationen der Neuen Mobilität, in: Proff, H.; Fojcik, T. (Hrsg.): Nationale und internationale Trends in der Mobilität, Wiesbaden 2016, 63-81.

Mobility 4ALL – Mobilität weiterdenken in Mittweida

Ralf Hartig, Alexander Knauer, Vivien Lorenz, René Härtel
Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

Mit diesem Beitrag wollen wir Ihnen einen Einblick zum gezielten Einsatz der Blockchain-Technologie für Mobilitätsanwendungen geben. Im Fokus steht dabei die Entwicklung der digitalen Identität für verschiedene Mobilitätsträger. Dabei werden heutige und zukünftige Anforderungen aufgezeigt und aktuelle Entwicklungen betrachtet. Mit der dezentralen Blockchain-Technologie soll beispielsweise ein Fahrzeug in der Lage sein, seine digitale Identität selbst zu verwalten, ohne von einem zentralen Identitätsdienstleister abhängig zu sein. Am Beispiel einer dezentralen Plattform für Sharing Mobility im ländlichen Raum wird die Bedeutung digitaler Identitäten für Mobilitätsträger erforscht und ein Ausblick auf zukünftig in der Geschäftswelt autonom handelnde Fahrzeuge gegeben. Es wird praxisnah gezeigt, wie die Blockchain-Technologie helfen kann, die Anforderungen an eine digitale Identität für Mobilitätsträger zu erfüllen und wie diese in einem nachhaltigen Geschäftsmodell Anwendung finden kann.

In this article, we want to give you an insight into the targeted use of blockchain technology for mobility applications. The focus will be on the development of the digital identity for various forms of transport. Current and future requirements will be highlighted and current developments will be considered. With decentralised blockchain technology, for example, a vehicle should be able to manage its digital identity on its own without being dependent on a central organisation providing identity services. Using the example of a decentralised platform for sharing mobility in rural areas, the importance of digital identities for vehicles of transport is explored and an outlook is given on vehicles that will act autonomously in the future business world. It will be shown in a practical way how blockchain technology can help to meet the requirements of a digital identity for mobility providers and how this can be applied in a sustainable business model.

1. Projektziele

Die Region Mittweida soll als Blockchain-Schaufensterregion etabliert werden. Als Beitrag dazu wird durch dieses Projekt zukunftsfähige Mobilität im ländlichen Raum entwickelt, die sicher, sozial, bezahlbar und an den Klimazielen ausgerichtet ist. Dazu ist es erforderlich, die Mobilität in der Region Mittweida neu zu denken. Im Rahmen des Projekts Mobility 4All soll insbesondere emissionsfreie Mobilität im Bereich der Sharing-Economy auf Basis der Blockchain-Technologie erforscht und getestet werden. Dabei nimmt die Region Mittweida die Rolle eines Reallabors für neue multimodale Mobilitätskonzepte ein.

Die Blockchain-basierte Sharing Economy ist Teil der Forschungskonzeption des Blockchain Competence Center Mittweida (BCCM). Die Bevölkerung erfährt, wie die Blockchain-Technologie in der Anwendung für einen sehr praxisnahen und dringend benötigten Anwendungsfall „Mobilität“ und „Sharing“ genutzt werden kann. Dies erhöht die Attraktivität und die Qualität des Standortes Mittweida und der sich im Umland befindlichen Attraktionen.

Darüber hinaus wird die Region Mittweida zum Kompetenzträger bei der Einführung Blockchain-basierender Mobilitätskonzepte in anderen Regionen mit vergleichbarer Situation.

Die operativen Kosten zum Betrieb von Sharing-Diensten sowie die Akzeptanz durch die Kunden sind kritische Faktoren für eine nachhaltige Etablierung solcher Angebote, insbesondere im ländlichen Raum.¹ Die Blockchain-Technologie soll insbesondere diese Aspekte berücksichtigen und einen signifikanten Mehrwert gegenüber den bisher am Markt verfügbaren Mobilitätsplattformen bieten. Im Rahmen des Forschungsvorhabens werden drei Schwerpunkte im Bereich von Sharing-Economy fokussiert:

1. **Mobilitätsträger werden mit einer Blockchain-basierenden Digitalen Identität ausgestattet.** Damit können die jeweiligen Mobilitätsträger alle relevanten Prozesse selbstständig und dezentral auslösen bzw. regeln und sind so nicht mehr auf eine zentrale Server-Plattform angewiesen. Diese Mobilitätsträger sind zukünftig in der Lage, in der Geschäftswelt autonom zu handeln.
2. **Dieser Ansatz ist durch Blockchain-basierte Digitale Identitäten für Nutzer zu erweitern.** Jene vereinfacht Registrierungs- und Anmeldeprozesse wesentlich, indem die Digitale Identität für Single-SignOn-Prozesse genutzt wird. Das bedeutet, dass Nutzer mit einer einzigen Identität die Dienste verschiedener (Mobilitäts-)Anbieter in Anspruch nehmen können.

3. **Die Rahmenbedingungen für erfolgreiches, Blockchain-basiertes Mobilitäts-Sharing im ländlichen Raum sollen untersucht und konzipiert werden.** Dabei soll eine hohe Auslastung von verfügbaren Ressourcen erreicht und über eine Einbindung von regionalen Partnern als Kilometer- bzw. Minuten-Sponsoren die Akzeptanz von Sharing-Angeboten gesteigert werden.

Ziel ist es, eine dezentrale Plattform „Mobility 4ALL“ in der Region Mittweida zu entwickeln und damit Blockchain-basierte Sharing-Mobility im ländlichen Raum zu etablieren. Nutzern soll langfristig ermöglicht werden, autonome Mobilitätsträger unkompliziert per App auszuleihen.

2. Digitale Identität der Mobilitätsträger

Für eine erfolgreiche Sharing-Mobility werden eine Prozess- und Ressourcenoptimierung mit einer Blockchain-basierten Infrastruktur essentiell sein. Eine besondere Rolle nimmt die Digitale Identität der Mobilitätsträger ein. Diese wird in der Lage sein, aktiv Stillstände zu melden oder gar vorherzusagen und liefert damit die Grundvoraussetzung für eine erhöhte Auslastung der Verkehrsmittel. Für die digitale Identität werden folgende Eigenschaften realisiert:

- **Identität:** Die Mobilitätsträger bekommen eine eigene digitale Identität. Mit dieser sind alle Informationen zum Mobilitätsträger verbunden. So sind für alle Nutzer transparent und unveränderbar, alle wichtigen Daten einsehbar. Das sind zum Beispiel Wartungen, Hauptuntersuchungen, Reinigung, Nutzungsdauer, Fahrleistung sowie Information zu Herstellung und Zulassung.
- **Kommunikation:** Die Mobilitätsträger sind permanent mit dem Blockchain-Netzwerk verbunden und stellen darüber Ihre aktuelle Position, Verfügbarkeit sowie weiter relevante Informationen den anderen Akteuren und Nutzern im Netzwerk zur Verfügung. Über dieses Netzwerk können die Mobilitätsträger aber auch Informationen zum aktuellen Bedarf von Mobilitätsleistungen abrufen oder eigenständig eine Wartung beauftragen und abrechnen.
- **Selbstverwaltung:** Die Mobilitätsträger sind autonom bei der Bereitstellung ihres Mobilitätsangebots und bei der Planung ihrer Auslastung. So sind sie in der Lage, auf sich ändernde Bedingungen zu reagieren und ihre Ladezyklen oder Wartungsintervalle in der Zukunft zu planen.
- **Gedächtnis:** Jeder Mobilitätsträger verfügt über einen vollständiges, durch das Blockchain-Netzwerk gesichertes Logbuch aller bisherigen und für die Zukunft geplanten Ereignisse. Diese sind manipulationssicher und untrennbar mit der Identität des Mobilitätsträgers verbunden. Somit lässt sich dessen gesamte Historie zweifelsfrei nachweisen.

Zwar lässt sich eine solche Digitale Identität auch durch

andere Technologien als der Blockchain, entwickeln, jedoch bietet letztere den Vorteil des unveränderbaren „Digitalen Zwillingings“. Dies bedeutet, dass der in der realen Welt physisch vorhandene Mobilitätsträger einen Repräsentanten in der digitalen Welt besitzt, welcher folgende Eigenschaften aufweist:

- Der digitale Zwilling ist unsterblich: Alle Ereignisse im Lebenszyklus eines Mobilitätsträgers verbleiben in der Blockchain, auch wenn dieser gestohlen oder verschrottet wird.
- Der digitale Zwilling wird gegen Manipulationen unverwundbar: Alle Vorgänge werden protokolliert, sind transparent einsehbar und können nicht manipuliert werden. Es findet kein Datenverlust entlang der gesamten Prozesskette statt.
- Der digitale Zwilling wird über Smart Contracts gesteuert: Der gesamte Service (Nutzung, Wartung, Reservierung, Bezahlung) wird automatisch, autonom und zuverlässig ausgeführt. Alle Aktivitäten im Zusammenhang mit den Buchungsprozessen werden automatisch gemäß den unter den Teilnehmern vereinbarten Geschäftsregeln ausgelöst.
- Der digitale Zwilling sorgt für Authentizität: Die Geschäftslogik ist durch alle einsehbar und für alle gleich. Sie sichert somit die Transparenz über die „Geschäftsbeziehung“.

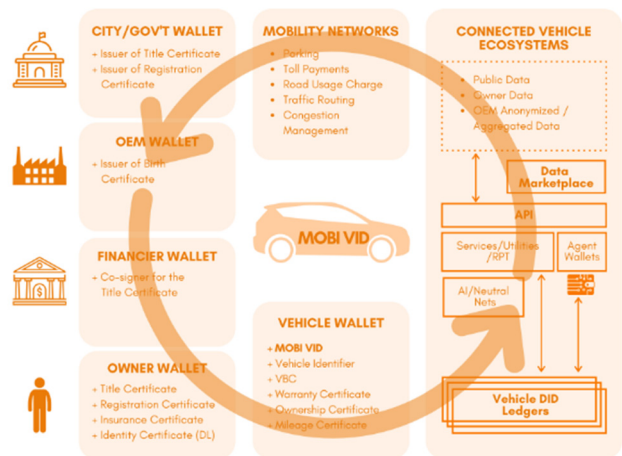


Abbildung 1- Eine Blockchain-basierte Plattform für vernetzte Fahrzeuge²

Für eine weitere Optimierung sollte die digitale Identität der Mobilitätsträger um die **Digitale Identität der Nutzer** erweitert werden. Damit ergeben sich zwei weitere konkrete und wichtige Forschungsfragen, die zu beantworten sind:

- **Interaktion:** Wie müssen die Digitalen Identitäten der Mobilitätsträger und der Nutzer miteinander agieren? Welche Informationen sind dabei auszutauschen?
- **Integration:** Wie ist die digitale Identität der Nutzer in Sharing-Angebote zu integrieren, wenn die Notwendigkeit separater Registrierungen bei unterschiedlichen Anbietern zukünftig nicht mehr bestehen soll?

In diesem Forschungsprojekt liegt der Fokus auf der Umsetzung eines Sharing-Angebotes. Im speziellen die Interaktion zwischen Mobilitätsträger und Benutzer. Die digitale Identität von Mobilitätsträgern kann aber die Grundlage einer ganzen Reihe weiterer Anwendungsfälle sein (siehe Abbildung 2).

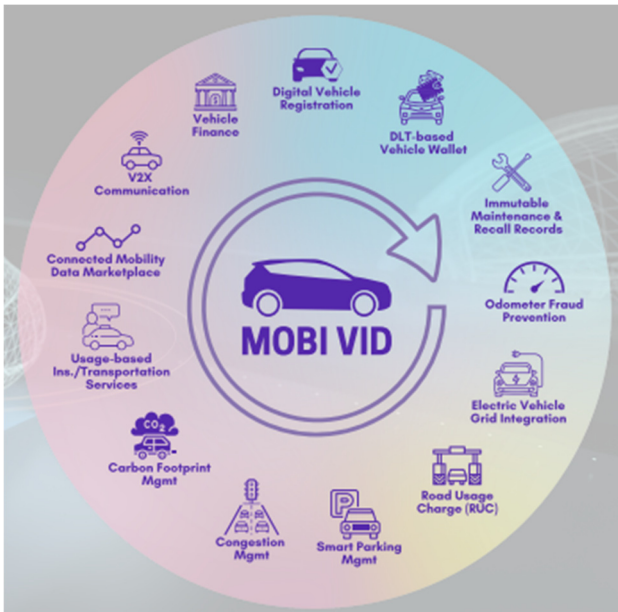


Abbildung 2 - Anwendungsfälle für die digitale Identität von Mobilitätsträgern³

3. Shared Mobility im ländlichen Raum

Sharing Economy bezeichnet eine Form der Ökonomie, in der sich Nutzergruppen von Eigentum, Dienstleistungen und Rechten selbst organisieren und dabei materielle Ressourcen und/oder Arbeitskraft je nach Verfügbarkeit einbringen. Der Begriff der Sharing Economy meint das systematische Ausleihen oder gegenseitige Bereitstellen von Gegenständen, Räumen oder Flächen, insbesondere auch durch Privatpersonen und Interessengruppen. Im Mittelpunkt steht der Gemeinschaftskonsum³. Eine geteilte Nutzung von ganz oder teilweise ungenutzten Ressourcen wird somit ermöglicht. Im Rahmen der Sharing Economy haben sich aufgrund digitaler Innovationen verschiedenste, vorwiegend zentralisierte, plattformgetriebene Geschäftsmodelle entwickelt. Die Blockchain-Technologie hat das Potential, die Infrastruktur für eine neue dezentrale Sharing Economy zu werden. Hohe Gebühren, Vertrauensprobleme und mangelnde Transparenz könnten der Vergangenheit angehören.

Der Forschungsschwerpunkt Shared Mobility im ländlichen Raum soll eine intensivere Auslastung untergenutzter Ressourcen forcieren, um bestehende Mobilitätsmittel wirtschaftlicher und nachhaltiger betreiben zu können. Bestehende Fahrzeugflotten von Unternehmen und Einrichtungen weisen oft hohe Stillstandszeiten auf. Im Rahmen des Projektes gilt es, zu überprüfen, ob diese

bereits vorhandenen Mobilitätsträger während dieser Zeit besser ausgelastet werden können, beispielsweise mit einer zusätzlichen Nutzung über Sharing-Angebote. Folglich wird eine Möglichkeit geschaffen, dass Firmen oder die Kommune ihren eigenen Fuhrpark optimieren können. Gemeinsam mit den potentiellen Nutzern (Unternehmen sowie Institutionen) sowie der Stadt Mittweida werden geeignete Standorte in Mittweida identifiziert, um die Auslastung und Akzeptanz des Sharing Angebotes zu erhöhen. Im Allgemeinen soll emissionsfreie Mobilität genutzt werden.

Einen innovativen Ansatz zur Akzeptanzsteigerung stellen dabei Kilometer- bzw. Minuten-Sponsoren dar. Sponsoren sind Vertreter des Einzelhandels (z.B. Kaufland, Simmel, Lidl, Netto, Bäckereien etc.), Banken oder Freizeit- und Kultureinrichtungen (Kletterwald, Sonnenlandpark Lichtenau). Die regionalen Partner belohnen in Form einer Gutschrift ihre Kunden für die Inanspruchnahme ihrer Dienstleistung. Für die Kilometersponsoren ergeben sich durch eine höhere Mobilität der Bewohner Mittweidas, Potenziale zur Neukundengewinnung und Kundenbindung. Insbesondere in der Anfangsphase mit geringer Bekanntheit können solche Anreize die Attraktivität und das Interesse in der Bevölkerung in Mittweida erhöhen. Zusätzlich bietet der Ansatz der Kilometer- bzw. Minuten-Sponsoren eine Chance für die Wirtschaftsregion Mittweida.

Im Forschungszeitraum wird für ein optimiertes Mobilitätsangebot in der Region Mittweida der Mobilitätsbedarf verschiedener Nutzergruppen erforscht. Diese Erkenntnisse können auch überregional bei anderen WIR!-Projekten wie beispielsweise „Smarte Mobilitätsketten im ländlichen Raum“ eingesetzt werden. Das Gesamtziel im Rahmen der Forschung ist es, einen Werkzeugkasten zu schaffen, der Werkzeuge für das Entwickeln von Blockchain-Anwendungen und Sharing-Modellen enthält. Dabei werden wissenschaftliche Ergebnisse in den Bereichen Sharing-Mobility im ländlichen Raum, Evaluierung von Geschäftsmodellen sowie technische Ergebnisse zur Blockchain-Technologie und Datensicherheit nutzbar gemacht.

Dabei soll u.a. analysiert werden, inwiefern die oft hohen operativen Kosten (insbesondere bei bestehenden Mikromobilitätsanbietern) mittels Blockchain reduziert werden können. Das Potenzial liegt dabei hauptsächlich in der Verringerung von Abrechnungs- und Verwaltungsprozesskosten. Außerdem gilt es zu überprüfen, inwieweit eine Blockchain-basierte Nachweisbarkeit von Wartungsintervallen und somit der Verkehrssicherheit das Vertrauen zwischen Nutzer und Anbieter sowie die Akzeptanz auf Nutzerseite erhöhen kann. Somit wird im Rahmen dieses Forschungsvorhabens nicht nur die Mobilität für den ländlichen Raum etabliert und weitergedacht, sondern auch die Grundlage für das Zusammenwirken von Digitalen Identitäten erforscht.



Abbildung 3 - Die drei Säulen des Forschungsprojektes Mobility 4All (Eigene Darstellung)

4. Regionale Verankerung

Aufgrund des hohen Potentials, die Ergebnisse des Forschungsprojektes im Nachgang in der Praxis für ein Sharing Angebot in Mittweida einzusetzen, werden parallel zu diesem Projekt mögliche Betreibermodelle evaluiert. Hierfür könnte sich zum Beispiel die Rechtsform einer Genossenschaft anbieten. In dieser können die technischen Entwicklungen sowie die Forschungsergebnisse nachhaltig eingesetzt werden und der Bevölkerung und den ansässigen Unternehmen einen Mehrwert bieten. Das Ziel besteht darin, ab 2023 ein wirtschaftlich tragfähiges Sharing-Konzept für Mittweida zu etablieren.

Solch eine Mobilitätsplattform könnte dann jeder Privatperson oder jedem Unternehmen in der Region Mittweida für das Mobilitätsangebot bereitstehen und stetig an die Bedürfnisse angepasst werden.

Mit einem lokal startenden Mobilitätskonzept werden für die Region Mittweida neue Potentiale gehoben. Zum einen aufgrund der technologischen Innovation mit der Digitalen Identität der Mobilitätsmittel, zum anderen durch ein neues und nachhaltiges Geschäftsmodell. Dieses in Mittweida startende Konzept kann auf die Region erweitert und auf andere ähnlich strukturierte Regionen angepasst werden. Allen in Mittweida lebenden Menschen wird die Möglichkeit gegeben, die Mobilität aktiv mitzugestalten und den eigenen Mobilitätsbedarf effizienter zu decken.



Abbildung 4 - Mobility 4All in Mittweida (Eigene Darstellung)

Danksagung

GEFÖRDERT VOM



Bundesministerium
für Bildung
und Forschung



Die Autoren danken dem Bundesministerium für Bildung und Forschung für die finanzielle Unterstützung sowie dem Projektträger Jülich für die Betreuung des Forschungsprojektes „Mobility 4All“ unter dem Förderkennzeichen 03WIR1314.

Literaturverzeichnis

- [1] Man muss privaten Autobesitz teuer machen (<https://www.businessinsider.de/gruenderszene/automotive-mobility/interview-car-sharing-studie-augustin-friedel/>) (aufgerufen am 02.03.2021)
- [2] VEHICLE IDENTITY I - TECHNICAL SPECIFICATIONS (https://dlt.mobi/wp-content/uploads/2021/02/MOBI-VID0003_TS_2019-compressedandlinked.pdf) (aufgerufen am 02.03.2021)
- [3] BLOCKCHAIN FOR VEHICLE IDENTITY BUSINESS WHITEPAPER (https://dlt.mobi/wp-content/uploads/2021/02/MOBI-VID0001_WP_2021-Version-2.0_compressedandlinked.pdf) (aufgerufen am 02.03.2021)
- [4] Gabler Wirtschaftslexikon (<https://wirtschaftslexikon.gabler.de/definition/sharing-economy-53876>) (aufgerufen am 02.03.2021)

Umstieg des ÖPNV auf alternative Antriebstechnologien im Regionalverkehr

Andreas Seltmann

REGIOBUS Mittelsachsen GmbH, Altenburger Straße 52, 09648 Mittweida

1. Einleitung

„Mobilität wird sich in den nächsten zehn Jahren mehr verändern als in vielen Jahrhunderten zuvor“

Andreas Scheuer

Bundesverkehrsminister

Der Pariser UN-Klimagipfel hat die Forderungen und Rahmenbedingungen für eine Nachhaltigkeit und den Umweltschutz klar definiert. So soll bis zum Jahr 2050 eine Klimaneutralität erreicht werden.

Der ÖPNV spielt dabei bei einer umweltschonenden Mobilität eine wichtige Rolle. Der VDV, Verband Deutscher Verkehrsunternehmen, hat dazu Mitte letzten Jahres in einer Tagung die aktuellen Konzepte vorgestellt. Zu diesem Zeitpunkt waren bereits ca. 400 batterieelektrische serienreife Fahrzeuge in Deutschland im Einsatz. Einige Städte, wie Stuttgart, Köln und Wuppertal sind in der Erprobung von Wasserstofffahrzeugen.

Für einen Weg zur Serie arbeitet die Branche sehr eng zusammen. So sind z.B. wichtige Details, wie die Ladeschnittstellen und -infrastruktur zu standardisieren.

Ein wichtiger Treiber zum Umstieg auf alternative Antriebstechnologien ist die „Clean Vehicles Directive“, deren Quoten zu den Mindestbeschaffungszielen ab August 2021 wirken. Dabei müssen bei europaweit aususchreibenden Neubeschaffungen mindestens 22,5 % der Fahrzeuge mit lokal emissionsfreien Antriebstechnologien ausgestattet sein. Die Umsetzung dieser begann bereits in vielen Großstädten, die ihren Fuhrpark auf batterieelektrische oder Wasserstoff-Busse umstellen. Hier verzeichnen sich deutlich Vor- und Nachteile der verschiedenen Konzepte.

Bei der Übernacht- oder auch Depotladung werden die Betriebspausen der Fahrzeuge genutzt, um die Batterien auf dem Betriebsgelände, zumeist konduktiv per Stecker, zu laden. Aufgrund der eingeschränkten Nachademöglichkeiten müssen die Batteriepakete entsprechend der benötigten Reichweiten angepasst werden. Die zu erzielenden Reichweiten sind jedoch sehr beschränkt. Der Aufbau der Ladeinfrastruktur muss der Anzahl der simultan zu ladenden Fahrzeugen in energetischen Aspekten der Energiebereitstellung der Energieversorger angepasst werden.

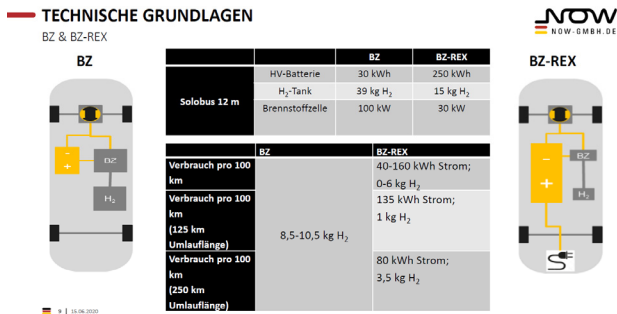
Ein weiteres Konzept ist die Gelegenheitsladung an den Haltestellen. Hier können kleinere Batteriekapazitäten in den Fahrzeugen verbaut werden, was einen erheblichen

Fahrzeugkosten- und Gewichtsvorteil mit sich bringt. Die Fahrzeuge werden mehrmals für einige Minuten an den vorgesehenen Haltestellen entlang der Linienführung geladen. Dies kann konduktiv über ein Pantografen-System oder über elektromagnetische Induktion erfolgen. Dieses Konzept hebt eine Reichweitenbeschränkung auf, sodass die Fahrzeuge einen ganztägigen Einsatz bewältigen können. Die Errichtung der Ladeinfrastruktur ist allerdings strikt an die Linienführung gebunden und erfordert einen immensen Ausbau an Infrastruktur mit erheblichen finanziellen Investitionen. Es müssen an den vorgesehenen Haltestellen die notwendigen Ladesäulen errichtet werden. Die Energienetzbetreiber müssen dort die benötigte Leistung, welche aufgrund der schnellen Ladung der Batterien, jedoch mit wesentlich höheren Strömen, zu Verfügung stellen. Zusätzlich muss zur Batterieconditionierung (Ausbalancierung und Temperierung) eine Ladeinfrastruktur wie beim Depotlader auf den Betriebshöfen installiert werden.

Einige Städte, in denen die Infrastruktur schon teils vorhanden ist bzw. diese ausgebaut wird, setzen Fahrzeuge mit Gelegenheitsladung durch partielle Oberleitungen ein. Diese Fahrzeuge nutzen direkt die elektrische Energie aus der Oberleitung zum Fahren und können gleichzeitig eine Hochleistungsbatterie laden, die es dann ermöglicht in einem begrenzten Radius ohne Oberleitungen die Fahrt fortzusetzen. Der Ausbau dieser Anlagen ist unter städtebaulichen Gesichtspunkten sowie den hohen Anschaffungsinvestitionen für Fahrzeuge und Infrastruktur sehr schwierig.

Die neuesten Entwicklungen der Fahrzeughersteller befassen sich mit Antriebstechnologien der Wasserstoffbrennstoffzelle. Hier sind schon einige Fahrzeugprototypen am Markt, die teils sehr gute Ergebnisse erzielen. Auch hier gibt es verschiedene Konzepte. Die derzeit häufigste Variante ist dabei die Nutzung der dementsprechend groß dimensionierten Brennstoffzelle als Hauptenergiequelle. Diese Fahrzeuge verfügen nur über eine kleine Hochleistungsbatterie (üblicherweise < 30 kWh), die der Speicherung der Rekuperationsenergie und der Abdeckung von Lastspitzen dient. Diese Fahrzeuge werden extern nicht geladen. Eine Ladeinfrastruktur muss somit nicht errichtet werden. Die erforderliche Energie wird nur über die Brennstoffzelle erzeugt. Der Wasserstoffverbrauch dieser Fahrzeuge ist dadurch jedoch relativ hoch. Eine Betankung kann aber an den technisch komplexen Wasserstofftankstellen innerhalb weniger Minuten erfolgen.

Einige Fahrzeughersteller setzen auf ein anderes Konzept. Sie rüsten ihre batterieelektrischen Busse (Batteriekapazitäten > 200 kWh) mit einer leistungstechnisch kleineren Brennstoffzelle als Erweiterung (Range Extender) aus. Diese Fahrzeuge werden z.B. über Nacht im Depot geladen und mit einer geringeren Menge an Wasserstoff betankt. Dieser wird dann genutzt um mit Hilfe der Brennstoffzelle die Batterie während der Fahrt zu laden, um die Reichweite dementsprechend zu erhöhen. Hier ist der Aufbau einer Ladeinfrastruktur und die Errichtung einer Wasserstofftankstelle auf den Betriebshöfen erforderlich.



Quelle: NOW „Leitfaden Einführung Wasserstoffbusse“

Die Einführung jedes der einzelnen Konzepte bedeutet große Veränderungen und Anpassungen an die Betriebsabläufe, Fahrzeugumläufe, Fahrpläne, Fahrzeug- und Servicestrukturen, Fahrzeugbedarfe, Betriebshöfe, Werkstätten und nicht zuletzt für das Personal.

2. Realisierung im Regionalverkehr

Anders als im großstädtischen Nahverkehr, mit z.B. eng getakteten Linien, engen Haltestellenabständen und durchgängigen Angeboten, ist der Regionalverkehr zu betrachten. Hier sind, besonders in den ländlichen Räumen, die Linienlängen und Haltestellenabstände wesentlich größer. Die Angebote erstrecken sich oft nur auf wenige Fahrten. Der Fahrzeugbedarf im Hinblick auf die Gefäßgrößen kann tageszeitabhängig stark schwanken. So kann es in den Früh- und Nachmittagsstunden durch den integrierten Schülerverkehr nötig sein, mit Fahrzeugen größer 12 m zu fahren, wobei in den Zwischenzeiten ein Kleinbus ausreicht. All dies muss bei einem Umstieg auf alternative Antriebe besonders beachtet werden. Wirtschaftlich und technisch ist es momentan nicht darstellbar, das Liniennetz für eine Gelegenheitsladung an den verschiedenen Haltestellen auszurüsten. Gerade im ländlichen Raum wären die dort benötigten Ladeleistungen der Schnellladesäulen durch den derzeitigen Netzausbau nicht zu realisieren. Deshalb kommt derzeit nur eine Depotladestrategie in Frage.



Bild: Ladesäule für zwei Fahrzeuge, Bh VHH, Hamburg Bergedorf

Hier zeigen sich nach den ersten Analysen die derzeit verfügbaren batterieelektrischen Fahrzeuge mit einer maximalen Reichweite von 150 – 200 km bei derzeit gefahrenen Tagesumläufen von ca. 250 – 350 km als nicht geeignet. Die im Einsatz befindlichen Dieselfahrzeuge könnten somit nicht 1:1 ersetzt werden. Die Anzahl der Fahrzeuge müsste sich demnach erhöhen, was wiederum einer Erhöhung der Platzkapazitäten der Betriebshöfe bedeuten würde.

Durch die Notwendigkeit einer Depotladung muss hierzu die Planung und Schaffung der dafür erforderlichen Infrastruktur auf den Betriebshöfen erfolgen. Außerdem müssen die Fahrzeugumläufe an die Reichweiten der Fahrzeuge angepasst werden.

Die Einführung eines Betriebshofmanagementsystems inkl. des Lademanagements ist zwingend notwendig, um einen gesicherten und geregelten Betriebsablauf und Fahrzeugeinsatz gewährleisten zu können. Für jede Linie bzw. für jeden Umlauf muss ein Anforderungsprofil in Hinblick auf Reichweitenbedarf, durchschnittlichen Verbrauch, mögliche Ladezeiten, eventuelle Umleitungen etc. erstellt werden. Der Soll- und Istzustand der Batteriekapazitäten und verfügbaren Reichweiten der Fahrzeuge muss ständig überwacht werden. Dazu sind die Fahrzeuge zwingend mit Datenmodulen und Schnittstellen auszurüsten. Die Daten müssen dann in Echtzeit an die Zentralen übermittelt werden. Abweichungen durch unvorhersehbare Ereignisse, unterschiedliche Fahrweisen des Personals, aber auch unterschiedliche Anforderungen der Jahreszeiten, die z.B. bedeutsam für Heizung und Klimatisierung der Fahrzeuge sind, müssen den Einsatzleitungen sofort angezeigt werden, damit diese entsprechend reagieren und die Umläufe der Fahrzeuge anpassen können. Das Betriebshofmanagement muss in Zusammenarbeit mit dem Lademanagement sicherstellen, dass die Fahrzeuge vor ihrem Einsatz, solange diese noch am Ladesystem angeschlossen sind, vorkonditio-

niert werden. Das bedeutet, dass die Fahrzeuge im Winter vorgeheizt und bei hohen Temperaturen im Sommer vorklimatisiert werden müssen, um diesen Energiebedarf nicht der Kapazität für das Fahren zu entziehen. Ebenfalls müssen besonders im Winter die Batteriepacks auf Betriebstemperatur erwärmt werden, um ihre volle Leistung entfalten zu können.

Ein Lademanagementsystem muss die erforderlichen Reichweiten und die dementsprechend benötigten Ladekapazitäten im Zusammenhang mit den möglichen Lade- und Einsatzzeiten berechnen. Bei gleichzeitiger Ladung mehrerer Fahrzeuge müssen die Ladeströme und Zeiten ausbalanciert werden, um Netzspitzen und eine Überlastung des Netzes bzw. des Netzanschlusses zu vermeiden.

Bei einer Planung zur Erweiterung in Richtung Wasserstofftechnologie erfordert dies einen weiteren immensen Aufwand zur Schaffung der dafür benötigten Infrastruktur, wie z.B. Tankstellen. Diese müssen unter größten Sicherheitsaspekten geplant und errichtet werden.

Eine weitere große Herausforderung ist neben der notwendigen Schaffung der Infrastruktur auf den Betriebshöfen und Fahrzeugabstellflächen hinsichtlich Ladeeinrichtungen, die Planung und Schaffung des erforderlichen Werkstattbedarfes. So muss auch in den Werkstattbereichen die notwendige Ladeinfrastruktur errichtet werden.



Bild: Dacharbeitsstände Betriebshof VHH, Hamburg Bergedorf

Durch den überwiegenden Verbau der Technik im Dachbereich der Fahrzeuge, besteht die Notwendigkeit der Einrichtung von Dacharbeitsständen (siehe Bild aus Hamburg). Der Einbau einer Kran-Bahn erweist sich z.B. bei einem Wechsel von Batteriepacks im Dach von großem Vorteil. Das bedeutet aber auch, dass die Werkstätten eine bestimmte Gebäudehöhe vorweisen müssen. Der Platzbedarf spielt ebenfalls eine bedeutende Rolle. So müssen Sicherheitsabstände zu den Fahrzeugen bei

den Wartungen und Reparaturen eingehalten werden. Havarieflächen sind für verunfallte Fahrzeuge auf den Höfen vorzuhalten. Bestimmte Ausrüstungen und Werkzeuge müssen für einen Umgang mit Hochvoltanlagen zugelassen sein.

Für einen erfolgreichen Umstieg auf alternative Antriebe ist allerdings nicht nur die Infrastruktur und Technik von großer Bedeutung, sondern auch das Personal, welches den täglichen Einsatz sichert. Alle Personale, die mit und an den Fahrzeugen arbeiten, müssen dementsprechend geschult werden. So müssen z.B. auch Reinigungspersonal eine Sicherheitsunterweisung erhalten. Gerade bei dem Fahrpersonal sind Schulungen für eine ökonomische Fahrweise besonders wichtig. Die Werkstattpersonale müssen für Hochvolttechnik und eventuell für den Umgang mit Wasserstoff qualifiziert werden.

3. Prüfung und Planung der RBM

Die REGIOBUS Mittelsachsen GmbH hat das Ziel, eine nachhaltige Mobilität in der Region zu schaffen und dazu gehört auch, diese ressourcenschonend und emissionsarm durch innovative Antriebe anzubieten. So prüfen und planen wir derzeit, wie uns ein Umstieg auf emissionsarme oder emissionsfreie Fahrzeuge unter technologischen und wirtschaftlichen Gesichtspunkten in den nächsten Jahren gelingt.

Ein erster Schritt könnte die Umstrukturierung unseres „kleinen Stadtverkehrs“ in Burgstädt sein. Dieser wird derzeit mit herkömmlichen Dieselfahrzeugen betrieben. Ein Umstieg auf batterieelektrisch angetriebene Kleinbusse ist aufgrund der dort schon teils angebotenen serienreifen Fahrzeuge denkbar. Diese können die Umläufe und die benötigten Reichweiten leisten. Die Schaffung der hierfür erforderlichen Infrastruktur ist durch den geringen Bedarf dieser PKW-nahen Technologie zu leisten.

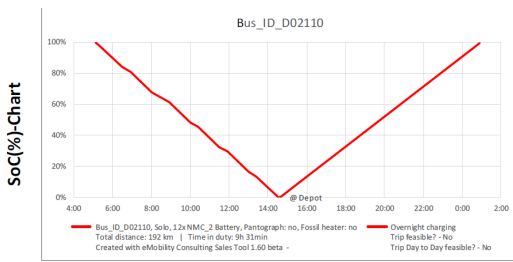
Weitaus schwieriger sehen wir den Umstieg im Regionalverkehr und in unseren größeren Stadtverkehren. Wie schon beschrieben, erreichen die derzeit angebotenen Fahrzeuge die von uns benötigten Reichweiten noch nicht.

Hierzu haben wir gemeinsam mit einem großen Bushersteller eine Machbarkeitsstudie, als erstes für zwei starke Linien und deren Umläufe, durchgeführt.

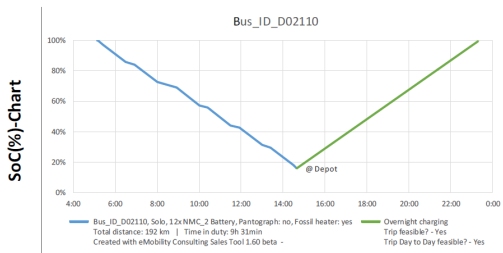
Dazu wurden rein batterieelektrische Fahrzeuge (Solo-Busse 12 m) zur Berechnung mit den jeweiligen Umläufen geprüft. Für das Ladekonzept wurde Ladung im Depot per CCS-Steckverbindung mit einer maximalen Ladeleistung von 150 kW gewählt. Der Start im Depot erfolgt mit voller Batterieladung. Die Gesamtkapazität der gewählten Batterien beläuft sich auf ca. 400 kWh.

Die Auswertung hat folgende Ergebnisse gezeigt:

Beispiele der geprüften Linie 657:



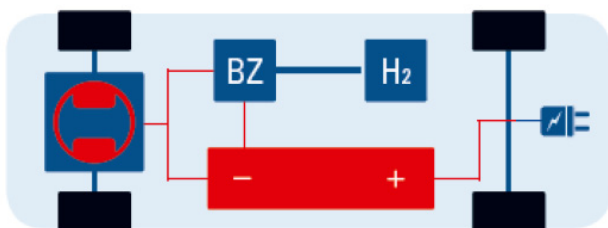
Das Beispiel 1 zeigt die Simulation eines Fahrzeuges ohne fossile Zusatzheizung. Der erforderliche Umlauf von 192 km wird nicht erreicht.



Das Beispiel 2 zeigt die identische Simulation eines Fahrzeuges allerdings mit einer fossilen Zusatzheizung. Hier könnte theoretisch der Umlauf bedient werden. Allerdings wird die Batterie weit entladen, was deren Leistungsfähigkeit und Lebensdauer schadet. Für einen Einsatz wären die Reserven zu knapp bemessen.

Eine Ausrüstung der Fahrzeuge mit größeren Batteriepacks ist aus wirtschaftlichen Gesichtspunkten nicht sinnvoll. Durch den höheren Gewichtsanteil würde man massiv an Fahrgastkapazität einbüßen.

Im Ergebnis der Machbarkeitsstudie haben wir gemeinsam mit dem Fahrzeughersteller festgestellt, dass hier der Einsatz von Fahrzeugen mit Brennstoffzelle als Range Extender vorzugsweise zu integrieren wäre. Diese würde den fehlenden Energiehaushalt kompensieren.



Quelle: NOW „Leitfaden Einführung Wasserstoffbusse“

Diese Fahrzeuge verfügen über eine große Batterie (>200 kWh). Diese wird über Nacht im Depot geladen und fungiert als Hauptenergiequelle. Die Brennstoffzelle sorgt für eine kontinuierliche Nachladung während des Betriebes und verlängert dadurch die Reichweite. Die Batterie nimmt zusätzlich die beim Bremsen durch Rekuperation zurückgewonnene Energie auf.

Nach den ersten veröffentlichten Ergebnissen liegt der Verbrauch dieser Fahrzeuge pro 100 km bei ca. 80 kWh Strom und 3,5 kg Wasserstoff.

Diese Werte und Daten nutzen wir derzeit, um einen genauen Bedarf der Fahrzeuganzahl, der erforderlichen Infrastruktur hinsichtlich Netzanschlusses und Energieverbrauch, Ladetechnik, Wasserstofftankstellen und Wasserstoffverbrauch zu ermitteln.

Die Ergebnisse dieser Daten gilt es dann in einem wirtschaftlichen Konzept zu verankern und gemeinsam mit unserem Gesellschafter einen Investitionsplan zu erarbeiten.

Literaturverzeichnis

- [1] DI Markus Schuster, Mag. Irene Steinacher, DI Christoph Link
Marktübersicht Elektro- und Wasserstoffbusse des Bundesministeriums für Klimaschutz, Umwelt, Energie, Mobilität, Innovation und Technologie Österreich
- [2] Ronny Mühlbach (Daimler AG)
Machbarkeitsstudie eCitaro für REGIOBUS Mittelsachsen vom 16.07.2020
- [3] Martin Schmitz (VDV)
Präsentation Clean Vehicles Directive (CVD)

Application of Green Hydrogen in Mobility Sector

Lorenzo Reyes-Bozo¹, Carlos Fúnez Guerra², Claudia Sandoval Yañez¹,
Guillermo Schaffeld¹

¹ Grupo de Investigación en Energía y Procesos Sustentables, Ingeniería Civil Química, Facultad de Ingeniería, Universidad Autónoma de Chile, Av. Pedro de Valdivia 425, Providencia, 7500912, Chile.

² National Hydrogen Center, Prolongación Fernando el Santo Street, 13500, Puertollano, Spain.

At a global level, different studies disclose that transport systems are responsible for 25% of CO₂ emissions. In the context of sustainable mobility, one of the challenges in the short term is associated with the research and improvement of alternative fuels, which should allow a fast decrease in the generation of greenhouse gases due to sustainable transport means. In this sense, green hydrogen can play a fundamental role. Green hydrogen is the basis for producing synthetic fuels, which can replace oil and its derivatives. Synthetic fuels or e-fuel are hydrocarbons produced from carbon dioxide (CO₂) and green hydrogen (H₂) as the only raw materials. H₂ or e-fuel could be used in many sectors (manufacturing, residential, transportation, mining and other industries). In this study, different applications of hydrogen are evaluated by techno-economic analysis. The main variable that affects the production of hydrogen and its derivatives is the cost of electricity. Considering the renewable energy potential of Chile, it is feasible to develop in Chile the green hydrogen production as an energy vector, which would be technically and economically viable, together with the environmental benefits.

1. Introduction

The density of population and scarce territorial planning of big cities historically have excessive use of private transport at the center of mobility. Under COVID-19 pandemic context, the use of private vehicles has increased to the detriment of public transport due to the risk of contagion perception. With more vehicles on the roads, one of the first consequences is the saturation of public streets by private vehicles with a greater use of fossil fuels and, therefore, air and noise pollution problems, which has a negative impact on the health of inhabitants.

The CO₂ emissions are growing continuously at a global level. For example, between 2007 and 2017, the CO₂ emissions have grown approximately 13.3%, from 28,984 million t to 32,840 million t, respectively [1]. Only the transport sector is responsible around 8.04 Gt CO₂ and the same represents about 25.0% of global CO₂ emissions and requires around 29.0% of global final energy consumption [2]. Then, it is necessary to conceptualize a sustainable mobility model considering different types of people, needs and the specific conditions of each city. The new model will allow people to go from one place to another in a more environment friendly, in an accessible, efficient, safe, and equitable way. One approach is to modify the current mobility management model, prioritizing the design of cities for pedestrian, bicycles, major and minor public transport, leaving shared transport in private vehicles and the one-person use of cars and taxis last.

In the context of sustainable mobility, one of the main challenges in the short term is associated with the research and improvement of alternative fuels, which

will allow a fast decrease in the generation of greenhouse gases due to sustainable transport means. In this sense, green hydrogen can play a fundamental role. Green hydrogen is the basis for producing synthetic fuels, which can replace oil and its derivatives. Synthetic fuels or e-fuel are hydrocarbons produced from carbon dioxide (CO₂) and green hydrogen (H₂) as the only raw materials [3]. The process requires energy, which must be provided from renewable sources (for example, solar and eolic energies). Synthetic fuels allow a significant reduction of CO₂ emissions compared to the emissions generated using fossil fuels. Based on life cycle analysis, the CO₂ emissions could be reduced by over 70%. Synthetic fuels also have a higher energy density compared to batteries, therefore, they could be offering a solution to those applications where electricity does not yet have an answer, for example, air and maritime transport.

Hydrogen as energy carrier or e-fuel could be used in many sectors (for example, manufacturing, residential, transportation, mining, and other industries). In this study, different applications of green hydrogen are evaluated by techno-economic analysis.

2. Methods

2.1 General technical description, cost estimation and sensitivity analysis

Different study cases are reported in this study. The first case discloses the techno-economic analysis of centralized hydrogen generation plant for use in the mobility sector. The second case analyzes the use of hydrogen obtained via electrolysis in railway transport and, the third case, analyzes the green ammonia

production in Chile. All studies consider the production of green hydrogen by renewable energies. The main processes/operational units considered in the production plant evaluated are based on current technologies that were quoted from industrial partners. The main equipments are: Electrolyser, hydrogen storage, hydrogen compressor system, hydrogen dispensers, etc. The cost of these equipments consider prices of the main European manufacturers.

The total estimation of production costs is based on the amount of capital costs (CAPEX) and operating costs (OPEX) for each study case. Capital costs include hydrogen production equipment, storage, compression, ancillary equipment, and civil works. The operational costs include electricity, water, equipment maintenance, staff costs and land leasing costs. The main economic indicators were Net Present Value (NPV), Internal Rate of Return (IRR) and pay-back. These parameters were estimated by standard project evaluation methods [4,5]. Sensitivity analysis was performed with the Oracle Crystal Ball tool to determine the effect of the main variables over the NPV.

3. Results

Chile has a high renewable potential. For example, in the Atacama Desert, the irradiation indexes range from 5 kWh/m² to 12 kWh/m² in winter and summer, respectively [6]. Then, this renewable potential could support photovoltaic and concentrated solar power technologies. Further, Chile has wind power to develop in-shore and off-shore projects. For example, at the Pacific Ocean the wind could reach average speeds around 10 m/s [6]. Then, this renewable potential will be managed to produce green hydrogen.

Nowadays, many countries have proposed different strategies to produce hydrogen and use it in different applications. The first question is related to how this hydrogen will be produced, in centralized or distributed systems. The first option is required to introduce the hydrogen as new energy vector to meet the demand for captive fleets. The centralized hydrogen production plant requires the transportation of H₂ to different hydrogen refueling stations (HRS). Other option is a distributed hydrogen generation system where the hydrogen is produced on-site in the HRS.

A centralized hydrogen production plant of 2 MW PEM electrolyser requires 2.2 MW of electricity and 15 L of raw water/kg hydrogen produced. The electrolyser will produce 863 kg/day of H₂, 3350 kg/day of oxygen and 21.6 MW/day of residual thermal energy.

The total investment cost of a centralized hydrogen plant is around 3,740,000€ and the economic indicators are: NPV 1,272,692€; IRR 14% and payback 9 years (Fig. 1). The analysis of sensitivity shows that the main variables that affect the Net Present Value (NPV) are the electricity cost and the hydrogen production plant size (Table 1).

The transportation sector is a big responsible of greenhouse gas emissions and the road transportation accounted around 70%. In this sense, new technologies based on hydrogen will play a key role. Then, variables such as technology and environmental awareness, financial status and adequate infrastructure will be studied to evaluate the use of green hydrogen. The environmental awareness and hydrogen availability to consumers (appropriate design of hydrogen production plant and hydrogen refueling stations), produce the acceptance of green hydrogen as environmentally friendly fuel [7].

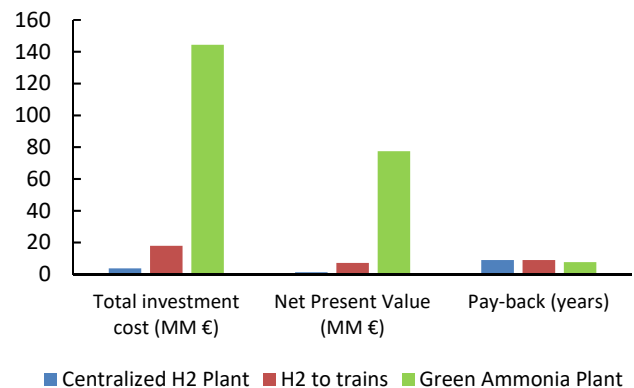


Fig. 1: Total investment and main economic indicators for the different studies cases.

The railway transport sector produces around 4.2% of global CO₂ emissions related to transport sector. The railway transport has social and environmental benefits, since reduces traffic congestion, is safer and the emissions are lower compared to other means of transport [2].

To satisfy the energy demand of a fleet of 20 trains, a 10 MW of polymeric electrolyser system is required. The electrolyser needs 11 MWh of electric energy and 15 L of raw water/kg hydrogen produced. The electrolyser will produce 200 kg/h of H₂, 1600 kg/h of oxygen and 8 MWh of residual thermal energy.

The total investment cost of hydrogen plant to 20 trains is around 17,960,000 € and the economic indicators are: NPV 7,115,391€; IRR 23.5% and payback 9 years (Fig. 1). The sensitivity analysis shows that the main variables that affect the Net Present Value (NPV) are electrolyser capacity or hydrogen refueling station capacity and hydrogen price (Table 1).

Other studies reported in the literature [11-13], show that the hydrogen is an energy vector that contributes to decarbonize with a high energy performance for energy and transport sectors. Particularly, it is possible the use of hydrogen generation and the integration of a hydrogen refueling station for fuel cell hybrid trains, both for passenger and freight trains. These authors reveal a fuel cell efficiency around 50% and a facility efficiency over 50%. Also, attractive financial indicators are disclosed (for example, IRR 19%). These results

confirm that the technical and economic opportunity to use green hydrogen in passenger and heavy-duty transportation.

Table 1: Sensitivity analysis of different study cases (adapted from [8-10]).

| Study Case | Centralized H ₂ Plant | H ₂ to trains | Green Ammonia Plant |
|--------------------------------|----------------------------------|--------------------------|---------------------|
| Variables | Contribution to variance (%) | | |
| Electricity Cost | 62.9 | -10.0 | 96.4 |
| Hydrogen Production plant size | 36.7 | -- | -- |
| Electrolyser Cost | 0.30 | -- | 0.9 |
| Operational hours/year | 0.00 | 5.5 | 1.0 |
| Hydrogen price | -- | 38.8 | -- |
| Electrolyser size | -- | 44.3 | 0.4 |
| Haber-Bosch cycle cost | -- | -- | 1.4 |
| Ammonia price | -- | -- | 0.0 |

Considering an energy transition, other clean fuels must be evaluated, while the hydrogen economy continues to develop. Recently, ammonia has been regarded as a hydrogen storage medium and it could be used as an energy vector. Ammonia is also used as a fertilizer, a chemical feedstock, a clean-burning fuel for transportation, a refrigerant fluid, and a power generator [10].

To produce green ammonia from green hydrogen a 150 MW PEM electrolyser is needed. In this case, the standard Haber-Bosch process is applied. The electrolyser needs 160 MW of electric energy and 15 L of raw water/kg hydrogen produced. The electrolyser will produce 2687 kg/h of H₂, 21582 kg/h of oxygen and 60 MWh of residual thermal energy.

The total investment cost of green ammonia production is around € 144,375,000 and the economic indicators are: NPV € 77,414,525; IRR 17.0% and payback 7.6 years (Fig. 1). The analysis of sensitivity shows that the main variable that affect the Net Present Value (NPV) is the electricity price (Table 1). Considering the latest energy tenders in Chile, renewable energies (mainly, solar and eolic) show the lowest prices. Then, with a low energy price of renewables, the green ammonia production in Chile is feasible with good economic indicators.

The above business cases show the role of hydrogen in the decarbonization in the mobility sector. Notwithstanding the above, the green hydrogen as energy vector could play a different role in many

economic sectors. For example, it could contribute to integrate large-scale renewable plants, as carrier of energy, decarbonize industry, transport and building heating and serve as feedstock to chemical and other industries [13].

4. Conclusions

The main findings are:

- The transportation sector is responsible for around 8.04 Gt CO₂ and the same represents about 25.0% of global CO₂ emissions and requires around 29.0% of global final energy consumption.
- Hydrogen has the potential to be a powerful enabler to different energy requirements because it offers a clean, sustainable, reliable, and flexible option.
- The business cases show that green hydrogen is an energy vector that could play a key role in the mobility sector. From a technical point of view, the green hydrogen is safe, is an efficient energy vector, the infrastructure is growing worldwide, is sustainable with environmental and social benefits. On the other hand, from an economic point of view, the projects of hydrogen production show acceptable economic indicators, such as, Net Present Value (NPV), Internal Rate of Return (IRR) and pay-back.

Acknowledgements

Grants to Dr. L. Reyes-Bozo from Universidad Autónoma de Chile ("Apoyo a Eventos Científicos Nacionales o Internacionales" and "Proyecto VRIP El 04-2020" both from Vicerrectoría de Investigación y Doctorados). Special thanks to the Faculty of Engineering who supplied financial support for this conference.

References

- [1] D.A. Hagos, E.O. Ahlgren, *Applied Energy*, 261 (2020), 114389.
- [2] C. Fúnez Guerra, L. Reyes-Bozo, E. Vyhmeister, J.L. Salazar, M. Jaén Caparrós, C. Clemente-Jul, *International Journal of Hydrogen Energy*, 2020 (*In press*).
- [3] International Energy Agency, *The future of hydrogen*, 2019, pp.199.
- [4] C. Dimian Alexandre, S. Bildea Costin, A. Kiss Anton. *Economic. Computer aided chemical engineering*, 35 (2014), 717-755.
- [5] V. Bartoov, P. Majerk, D. Hrakov, *Procedia Economics and Finance*, 24 (2015), 68-75.
- [6] E. Vyhmeister, C. Aleixendri-Muñoz, J.M. Bermúdez Miquel, J. Pina Moya, C. Fúnez Guerra, L. Rodríguez Mayor, A. Godoy-Faúndez, P. Higuera, C. Clemente-Jul, H. Valdés-González, L. Reyes-Bozo, *Journal of Cleaner Production*, 149 (2017), 999-1010.

- [7] D. Apostolou, S.N. Welcher. *International Journal of Hydrogen Energy*, 2020 (*In press*).
- [8] C. Fúnez Guerra, M. Jaén Caparrós, B. Nieto Calderón, V. Sendarrubias Carbonero, E. Nieto Gallego, L. Reyes-Bozo, A. Godoy-Faúndez, C. Clemente-Jul, E. Vyhmeister, *International Journal of Hydrogen Energy*, 43 (2018), 11793-11802.
- [9] C. Fúnez Guerra, L. Reyes-Bozo, E. Vyhmeister, M. Jaén Caparrós, J.L. Salazar, A. Godoy-Faúndez, C. Clemente-Jul, D. Verastegui-Rayo, *International Journal of Hydrogen Energy*, 45 (2020), 5112-5121.
- [10] C. Fúnez Guerra, L. Reyes-Bozo, E. Vyhmeister, M. Jaén Caparrós, J.L. Salazar, C. Clemente-Jul, *Renewable Energy*, 157 (2020), 404-414.
- [11] F. Piraino, M. Genovese, P. Fragiaco. *Energy Conversion and Management*, 228 (2021), 113650.
- [12] P. Fragiaco, M. Genovese. *Energy Conversion and Management*, 223 (2020), 113332.
- [13] G. Matute, J.M. Yusta, L.C. Correas. *International Journal of Hydrogen Energy*, 44 (2019), 17431-17442.
- [14] Hydrogen Council, *How hydrogen empowers the energy transition*, 2017, pp. 19.

Autonomes und Vernetztes Fahren auf der Schiene - ATLAS-Modellprojekt der HS Mittweida

Christian Schulz, Jan Thomanek
Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

Autonome, fahrerlose Einzelfahrer bei der Bahn können ein Ansatz sein, um wieder mehr insbesondere Güterverkehr von der Straße auf die Schiene zu holen. Hintergrund, Probleme und Lösungsansätze werden aufgezeigt. Dann erfolgt eine Vorstellung der in den letzten Jahren an der Hochschule Mittweida vor allem im Rahmen studentischer Projekte entwickelten Modellanlage zu ATLAS – „Asynchroner Transport-, Logistik- und Automatisierungsmodus auf der Schiene“. Das System wurde bereits auf der IWKM 2018 vorgestellt, seitdem gibt es signifikante Weiterentwicklungen insbesondere im Rahmen studentischer Arbeiten. Diese werden auch im Bezug zur realen Technik des autonomen und vernetzten Fahrens beschrieben und Perspektiven aufgezeigt.

1. Einleitung

Das fahrerlose autonome und vernetzte Fahren auf der Straße ist eines der hochaktuellen technischen Themen, an denen mit Hochdruck gearbeitet wird und das auch in den Medien sehr präsent ist. Es laufen inzwischen weltweit Testfahrten auf öffentlichen Straßen und Autobahnen, und in absehbarer Zukunft sind gewaltige Umwälzungen im Bereich der Mobilität zu erwarten.

Das autonome Fahren auf der Schiene führt dagegen ein Schattendasein, obwohl dort die Bedingungen eigentlich viel günstiger sind. Die Schiene bietet hervorragende Voraussetzungen für einen flexiblen, automatisierten Transport und damit für einen deutlich höheren Anteil am Verkehrsaufkommen, als es derzeit der Fall ist. [1]

Die Thematik wurde bereits auf der letzten IWKM im Herbst 2018 diskutiert [2]. Die grundsätzlichen Überlegungen zum autonomen Fahren auf der Schiene und den dazu notwendigen Paradigmenwechseln im Bahnverkehr sollen im Folgenden etwas gestrafft erneut dargestellt werden. Schwerpunkt sind dann die Fortschritte bei der ATLAS-Demonstrationsanlage der HS Mittweida in den letzten drei Jahren. Es sind in diesem Zeitraum etwa 20 aufeinander aufbauende studentische Arbeiten zu der Thematik entstanden, die das Projekt wesentlich vorangebracht haben. Abschließend sollen Ideen und Visionen zum autonomen und vernetzten Fahren auf der Schiene dargestellt werden.

2. Hintergrund, Probleme und Lösungsansätze

Ein kurzer Blick in die Historie: Die rasante Entwicklung der Eisenbahn etwa seit Mitte des 19. Jahrhunderts war unmittelbar verknüpft mit der industriellen Entwicklung. Die entscheidende Innovation war damals der Antrieb in Form der mobilen Dampfmaschine. Es ergab sich damit die Grundstruktur des Zuges mit Lokomotive und antriebslosen Wagen, die Züge wurden immer schneller, länger und schwerer. Das neue Verkehrsmittel war dem damaligen Fernverkehr mit Pferdewagen haushoch überlegen und setzte sich in atemberaubendem Tempo durch. Es entstand rasch ein dichtes Netz von Eisenbahnstrecken. Das damalige Bautempo kann uns heute

noch in Erstaunen versetzen und zeugt von der Faszination, die von dem neuen Verkehrsmittel ausging.

Seit einigen Jahrzehnten ist eine zunehmende Verlagerung sowohl des Personen- als auch des Güterverkehrs auf die Straße zu beobachten. Die Bahn hat zumindest in Deutschland ihre Rolle als dominierendes Verkehrsmittel längst eingebüßt. Neben anderen sind drei wichtige Ursachen technischer Art zu nennen:

Erstens fahren bei der Bahn nach wie vor hauptsächlich Züge mit Lok und antriebslosen Wagen. Daraus ergibt sich der gesamte Betriebsablauf auf der Schiene, der im Vergleich zur Straße viel zu schwerfällig, langsam und unflexibel ist und gewaltige Stahlmassen bewegt. Zweitens geht die gesamte Steuerungs- und Sicherheitsphilosophie von dem Prinzip „schlaue Strecke – dummer Zug“ aus, auch wenn es hier in den letzten Jahren mit dem ETCS-System eine technische Weiterentwicklung gegeben hat [3]. Auf Grund der langen Bremswege sind riesige Sicherheitsabstände nötig. Damit sind „ausgelastete“ Schienenwege im Vergleich zu einer befahrenen Autobahn praktisch leer (Bild 1). Drittens sind viele Ziele mit der Bahn nicht direkt erreichbar, das bedeutet ein Umsteigen von Personen und ein Umladen von Gütern, beides entfällt beim Verkehr auf der Straße.



Bild 1: Leere Schiene - Volle Straße [4]

Entscheidender Lösungsansatz für mehr Flexibilität ist der Übergang zu selbstfahrenden, fahrerlosen Einzelwagen (Einzelfahrer) und damit der Verzicht auf die Bildung von Zügen und einen Fahrplan. Ein Antrieb ist nichts Besonderes mehr, und es gibt deshalb keinen technisch vernünftigen Grund, an der Grundstruktur des Zuges festzuhalten. Zum anderen sind die Informations-, Kommunikations-, Automatisierungs- und Sensortechnik soweit fortgeschritten und die Bedingungen für ein automatisiertes Fahren auf der Schiene so günstig, dass ein Fahrer keineswegs notwendig ist.

Wenn auf eine Zugbildung verzichtet wird, können die Wagen wesentlich leichter gebaut werden. Damit sinkt der Energieverbrauch beim Beschleunigen, und die Bremswege verkürzen sich. Verringerte Achslasten bedeuten geringere Ansprüche an den Schienenweg und an Brücken und damit unmittelbar geringere Kosten.

Beim Antrieb könnte vieles von den Entwicklungen für Elektroautos übernommen werden, wobei die technischen Bedingungen auf der Schiene wesentlich günstiger sind als auf der Straße. Die Reibung ist geringer, und es bestehen ideale Möglichkeiten für eine 100%ige vorausschauende Fahrweise mit effizienter Energierückgewinnung beim Bremsen – das geht auf der Straße nur ansatzweise.

Auch die Steuerung der Fahrzeuge kann vieles von dem übernehmen, was gegenwärtig für das autonome Fahren auf der Straße entwickelt und erprobt wird. Das betrifft sowohl die Sensorik (Kamera, RADAR, LIDAR, Ultraschall) als auch die Kommunikation der Fahrzeuge untereinander und mit der Infrastruktur. Wichtig ist dabei, dass die Fahrzeuge untereinander eine Art verteilte Steuerung bilden. Sie wählen ihre Strecke und steuern sie an (insbesondere die Weichen) und nicht umgekehrt, also ein Wechsel zum Prinzip „schlaue Fahrzeuge – dumme Strecke“.

Der gesamte bisherige Betriebsablauf der Bahn wäre grundlegend umzugestalten. Dabei gibt auf Grund des relativ abgeschlossenen Schienensystems weniger Unwägbarkeiten und es ist keine permanente Lenkung der Fahrzeuge erforderlich. Hauptpunkte sind die Abstandssteuerung und Kollisionsvermeidung, das Weichenstellen und die kooperative Fahr- und Streckenplanung.

Eine unabhängige vollständige Überwachung der Strecke kann mit Kamerasystemen und Bildauswertung erfolgen. Damit wäre redundant zu den vorliegenden Informationen der Positionen und Fahrzeugtrajektorien eine weitere Sicherheitsebene möglich, die auch unerwartete Hindernisse und andere Risiken behandelt. Der jetzige Bahnbetrieb geht von einem abgeschlossenen System ohne Fremdeinflüsse aus und kennt als Sicherheit gegen Störungen vor allem die Aufmerksamkeit des Lokführers. Bei unerwarteten Hindernissen auf der Strecke kommt es immer wieder zu schweren Unfällen.

3. ATLAS-Demonstrationsanlage - Überblick

ATLAS ist in diesem Kontext die Abkürzung für „Asynchroner Transport-, Logistik- und Automatisierungsmodus auf der Schiene“ und auch die Bezeichnung einer Demonstrationsanlage, die im Rahmen studentischer Projekte 2006/2007 entstand ([5] [6]) und nach einer längeren Pause seit etwa drei Jahren weiterentwickelt wird. Ziel dieser Anlage ist es, die genannten Lösungsansätze und Prinzipien zu demonstrieren und zu testen. Als Grundlage wurde das LGB-System [7] gewählt, das auf Grund der Spurweite von 45 mm und den entsprechend großen Fahrzeugen entspannte Bedingungen für eigene elektronische Komponenten bietet. Das Atlas-System besteht aus den Gleismodulen und den Fahrzeugen.

Die Gleismodule sind ein-, zwei- oder dreigleisig. Sie haben eine Basislänge von 1800 mm, daneben gibt es Kurven und kürzere Ausgleichsmodule sowie Endmodule mit Puffern. Insgesamt gibt es gegenwärtig 48 Module mit einer Streckenlänge von über 120 m. Sie enthalten 42 Weichen und 155 Balisen, die elektronischen „Kilometersteine“ [8]. Mit diesen Modulen kann relativ schnell eine große Anlage variabel aufgebaut werden. Einen Eindruck von einer aufgebauten Anlage vermittelt Bild 2:



Bild 2: ATLAS-Teilanlage

Die vorhandenen 6 Eigenbau-Fahrzeuge [6] basieren auf einem LGB-Antriebsmodul. Sie werden derzeit massiv überarbeitet, es erfolgt dabei eine konsequente Trennung in low-level- und high-level-Ansteuerung. Dabei wird die low-level-Steuerungselektronik neugestaltet (siehe 4.). Die Positionserkennung, Abstandssicherung und die Sonderfunktionen Licht und Sound werden neu gestaltet. Hinzugefügt werden eine Weichenansteuerung vom Fahrzeug aus und eine „Handsteuerung“ über bluetooth und Smartphone-APP. - Die high-level-Ebene gab es bisher noch nicht. Hier geht es um die übergeordnete Geschwindigkeits- und Richtungssteuerung, die Vernetzung der Fahrzeuge untereinander und Szenarien des kooperativen Fahrens (siehe 5.).

4. Gleismodule und Low-Level-Ansteuerung

Die Gleismodule haben als Basis-Infrastruktur folgende Funktionen zu erfüllen:

- Stromversorgung aller Komponenten
- Positionsbestimmung über Balisen
- Richtungssteuerung über Weichen

In der Vorgängeranlage entsprachen **Stromversorgung** und zentrale Weichenansteuerung der üblichen Modellbahntechnik mit DCC-Protokoll [5]. Hier erfolgte eine konsequente Änderung, die Gleise führen nur noch 24V-Gleichspannung zur Versorgung aller Komponenten.

Die Balisen als „elektronische Kilometersteine“ dienen der eindeutigen **Positionsbestimmung** der Fahrzeuge. Es sind im ATLAS-System kleine, unter den Schwellen eingebaute Platinen [9] mit einem Infrarot-Sender, der im 7-ms-Zyklus einen Code ausgibt (Bild 3). Dieser kann von passierenden Fahrzeugen empfangen werden, dazu befinden sich an beiden Fahrzeugenden die zugehörigen IR-Empfänger.

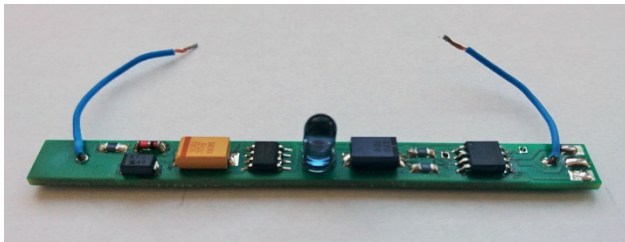


Bild 3: Infrarot-Balise im ATLAS-System

Der IR-Sendecode verwendet eine Trägerfrequenz von 56 kHz und eine Manchester-Codierung für die individuelle 10-Bit-Balisennummer mit Startbit und Paritätsbit, siehe Bild 4. Das Signal wird von einem kleinen ATtiny25-Mikrocontroller erzeugt, der die IR-Sende-LED ansteuert. Jedes Gleismodul enthält mindestens zwei Balisen, die größeren deutlich mehr. Der typische Abstand ist etwa 0,5..1,0 m. Da der Abstand zwischen Balise und IR-Empfänger im Fahrzeug nur 30 mm beträgt, wurde die IR-Sendeleistung sehr gering gehalten.

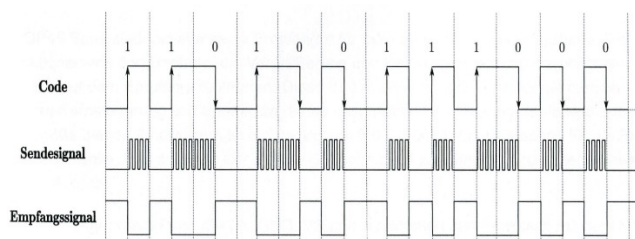


Bild 4: IR-Codierung des Balisensignals [9]

Als Empfänger im Fahrzeug wird der TSOP4856 verwendet [10], der für die 56-kHz-Trägerfrequenz ausgelegt ist. Während einer Überfahrt mit der maximalen Fahrzeuggeschwindigkeit von etwa 1 m/s = 1 mm/ms wird der Balisen-Code 5...8x sicher empfangen, so dass eine verlässliche Positionsbestimmung erreicht wird. Die Zwischenpositionen können dann anhand der Fahrzeuggeschwindigkeit interpoliert werden.

Die **Richtungssteuerung** der Weichen erfolgt vom Fahrzeug aus über ein IR-Signal mit 40 kHz Trägerfrequenz zur Vermeidung von Störungen durch die Balisen. Es können bis zu vier Fahrstraßen auf dem aktuellen Gleismodul geschaltet werden, codiert über ein Tastverhältnis 1:4, 2:3, 3:2 oder 4:1 im 5-ms-Zyklus [11].

Das IR-Signal wird über TSOP4840 [10] empfangen, die jeweils in den Einfahrten von Gleismodulen mit Weichen in einer Schwelle angeordnet sind. In die Gleismodule ist eine Elektronik auf Basis des Mikrocontrollers M16C eingebettet, die die Signale auswertet und die zur entsprechenden Fahrstraße gehörenden Weichen ansteuert. Die elektrische Ansteuerung der LGB-Weichenantriebe verlangt eine H-Brücke, da umgepolt werden muss. Es wurde der ZXBM5210 verwendet [12].

Bild 5 zeigt das Gleismodul 332530 mit 7 Balisen (blau), 5 IR-Empfängern (rot) und 5 Weichen als Beispiel. Es gibt 1, 2 oder 4 Fahrstraßen je nach Einfahrtgleis [13].

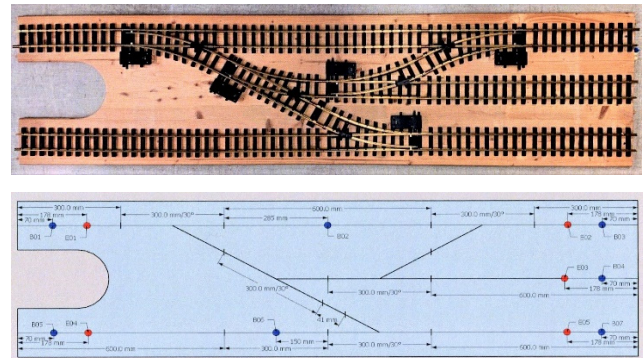


Bild 5: Gleismodul 332530; blau: Balise, rot: IR-Empfänger [13]

Bei Einfahrt in ein Weichenmodul wird die entsprechende Balise detektiert und dann das IR-Signal für die Fahrstraße gesendet. Wichtig ist dabei der Abstand zwischen IR-Empfänger und erster Weiche der Fahrstraße, die ja gestellt sein muss, wenn das Fahrzeug sie erreicht. Messungen [14] ergaben etwa 30 ms für das Weichenstellen (Bild 6), dazu kommen etwa 20 ms für die sichere mehrfache Detektion des IR-Empfangssignals. Das bedeutet einen Fahrzeugweg von etwa 50 mm bei 1 mm/ms. Der Mindestabstand zwischen IR-Empfänger und Weiche liegt bei allen Modulen über 300 mm, somit besteht eine gute Sicherheitsreserve.

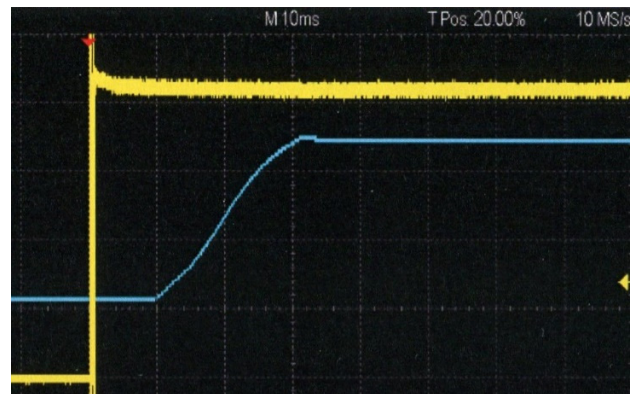


Bild 6: Weichenantrieb; gelb: Schaltsignal, blau: Bewegung [14]

Alle **Gleismodule** sind mit ihrer Topologie, ihren Abmessungen, Balisenpositionen und -nummern sowie den Weichenpositionen und möglichen Fahrstraßen in einer Datenbank systematisch erfasst. Das bildet eine wesentliche Basis für das unter 5. beschriebene statische Umfeldmodell.

Die **Fahrzeuge** (Bild 7) haben als Kern der low-level-Steuerung einen M16C-Mikrocontroller [15], der alle Sensoren auswertet und alle Aktoren ansteuert. Es gibt jeweils vorn und hinten am Fahrzeug einen Infrarotsender für die Fahrstraßen und einen Infrarotempfänger für die Balisensignale. Außerdem ist ein optischer SHARP-Abstandssensor in beide Fahrrichtungen vorhanden. Die Fahrzeuge haben vorn und hinten jeweils zwei Mehrfarb-LEDs zur Signalisierung von Richtung und Fahrstraße, und es ist ein Beeper für beliebige akustische Signale vorhanden. Der Motor wird über ein H-Brücke in Geschwindigkeit und Richtung angesteuert.



Bild 7: ATLAS-Fahrzeug mit IR-Abstandssensor und LEDs

Auf der low-level-Ebene sind als autarke Funktionen geplant: eine Abstandsregelung oder ein Notstopp aufgrund des Abstandssensors, das Einschalten des IR-Fahrstraßensignals bei Überfahren des IR-Empfängers im Gleis sowie eine Geschwindigkeitsregelung mit Ausgleich der Fehler durch Spannungsabfall im Gleis.

Die high-level-Ansteuerung für Fahrziel, Geschwindigkeit und die Sonderfunktionen ist in zwei Varianten geplant. Zum einen laufen die Arbeiten zum autonomen und vernetzten Fahren (siehe 5). Die Funktionalität kann aber auch durch den Menschen übernommen werden, mit dem Nebeneffekt eines höheren Spielwertes der Anlage. Nur automatisch Fahren ist langweilig, ein Mischbetrieb durchaus reizvoll. Dazu wird über eine bluetooth-Schnittstelle ein UART-Protokoll an den M16C übertragen und es entstand eine entsprechende Smartphone-APP [16]. Diese hat einen Kinder-Spielmodus mit den Fahrstraßen A, B, C und D, den Sonderfunktionen F1, F2 und F3, Notstopp, Beep, Vor/Rück und Geschwindigkeitssteuerung ähnlich einer Modellbahn. Es gibt auch einen Entwicklermodus mit zusätzlichem Zugriff auf die internen Daten wie UART-Buffer, Balisennummern, bluetooth-Feldstärken und Verbindungsqualität.

5. High-Level-Plattform und Umfeldmodell

Selbstfahrende Fahrzeuge benötigen für die Berechnung autonomer Fahrfunktionen eine entsprechende leistungsfähige aber auch energieeffiziente Hardwareplattform. Um den Anforderungen an funktionelles und sicheres KI-gesteuertes Fahren zu genügen, wird das ATLAS-Fahrzeug um eine High-Level-Plattform auf der Basis des NVIDIA® Jetson™-Systems erweitert. Das Jetson Nano-Modul repräsentiert einen kostengünstigen KI-Computer, um Daten hochauflösender Sensoren zu verarbeiten und neuronale Netze zu betreiben unter Berücksichtigung der notwendigen Energieeffizienz eingebetteter Systeme. Die strikte Trennung zwischen den autonomen Fahrfunktionen und den Low-Level-Funktionen zur Motorsteuerung und Gleiskommunikation ermöglicht die Wiederverwendung bereits entwickelter Hard- und Software.

Die Systemarchitektur des neuen ATLAS-Fahrzeugs besteht damit im Kern aus zwei Komponenten (Bild 8):

- High-Level-Komponente zur Verarbeitung und Fusion der Sensordaten, Lokalisierung innerhalb der Gleisanlage, Berechnung der Fahrroute und Kommunikation mit anderen Fahrzeugen sowie Infrastrukturelementen
- Low-Level-Komponente zur Ansteuerung der Motoren, Kommunikation mit den Gleismodulen für Lokalisierung und Weichensteuerung

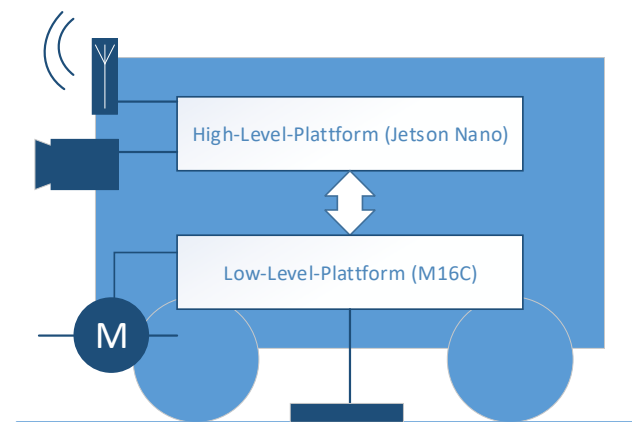


Bild 8: Systemarchitektur ATLAS-Fahrzeuge

Die Kommunikation zwischen High- und Low-Level-Plattform erfolgt über eine serielle Schnittstelle (UART), wobei zyklisch im 10 ms-Raster ein Speicherabbild u.a. mit folgenden Informationen übertragen wird:

Low- → High-Level-Plattform:

- Zuletzt überfahrene Balisennummer
- Motoransteuerung IST

High- → Low-Level-Plattform:

- Motoransteuerung SOLL
- Weichenansteuerung
- Akustisches/Optisches Signal

Grundlage aller automatisierter Fahrfunktionen bildet unter Anwendung verschiedener Sensortechnologien eine hinreichend detaillierte Kartierung des Fahrzeugumfeldes. Ein solches **Umfeldmodell** beinhaltet eine vollständige Abbildung aller Objekte, Hindernisse und Freiräume. Das Umfeldmodell dient als Datenbasis für die nachfolgenden Algorithmen zur Situationsanalyse und Regelung.

Mit einem einzigen Sensor ist es allerdings unmöglich, eine Umgebung komplett zu erfassen. Im ATLAS-Fahrzeug kommen neben Infrarotsensoren für die einfache Abstandserkennung auch Kamerasysteme für detaillierte Hinderniserkennung zum Einsatz. Des Weiteren sind die ATLAS-Fahrzeuge zur Kommunikation mit anderen Fahrzeugen mit einer WLAN-Funkschnittstelle ausgerüstet. Dadurch ist es möglich, Informationen zur Verkehrssituation, die außerhalb des Sichtbereichs der Board-Sensoren liegen, zu erhalten. Dabei werden keine Unicast-Botschaften versendet, sondern die Kommunikation erfolgt im Broadcast-Betrieb. Dabei versenden die Fahrzeuge mit einem Raster von 1-10 Hz sogenannte Beacons, welche die aktuelle Position, Fahrtrichtung sowie die Fahrzeug-Identifikation beinhalten. Alle Fahrzeuge empfangen dies Beacon-Frames und ergänzen die Inhalte ihres eigenes Umfeldmodells mit den Daten der empfangenen Botschaften.

Analog zu den Umfeldmodellen autonomer Straßenfahrzeuge, welche neben den dynamischen Objekten und Hindernissen auch statische Kartendaten zur Identifikation möglicher Fahrstraßen beinhalten, muss im Umfeldmodell der ATLAS-Fahrzeuge der aktuelle Gleisplan abgebildet sein. Basis dieses Umfeldmodells bildet dabei ein fahrzeugunabhängiges kartesisches Koordinatensystem (Metrisches Weltkoordinatensystem). Da die ATLAS-Fahrzeuge sich auf einer Ebene bewegen, kann die z-Koordinate vernachlässigt werden. Der Koordinatenursprung kann willkürlich festgelegt werden. Hinsichtlich der Auflösung des Umfeldmodells ist ein Kompromiss aus Speichernutzung und notwendiger Genauigkeit zu finden. In der aktuellen Umsetzung wurde eine Auflösung von 10 mm gewählt, was bei einer Ausdehnung der Gleisanlage von 40 m x 20 m und einer Zelltiefe von 8 Bit insgesamt 7,63 MB Speicherplatz bedeuten. Eine jede Zelle mit der Ausdehnung 10 x 10 mm trägt bitcodiert die folgenden statischen Informationen zum Gleisplan:

- Vorhandensein Gleis
- Vorhandensein Balise und Balisen-ID
- Vorhandensein Weiche

Diese statischen Gleisplaninformationen werden in einer ersten Ausbaustufe über eine Konfigurationsdatei in die High-Level-Plattform der ATLAS-Fahrzeuge geladen. In einer zukünftigen Umsetzung sollen die Fahrzeuge den gegebenen Gleisplan selbst erlernen und somit das Umfeldmodell entsprechend ausfüllen. Anhand der vom Fahrzeugcontroller (Low-Level-Plattform) enthaltenen

Informationen zu passierten Balisen und der ermittelten Fahrgeschwindigkeit kann die eigene Position (Ego) innerhalb des Umfeldmodells berechnet werden. Des Weiteren werden mittels Bordsensorik sowie mittels der empfangenen Positionsdaten weiterer Fahrzeuge die dynamischen Informationen zum aktuellen Verkehrsszenario in der Gleisanlage komplettiert (vgl. Bild 9). Die eigene sowie die Positionen gegebenenfalls weiterer Fahrzeuge bilden die Grundlage für die High-Level-Fahrfunktionen wie Pfadplanung und Manöverdurchführung.

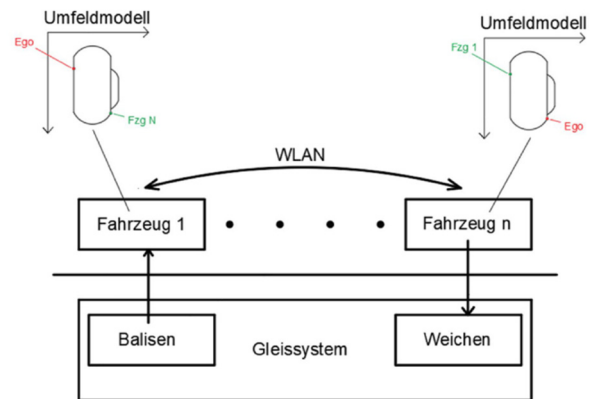


Bild 9: Visualisierung der Umfeldmodelle mit statischen (Gleisplan) und dynamischen (Fahrzeugpositionen) Informationen [17]

6. Reale Bahn – Überlegungen und Visionen

Technisch erscheint das Gesamtkonzept des Einzelfahrers ohne weiteres machbar, alle notwendigen technischen Komponenten von der Sensorik über die Antriebskonzepte und die Telematik-Infrastruktur bis hin zum Echtzeit-Management großer Datenbanken existieren. Es fehlt jedoch eine umfassende, mindestens europaweit durchgesetzte Gesamtlösung. – Es ist nicht möglich und auch nicht notwendig, den gegenwärtigen Bahnbetrieb mit seinen Schienenfahrzeugen, der Sicherheitstechnik und den Prinzipien der Verkehrsführung schlagartig „abzuschalten“. Über eine Integration neuer Lösungen und die Koexistenz von herkömmlichen Zügen und Einzelfahrern muss intensiv nachgedacht werden.

Die Organisationsstruktur der Bahn könnte sich an die im Straßenverkehr anlehnen. Das bedeutet ein öffentlich unterhaltenes Schienennetz einschließlich der notwendigen Telematik-Infrastruktur sowie ein vorwiegend privat bewirtschaftetes rollendes Material für den Personen- und Güterverkehr. Es ließe sich ein präzises und gerechtes Abrechnungssystem einführen, da die jeweilige Nutzung des Schienennetzes detailliert erfassbar ist.

Es ist notwendig, eine grundlegend neue Sicherheitsphilosophie für den Schienenverkehr zu entwickeln. Eine wesentliche Komponente sollte die Kamera-Komplettüberwachung der Strecken mit Bildauswertung werden. Dabei muss ein Nebeneinander von „klassischem Zugverkehr“ und Einzelfahrern für viele Jahre möglich sein.

Die „letzten Meter“ von der Schiene bis zum Ziel müssen effizient überbrückbar sein. Entweder kann der Einzelfahrer selbständig die Schiene verlassen, oder es gibt ein Shuttle-System zum Kurzstrecken-Transport der Schienenfahrzeuge auf der Straße. Bahnhöfe sind nicht notwendig, sie reduzieren sich auf diese Schnittstellen zwischen Schiene und Straße.

Es wäre intensiv darüber nachzudenken, ob das Schienennetz neben dem Transport nicht auch andere Funktionen übernehmen kann, etwa:

- Schienennetz als komplexes Logistikzentrum mit Lagerhaltung und exakter „just-in-time“-Anlieferung (genauer als auf der Straße); damit verbunden wäre eine vernünftige Nutzung der heute praktisch brachliegenden Rangier- und Güterbahnhöfe und der Nebenstrecken
- Produktion auf der Schiene sowie Firmen, die flexibel ihren Standort wechseln und zu den Aufträgen oder auch zu den Arbeitskräften fahren (Baufirmen, „fahrender Schlachthof“ statt tagelanger Tiertransport ...)
- Urlaub, Ausflüge (statt Bus), Schulfahrten mit rollender Jugendherberge und mitgeführtem Klassenraum, Ausstellungen, mobile Tagungen, rollende Impfzentren ...

Die Bahn hat das Potential auf einen wesentlich größeren Anteil am Personen- und insbesondere am Güterverkehr, wenn die „Erblasten“ des 19. Jahrhunderts überwunden werden und ein Paradigmenwechsel stattfindet. Interessant ist auch die weitere Nutzung des ausgedünnten und verfallenden Nebenstreckennetzes, da die Anforderungen an die Strecke bei wesentlich geringeren Achslasten viel unkritischer sind.

Vielleicht können die dargestellten Gedanken zu einer Erneuerung der Bahn beitragen, und die ATLAS-Anlage kann dazu nicht nur technisch, sondern auch mit ihrem werbewirksamen Spielwert ihren Teil liefern.

Danksagung

Die Autoren bedanken sich bei den beteiligten Studentinnen und Studenten, die mit ihren Abschlussarbeiten dieses Projekt vorangebracht haben und weiter voranbringen, auch unter schwierigen Corona-Bedingungen.

Literaturverzeichnis

- [1] R.H. Rauschenberg: Potentiale für die Verringerung der externen Effekte des Verkehrssektors durch einen dezentralisierten und automatisierten Gütertransport der Bahn; Goethe-Universität Frankfurt/Main; Dissertation 2006.
- [2] C. Schulz: ATLAS-Demonstrationsanlage für autonomes Fahren auf der Schiene; Scientific Reports 3,2018 ISSN 1437-7624 S. 192-195
- [3] H. Schilling: Trainguard – ETCS-Lösungen für Europa; Vortrag auf der 18. IWKM der Hochschule Mittweida am 10.11.2006, Tagungsgruppe Automatisiertes Fahren auf der Schiene.
- [4] ADAC-Motorwelt 8/2001
- [5] T. Krauß: Versuchs- und Demonstrationsanlage für Projekt ATLAS; Hochschule Mittweida, Fachbereich Informationstechnik & Elektrotechnik; Diplomarbeit 2006
- [6] R. Schneider: Komponenten für Projekt ATLAS; Hochschule Mittweida, Fachbereich Maschinenbau, Diplomarbeit 2007
- [7] Firmen-WEB-Seite zu LGB unter www.lgb.de
- [8] Z. Hou: Gleisaufbau und IR-Verkabelung der ATLAS-Gleisanlage; Hochschule Mittweida, Fakultät Ingenieurwissenschaften; Bachelorarbeit 2020
- [9] X. Ma: Lokalisierung von ATLAS-Fahrzeugen über Infrarot-Balisen; Hochschule Mittweida, Fakultät Ingenieurwissenschaften; Bachelorarbeit 2019
- [10] Vishay: TSOP22., TSOP24., TSOP48.,TSOP44 IR Receiver Modules for Remote Control Systems; Document Number 91000; 02-Oct-12
- [11] Y. Wang: Prototyp für ATLAS-Weichen-ansteuerung über Infrarot-Schnittstelle; Hochschule Mittweida, Fakultät Ingenieurwissenschaften; Bachelorarbeit 2019
- [12] Diodes Incorporated: ZXB5210 Reversible DC Motor Drive with Speed Control; Document Number DS36765 Rev. 1-2; December 2013
- [13] X. He: Gleisaufbau und Balisen-Verdrahtung der ATLAS-Gleisanlage; Hochschule Mittweida, Fakultät Ingenieurwissenschaften; Bachelorarbeit 2020
- [14] T. Shen: Funktionsmuster für ATLAS-Weichen-ansteuerung über Infrarot-Schnittstelle; Hochschule Mittweida, Fakultät Ingenieurwissenschaften; Bachelorarbeit 2018
- [15] Mitsubishi Electric Corporation: User's Manual M16C/62 Group Rev.C2; First Edition July 2000
- [16] T. Mahfoud: App zur Ansteuerung von ATLAS-Fahrzeugen über Bluetooth; Hochschule Mittweida, Fakultät Ingenieurwissenschaften; Masterarbeit 2020
- [17] Y. Xu: Spezifikation und Implementierung eines Fahrzeugumfeldmodells im Rahmen des ATLAS-Projektes; Hochschule Mittweida, Fakultät Ingenieurwissenschaften; Bachelorarbeit 2019
- [18] K. Jbaili: Implementing and evaluating a tracking-by-detection algorithm for a camera monitoring system of ATLAS; Hochschule Mittweida, Fakultät Ingenieurwissenschaften; Masterarbeit 2020

Implementing and Evaluating a Tracking-By-Detection Algorithm for a Camera Monitoring System

Khaled Jbaili, Jan Thomanek

Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

Abstract

In this paper, we designed, implemented, and tested a special surveillance camera system based on a combination of classical image processing algorithms. The system's sub-objective consists of tracking experimental vehicles driving on a defined trajectories (Rail) in real time. Furthermore, it analyzes the scene to collect additional vehicles & rail-related information. The system then uses the gathered data to reach its main objective which confines oneself in independently predicting vehicles collision. Consequently, we propose a hybrid method of detecting and tracking ATLAS-vehicles efficiently. To detect the vehicle at the beginning of the video, periodically every n -frame, and in the case where the tracked vehicle has been lost, we used Histogram Back-Projection. By contrast, Kernelized correlation filter is used to track the detected vehicles. Combining these two methods provides one of the best trade-offs between accuracy and speed even on a single processing core. The proposed method achieves the best performance compared with three different approaches on a custom dataset.

1. Introduction

There is no doubt that Object Detection has always been a key technology behind many applications like video surveillance, image retrieval systems and Advanced Driving Assistance Systems (ADAS). In addition to this, it is considered as a foundation of several tasks, such as instance segmentation, image captioning, and object tracking. To understand object detection problem, it necessary to define it more formally. In the foremost place comes object recognition which is considered as a general term that describes collection of related Computer Vision (CV) tasks and involves identifying objects in digital photographs starting from image classification and ending in object segmentation. Image classification involves predicting the class of one object in an image, whereas Object localization refers to identifying the location of one or more objects in an image. Object detection combines these two tasks in which it localizes and classifies one or more objects in the image. The previous statement can be summarized with the following question: "What Objects are where?"

Going back to object detection where some input data such as a camera image or Lidar point cloud is given. From this one first needs to determine how many objects of interest are present. For instance, if one is trying to detect vehicles and predestines in a traffic situation, then for each detected object, one determines the object type and the object shape.

Object type refers to the class of the object and object shape represents the enclosure of the detected object of interest or detection hypothesis. There are many detection hypotheses used for labeling a detected object. In most cases, the object shape is determined by a bounding box (bbox) in 2D or 3D.

Another research field which is most salient and explored amongst several CV disciplines and plays an important role in this work is Visual Object Tracking (VOT). Object Detection as defined previously involves determining semantic

existence of object instances of a specific class in an image or every video frame, e.g., cars, boats, and humans.

In contrast, VOT can be defined as the estimation and prediction of object trajectory in a video frame as it moves from one frame to another based on dynamics. In other words, the use of dynamics, such as object's velocity, direction, location history, and similarity among consecutive frames, is the fundamental difference between just doing detection and tracking. This can be seen in Figure. 1, where a person walks on the pedestrian's path. Due to the nature of human locomotion, the human Body changes shape from Frame to Frame.

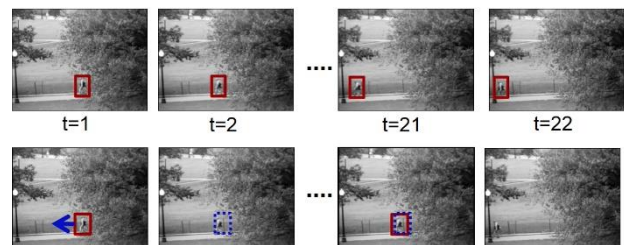


Figure 1: Detection vs tracking. In the first row, a detection occurs in every frame. In the second row a tracking occurs (dotted bounding box) in every frame whereas detection occurs every n -frame.

In this study, we propose to utilize a detector alongside a tracker to mitigate its shortcomings. Specifically, The detector acts as a localization initializer and a self-correction mechanism periodically and whenever the tracker loses the target.

The resulting method delivers good segmentation accuracy. It is also lightweight enough to run on single-core central processing unit (CPU) or even edge computing boards. The performance of this method in detecting & tracking ATLAS-vehicles was tested with a dataset we created. Experiment results indicate that the proposed method is suitable for real-time tracking and detection application with very much affordable computation requirements.

1.1 Scope on ATLAS

The project germ: "Asynchroner Transport-, Logistik- und Automatisierungsmodus auf der Schiene" (ATLAS)- [1] represents a demonstration plant (100 m long of total rail-models), on which lightweight individual self-driving vehicles drive. The concept was first launched in 2006 at the University of Applied Science in Mittweida-Germany by Prof. Christian Schultz and is currently being developed. The purpose of this project is not only to introduce a new control principle of the train which in turn is based on distributed control instead of central one, but also its core concept starts from the idea of abandoning the train's manufacture that has been used since the middle of 19th Century as basis structure for train composition along the way until reaching the fully automated level. This leads to a huge reduction in power consumption by acceleration and the breaking distances are shortened, unlike the current train operation where huge safety distances are necessary, and this is why one sees an empty railroad and crowded highway.

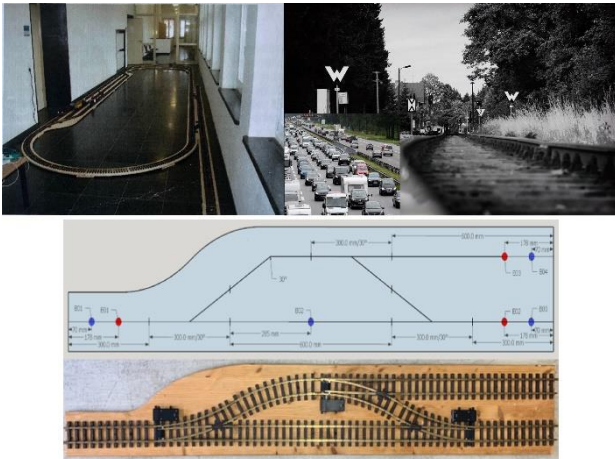


Figure 2: (Top-left) Part of ATLAS's demonstration plant. (top-right) fully occupied highway while empty railroad which expresses the Project's motivations. (Bottom) technical view of an ATLAS submodule. Blue points are position indicator (Balise) and red points represent Infrared-receivers for controlling the switches.

The project also aims to implement and invest those automation driving algorithms on the railroad, since the constraints and demands are much more convenient than the ones on the street. E.g. no permanent steering of the vehicles is required. The main points to be considered are distance control, collision avoidance, controlling the Railroad switches, and cooperative driving and route planning. Unlike the current European Train Control System (ETCS), the vehicles in ATLAS should be able to take full control of the route, accomplish the previously mentioned tasks ideally reverse the philosophy "stupid train, clever route" into "stupid route, clever train"

Therefore, the vehicles in ATLAS are equipped with sensors for sensing the environment. These are, Distance Sensor Module (DSM), Jetson Nano which represents the environmental model and is responsible for route planning as well as communication among the vehicles via vehicle to vehicle

wireless link (V2V), a mainboard (M16C) for controlling the engine, and the Infrared (IR)-system.

The Rail-model of ATLAS is provided with IR-senders called Fr. Balisen (milestone). These are position transmitter that broadcast periodically 16-bit code, which represents the absolute position of the vehicle. This means that up to 65536 Balise can be installed. Currently, there are 130 Balisen installed on the Demonstration plant. When a vehicle drives above the Balise, it receives, processes, and shares its position with other neighbor vehicles.

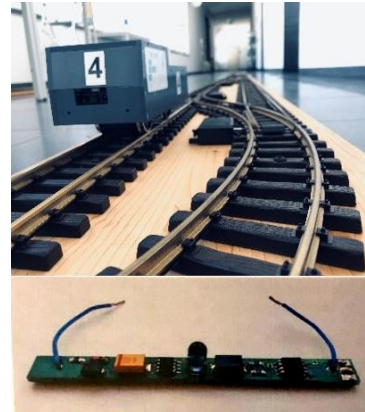


Figure 3: (top) ATLAS-Vehicle model. (bottom) Balise-module.

1.2 Viewpoint

The camera network consists of multiple static cameras (at least two) with relatively non-overlapping views. This network operates independently from ATLAS-system. All cameras are connected through a usb3-hub to the computing unit. The cameras are placed on Tripod and ladder in such a way that a bird's eye view of the scene is obtained. Thus, it is expected that both have relatively the same rotation vector with respect to the world origin. This is very convenient compared to cameras with mutual view and different rotations, since in our case the object of interest (OOI). That is the ATLAS-vehicle that cannot be present in both camera's Field of View (FOV) at the same time. Therefore the problem of associating objects belonging to the same class is improbable.

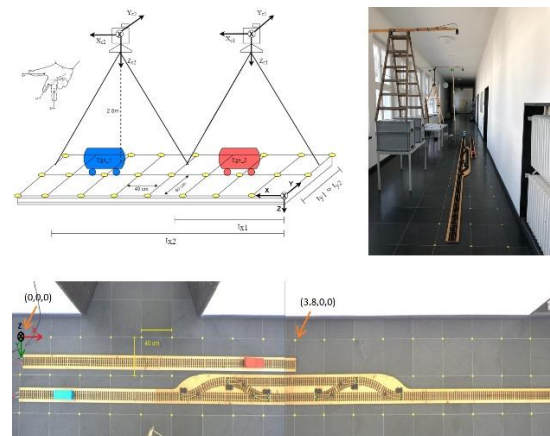


Figure 4: Side and top view of the of camera system and ATLAS respectively. Cameras are placed at 2,8m high from the scene.

It is important to refer that the world axes directions are set using the convention of the clockwise axes-system (right-hand rule) which makes it aligned with the camera coordinate system. In figure 4, we can see yellow points placed on the world plane in certain positions. These points will be used along with their image points correspondences for estimating the camera pose (Rotation and Translation of both cameras) which results, by incorporating the intrinsics as well as homography between the world and both sensor plane respectively.

2. Overall System Flowchart

The developed camera system can:

- Detect and track any ATLAS-vehicle located in the FOV of any camera.
- Identify the direction of every tracked vehicle.
- Reconstruct the planar-world coordinates of the tracked vehicle.
- Determine the status of each relevant switch with respect to every vehicle being tracked.
- Assemble the independent camera views together to form the bird's eye view mosaic and display the results on a simulation plane.
- Exploit all above information in detecting collision for our scenarios.

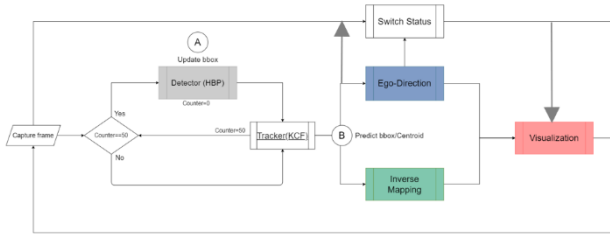


Figure 5: Block diagram of the whole process.

2.1. Proposed Method

In this study, we propose a hybrid method to track ATLAS's self-driving vehicles in real-time. For that reason, the Kernelized Correlation Filter (KCF)- [2], which is particularly fast at tracking, is used along with detection approach, Histogram Back-Projection (HBP)- [3], which in turn uses color information and the geometry of the OOI, namely the Rectangularity, Area and Ratio.

2.1.1. HBP Detector

CV tasks related to object detection expose the answers to the dichotomy: "What/where?". "what" means the ability to identify or classify objects from different object categories in each input image. Algorithms such as histogram intersection, supported vector machine (SVM) and other machine and deep learning- based classification solve this problem. On the other hand, "where" abbreviates the question, where in the target image is the object, one is looking for?

HBP answers this question considering that the image is colored. Given an input target-image which contains the OOI to be localized, the colors which appear in other objects besides the OOI are de-emphasized and they are less likely

to distract the search mechanism. On the contrary, the pixel values belonging to the OOI are enhanced. Therefore, each vehicle must have a discriminative color representing its ID.

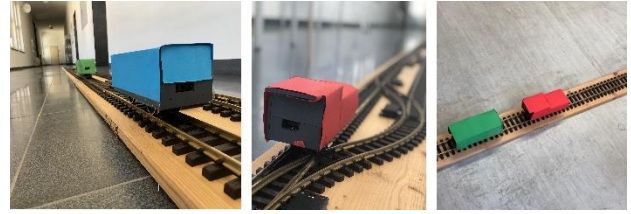


Figure 6: ATLAS-vehicles, each with a different color.

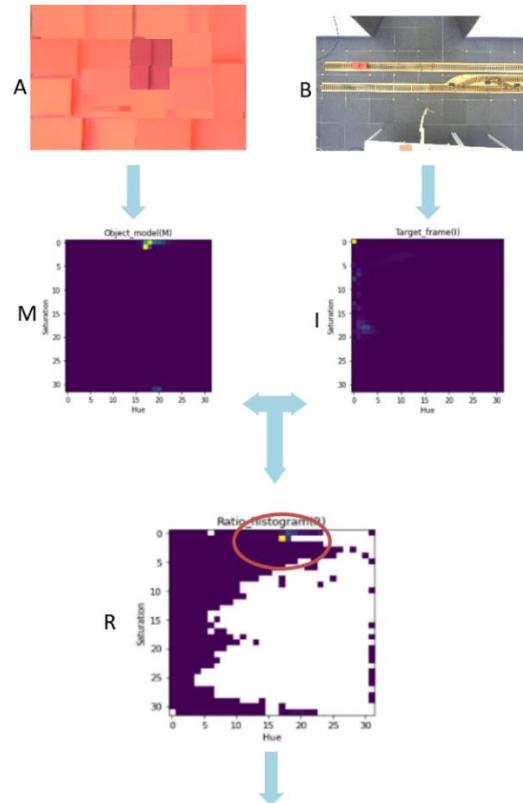
Let M be the calculated 2-dimensional color histogram of the object model (figure 7-A) which we search for, and I the histogram of the target image (figure 7-B). Consequently, a third histogram R , which is the ratio of M divided by I , is computed.

$$R = \frac{M}{I} \quad (1)$$

This histogram contains pixel-values distribution for the object model only. Then the Ratio histogram is back-projected onto the image, that is, the image values are replaced by the values of R that they index. By normalizing pixel values between 0 and 1, each resulted pixel value of the final image represents the probability belonging to the object model, given that there is an object.

$$p(x|obj) = \frac{p(x \cap obj)}{p(obj)} \quad (2)$$

All histograms were created based on HSV-color space using the Hue and Saturation channel



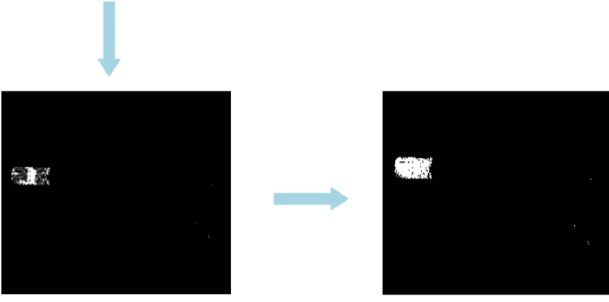


Figure 7: overall visual workflow of the HPB-detection algorithm.

The HBP is then followed by three post-processing stages. These are, refining the obtained mask-image, object representation, and further improving the detection accuracy through shape descriptors that best describes the target object. In the refining process, we worked on a binary image (values range between 0-255) which results from HBP-pipeline. First, a convolution with radius D , which has relatively the same area as the expected area subtended by the OOI, is performed to improve the implicit segmentation.

$$Obj_{mask} = Obj_mask * D \quad (3)$$

Further filtering steps are then performed such as removing bit-wise false negatives (FN) and salt noise in the mask image. These operations which are used for enhancing the detected object are called morphological transformations. E.g., closing and dilation.

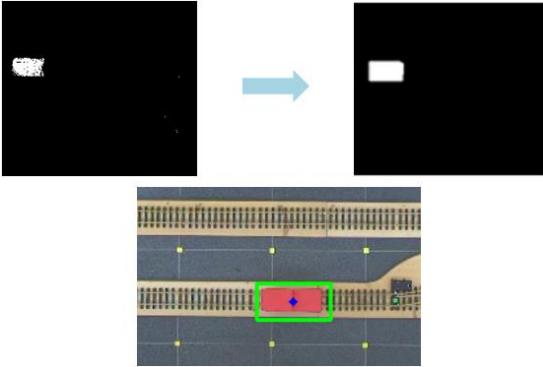


Figure 8: Object representation.

Second, the binary image which contains the enhanced OOI is passed to the feature representation process. In this step, the OOI is represented using primitive shapes. We used Centroid for displaying the OOI on the simulation image. At the same time, a rotated bounding box is implicitly calculated and used for further discriminating the OOI from other false positives (FP) objects that, because of color similarity, remain in the output binary image. The OOI's area, due to the fact that the vehicles are observed from Bird's-eye view, varies in certain range and relatively preserves its shape which in this case a rectangle. Hence, the OOI's rectangularity according to Podczek's shapes- [4] is obtained through the following equation

$$Rectangularity = \frac{a}{wh} \quad (4)$$

where,

a is the object area, and w and h are the width and height of the smallest bbox containing the object respectively.

2.1.2. KCF Tracker

As the name implies, correlation between two samples is calculated. The correlation value is in the range of X and Y and maximizes when samples are identical. This is the principle of every correlation filter and can be adapted to the tracking problem. Visual tracking methods based on correlation filters are designed to produce the highest response when the filter is applied on the object to be tracked. Unlike all of the tracking-methods such as MOSSE [5], MIL-Track [6] and STRUCK [7] which had one thing in common: a "sparse sampling strategy" that selects p random sub-windows from the Region of Interest (ROI), KCF uses "dense sampling" which takes all sub-windows in the ROI and train a ridge-regression classifier with all samples, and thus it allows a more efficient training. To reduce the computational complexity of the learned filter, and thus to speed up the method, the properties of circulant matrix-[8] are used. Additionally, the calculations are made in the frequency domain using Fast Fourier Transform (FFT) to quickly incorporate information from all sub windows, without iterating over them. KCF also utilizes the kernel trick to improve the accuracy.

let y_i represent a target and $f(z) = W^T z$ be the function that minimizes the squared error over samples X_i and their regression targets y_i . The KCF briefly solves the following ridge regression optimization problem,

$$\min_w \sum_i (f(X_i) - y_i)^2 + \lambda \|W\|^2 \quad (5)$$

Where λ is a regularization parameter that controls overfitting. The solution of the problem in the frequency domain corresponds to

$$\hat{W} = \frac{\hat{X}^* \odot \hat{Y}^*}{\hat{X}^* \odot \hat{X} + \lambda} \quad (6)$$

where \hat{W} indicates the Fourier transform of W , \hat{X}^* is the complex conjugate of \hat{X} , and \odot denotes the element-wise product. We can easily obtain W in the spatial domain by using inverse Discrete Fourier Transform (DFT). Since the problem is solved in the frequency domain, the method is extremely fast and thus makes it ideal for practical applications. Consequently, this method is one of the fastest in the literature, robust against to translation and scale variances, and provides high accuracy in general. However, it only searches for the object in a region of interest, and hence it fails when the object is occluded or has left the field of view, as it does not have a self-correction mechanism. This is where the HBP takes its place.

2.1.3 The Proposed Tracking-By-Detection Method

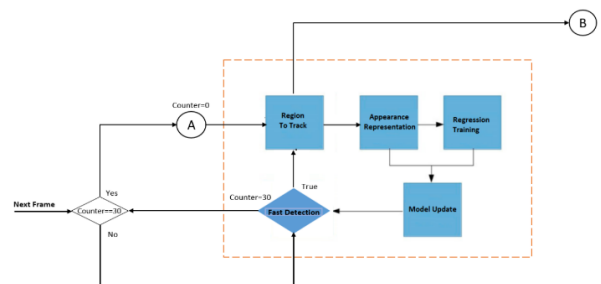


Figure 9: Block Diagram of the Hybrid-KCF Tracker.

As seen from the previous figure, a fast detection occurs implicitly inside the tracker block to generate the new region to track. The detection through HBP on the other hand occurs periodically (e.g., every 30 frame). The tracker can also call the detection earlier when it fails to fast detect the object. This uncovers the fact that the tracker should be accurate enough so that it doesn't invoke the detection very frequently. The periodical occurrence of the HPB-detection is chosen based on multiple experiments.

2.2. Determining Direction

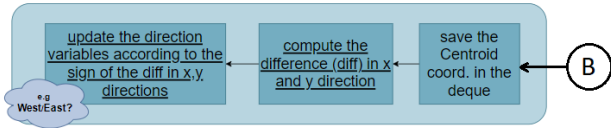


Figure 10: Block diagram of direction-determination process. B is the incoming centroid from Tracking Pipeline.

For computing the direction, the movement of the vehicle (if any) is essential. In case of no movement, the direction is not updated and the vehicle's direction is considered as "None". When the vehicle moves, its direction of movement can be obtained by computing dx and dy, namely the deltas (differences) between the x and y coordinates of the current frame and a previous one.

However, using the current and the previous frame is a bit of an unstable solution. Unless the object is moving very quickly, the deltas between the (x, y)-centroid coordinates will be very small. Thus, when computing These Deltas directly to report direction, then the results are extremely noisy, implying that even small changes in trajectory would be considered as a direction change. In fact, these changes are so small that they would be near invisible to the human eye (or at the very least, trivial) and reporting as well as tracking such small movements is not of interest.

Instead, reporting the direction of larger object movements is much more likely of interest — Hence, the deltas between the OOI-centroid coordinates of the current frame and a frame preceding the current with relatively large time-period is computed. Performing this operation helps reduce noise and false reports of direction change.

2.3. Location Reconstruction on Planar World

Any point in the 3D-space is mapped onto a sensor plane through perspective Transformation process (PTP) which can be expressed as 3x4 projection matrix P.

$$P = KR[I | -t],$$

$$R = \begin{pmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{pmatrix}, K = \begin{pmatrix} f_x & 0 & x_u \\ 0 & f_y & y_v \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} t_x \\ t_y \end{pmatrix} \quad (7)$$

where, K is the camera matrix containing the Intrinsics, R and t are the rotation and translation matrices of the camera with respect to the world origin and both are the Extrinsic.

If the points in the 3D-space lie on a plane, then they and their corresponding points on the image plane are related

by a planar-homography, and this is the key idea behind obtaining the planar world coordinates of the vehicle's centroid.

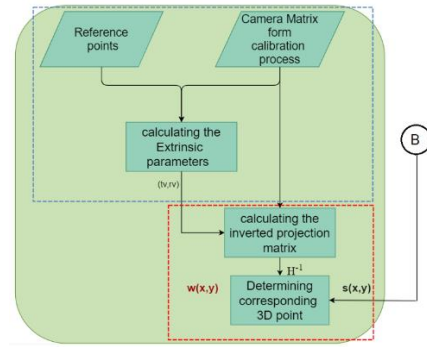


Figure 11: Block diagram of the location reconstruction's process. (tv, rv) are translation & rotation vector of the camera respectively. H⁻¹ denotes the inverse homography. S is the 2D-vehicle's centroid and W is the corresponding world point with z = 0 (planar world).

The pipeline illustrated in the above figure consists of two main processes:

- Camera Calibration
- Image Plane to World Plane transformation (Planar-Homography)

2.3.1 Camera calibration

This process involves finding the Intrinsics, Extrinsic and distortion coefficient of the cameras. For estimating the Intrinsics of both cameras, we have performed Zhang's Method on Matlab, which uses known calibration points located on a checkerboard and find correspondences between these points when they are in different positions. Regarding the extrinsic, we have created a grid of world points with a constant known distance between them. For the sake of convenience, we have exploited the floor's -tile grid which has 40x40 centimeter square tile. Hence, the intersections of the tiles are the locations of the world points. To find the corresponding 2D-image points effortlessly, we have applied thresholding techniques on the image containing the grid's tiles so that only world points remain visible and then used OpenCV functions to get the 2D-pixel coordinates of the control points. Missing points can be found by manually locating the pixel locations (see figure 12).

| | | Camera matrix in H.c | | |
|----------|--|---|---|--|
| Camera 1 | | $\begin{pmatrix} 4,3807e+03 & 0 & 1,0059e+03 \\ 0 & 4,3782e+03 & 6,9702e+02 \\ 0 & 0 & 1 \end{pmatrix}$ | | |
| | | Rotation vector | Translation vector | |
| | | $\begin{pmatrix} -0,02823308 \\ 0,08623225 \\ 0,01563199 \end{pmatrix}$ | $\begin{pmatrix} -5,47884374 \\ -3,08581371 \\ 24,15112048 \end{pmatrix}$ | |
| Camera 1 | | $\begin{pmatrix} 4,4636+03 & 0 & 9,5639e+02 \\ 0 & 4,4642e+03 & 6,6447+02 \\ 0 & 0 & 1 \end{pmatrix}$ | | |
| | | Rotation vector | Translation vector | |
| | | $\begin{pmatrix} -0,04811656 \\ -0,11802668 \\ -0,11802668 \end{pmatrix}$ | $\begin{pmatrix} -8,95359027 \\ -3,01504595 \\ 24,83908461 \end{pmatrix}$ | |

Table1: Results of calibration parameters for both cameras respectively.

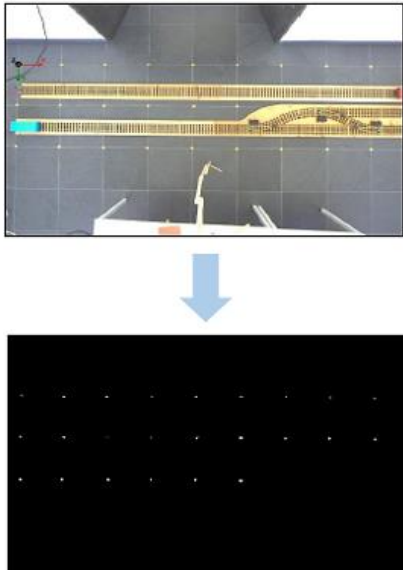


Figure 12: Finding correspondences (white spots on the upper image) of world points (yellow spots on the latter image).

2.3.2 Image Plane to World Plane transformation

Since the world coordinate system can be set anywhere, it can be conveniently positioned on the plane, such that it has zero Z-coordinate ($Z = 0$). This choice reduces the projection matrix, to

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} & h_{13} & h_{14} \\ h_{21} & h_{22} & h_{23} & h_{24} \\ h_{31} & h_{32} & h_{33} & h_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} h_{11} & h_{12} & h_{14} \\ h_{21} & h_{22} & h_{24} \\ h_{31} & h_{32} & h_{34} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} = H \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}, \quad (8)$$

where, H is 3x3 planar-homography matrix that represents the projective relationship between the world plane and the corresponding image-plane points. The equality above is up to a scale factor s and as a result the planar-homography matrix has 8 degrees of freedom.

Since this matrix H can be inverted, it follows that if H is completely known, the world coordinates of any image point can therefore be recovered through the following equation:

$$\begin{bmatrix} X/s \\ Y/s \\ 1/s \end{bmatrix} = H^{-1} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \Leftrightarrow W = H^{-1}C^T \quad (9)$$

where,

W is the reconstructed world coordinate of the detected centroid C. The above equation is expressed in homogeneous coordinates.

2.4. Detection of the Switch Status

Determining switch status means, how a switch is connected with respect to each vehicle located on its trajectory and being observed by the system, e.g., Left = L/Right = R connected. To this end, we have exploited two variables:

- the vehicle's relative direction, and

- a Switch-Flag (SF).

The SF is an indicator that returns 1 whenever a switch alters its state in such way so that the white spot placed on the notch is visible for the system (cf. see fig. 13-top). However, detecting such a small area in an image (5-6 pixels) from a relatively far working distance is not straightforward when it comes to the camera specifications, which consist of resolution, sharpness, focal length, sensor dimensions, and angle of view. Meaning to say, they must be examined so that the system can still detect the ROI from the defined working distance and under different lightening conditions. In the research project [9], we have examined all hardware components required for the system and therefore they are out of the scope of this work. The procedure for detecting the white area on the notch is shown in below figure:

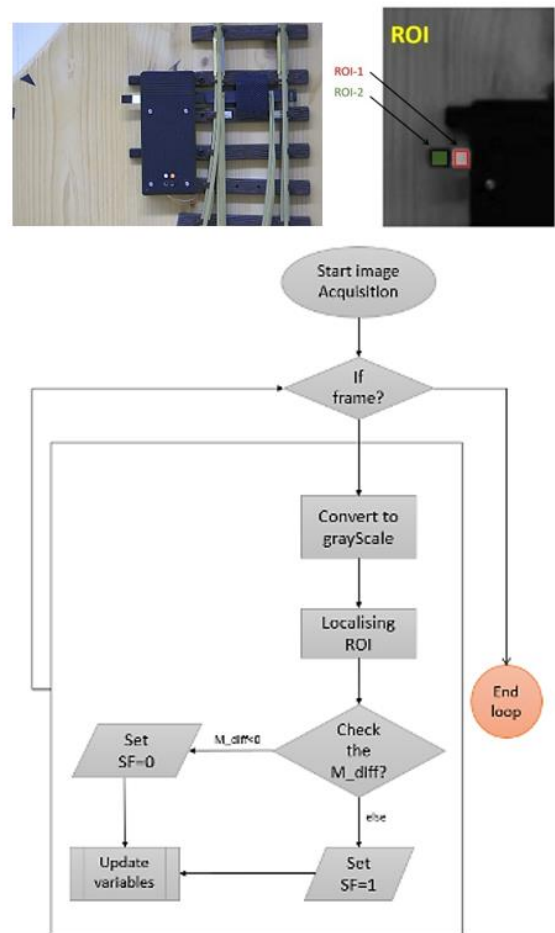


Figure 13: (Top)-ATLAS-Switch Model showing extraction of ROI. (Bottom)-Block Diagram of Detecting white spot on rail switch.

For the generality of the previous procedure, it must be assured that all white spots are placed on the same side for all rail-modules when they are visible! Thus, we could create a Look-Up-Table (LUT) that consists of all possible (considered) directions and SF's state along with the switch status accordingly.

For instance, the blue vehicle shown in figure 14 is heading "North" and the SF of the switch located a head is 1, then according to the LUT, this switch is Left-connected with respect to the blue vehicle.

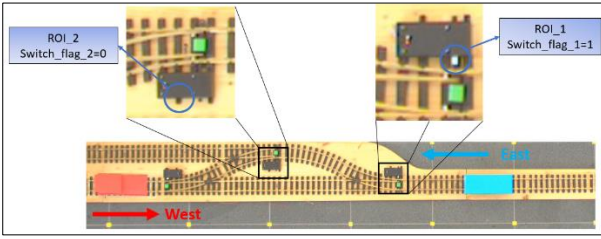


Figure 14: an illustration showing how a combination of SF and vehicle's direction gives an output that indicates the current Status of a certain switch relative to a relevant vehicle.

| Vehicle Direction | Switch flag | Status |
|-------------------|-------------|--------|
| North | 0 | R |
| North | 1 | L |
| South | 0 | L |
| South | 1 | R |
| West | 1 | L |
| West | 0 | R |
| East | 1 | R |
| East | 0 | L |

Table 2: Look Up table.

2.5. Visualisation

This step capsules the work into one uniform view. In other words, all camera's FOVs as well as the extracted scene-related information are transformed onto one view (simulation image). For this purpose, we have used the same concept explained in section 2.3.2.

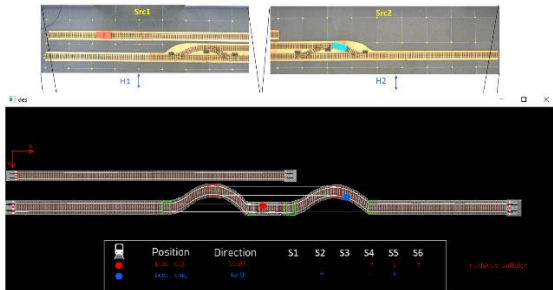


Figure 15: Transformation of Src1 & Src2 onto simulation-plane and thus achieving a central viewing monitor. In the Dst-image, information about the tracked vehicles (colored dots) such as, their position with respect to world origin and directions according to compass direction are shown. S1 till S6 are switch status with respect to every vehicle.

3. System Evaluation

The evaluation process can be split into three subtitles:

- Evaluation of the adopted Detection-algorithm.
- Evaluation of the adopted Tracking-algorithm.
- Evaluation of the reconstructed world coordinates of the vehicles being tracked.

Speaking of dataset, a 20 video sequences with 1280x720 resolution were collected. These videos have attributes that concerns different lightening conditions such as sunny & cloudy weather, with and without artificial room lights,

shadows, and illumination variation (IV). The dataset is then split into train and test data. We call it training data because the adopted HBP learns in the sense that it memorizes the training data.

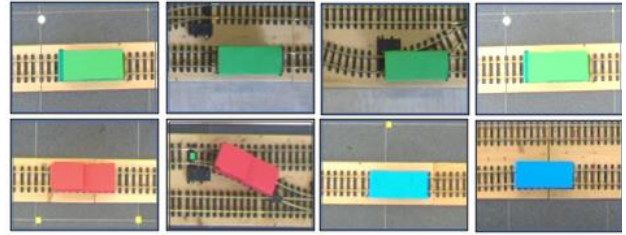


Figure 16: Snapshots from the collected dataset showing different lightening conditions.

A. We have evaluated the HBP against three other approaches, each of which has different methodology and level of complexity, however, all use color component as cue feature representation. These are: Basic Color Range Detector using HSV-color space (CRD), Color Component Ratio (CCR), and Gaussian Mixture Model (GMM). The four approaches are evaluated based on precision, which in this case defined as "how accurate the algorithm can discriminate the OOI", and Speed "number of frames per seconds (FPS) the algorithm can process". Due to unavailability of the optimal ground-truth data, the precision-evaluation is performed through the following method:

Based on the fact that an axis-aligned bounding box or a centroid of the OOI is obtained based on a binary-mask image resulted from an implicit segmentation and that the scene is being observed from Bird's-eye view camera, the OOI would always have a rectangular shape considering a perfect segmentation (under labor-conditions). Hence, a candidate blob that far deviates from rectangle-shape is considered a poor score. In other words, the more the segmented blob deviates from being a rectangle the less precise the detection is.

For making the evaluation-process rational, we have additionally defined a range-threshold for the area of the detected blob. So, for instance, if a detected blob's area falls outside the threshold, it will not be accepted, even though it has a high rectangularity score.

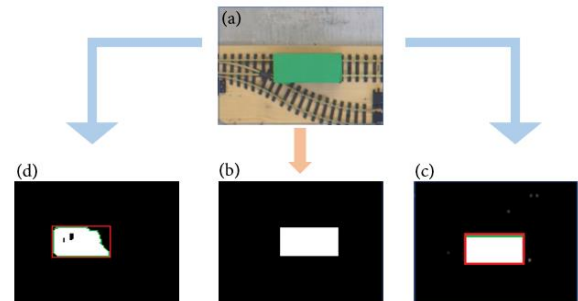


Figure 17: (a)-input image, (b)- the optimal output mask in which the target object has 100% rectangularity and the blob area is in-between reasonable threshold. (c)-output mask showing slight deformation from a rectangular shape. (d)- output mask showing a large deformation.

| Algorithm | Feature | Mean Rect | Mean FPS |
|---------------|--|-------------|-----------|
| CRD | color component (raw pixels) | 0,42 | 30 |
| CCR | | 0,75 | 7 |
| GMM | | 0.94 | 0.7 |
| HPB (adopted) | color component & shape descriptors | 0,79 | 34 |

Table 3: The table shows Summary of experiment results on 15 Videos. The reported quantities are averaged over all videos. Reported speeds include feature computation and post-processing step.

Looking at the previous table, the HPB outperforms the other candidates regarding speed, despite processing additional features like rectangularity, area, and ratio. On the other hand, it comes in second place after GMM regarding the average rectangularity, which is reasonable, because GMM does not only soft-assign each pixel in the image to the defined clusters with a probability value (weight), but its parameters are also efficiently pre-trained on multiple color distributions of the OOI.

B. For tracking-process, the implemented KCF-tracker is evaluated among other commonly trackers (state-of-the-art at its time). The best way to evaluate trackers is still a debatable subject. However, Babenko et al. [10] argue for the use of precision plots. These plots show, for a range of distance thresholds, the percentage of frames that the tracker is within that distance of the ground truth, and thus are easy to interpret. For obtaining direct results, the outcome from [11] is adopted.

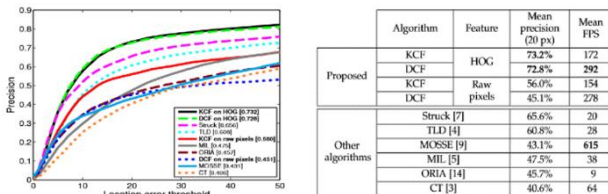


Figure 18: (left)- Precision plot for all 50 sequences. (right)-Results summary over 50 videos dataset [11].

By looking at [11], it is shown that Struck operates on many kinds of features and a growing pool of support vectors. Track-Learn-Detect (TLD) is specifically geared towards re-detection, using a set of structural rules with many parameters.

Despite this unevenness, KCF can reach competitive performance when operating on raw pixels as seen in the table of figure 18. Replacing raw pixel information with HOG features allows the KCF) to surpass even TLD and Struck, by a relatively large margin.

As we are looking for a tracker that gives a compromising result between precision and speed, it is shown that KCF is most suitable for the work's use case. In order to realize the difference in performance, namely speed between using pure Detection and the proposed Tracking-By-Detection approach and since the tracking speed is a crucial factor in our use case, the speed of both approaches is therefore plotted on multiple videos from the collected dataset.

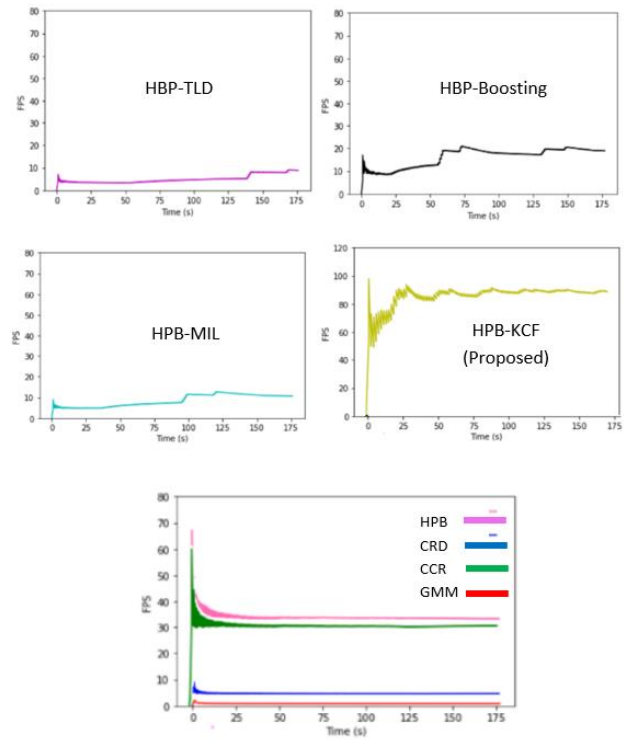


Figure 19: Comparing Hyper-Tracking against pure Detection. The evaluation process is made based on data taken from one camera.

The above figure shows that, integrating KCF in the overall process not only outperforms the pure detection, but also other available trackers by large margin. In fact, an average of 400 fps can be reached by reducing the size of the tracked region, however, it reduces the precision-performance.

C. The last part of the evaluation confines oneself in the following question: How accurate are the world coordinates of the ATLAS-vehicles obtained from the Detect-Track-Reconstruct pipeline with respect to the ground truth-coordinates?

The ground-truth/actual positions were taken using a caliper device. Considering the distance between every two yellow control points in x and y direction as reference and equals to 40cm, direct results can be obtained. On the other hand, an error-deviation in millimeter-range caused from the previous convention is expected.

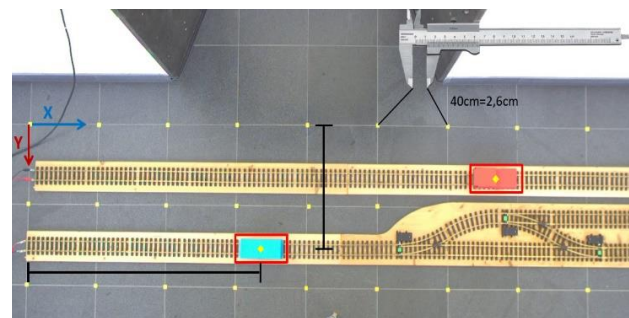


Figure 20: An illustration showing how ground-truth (actual locations) are estimated.

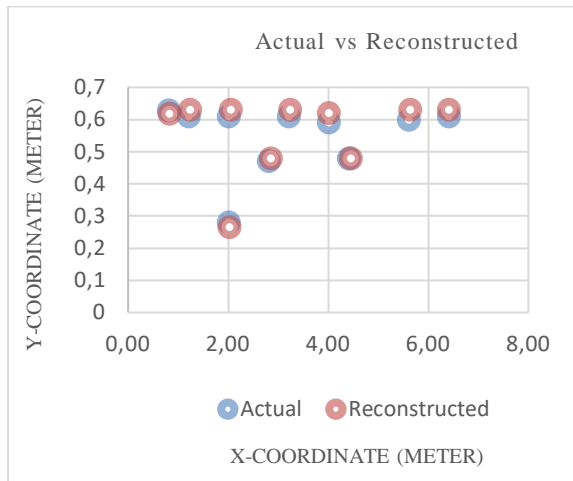


Figure 21: The scatter chart plotting ten different reconstructed coordinates of ATLAS vehicles compared with their observed (actual) ones respectively.

Based on the previous convention, and on the experimental results depicted in figure 20, it is observed that the Euclidean offset between the actual and reconstructed points for both x and y coordinates are very low which implies a very good performance of the process. Mean Squared Error (MSE) was taken as an evaluation metric. The results were (0,00092, 0,000379) meter for x, y world-coordinates respectively.

4. Collision Detection

Detecting a collision is, due to lots of variables that plays an important role by decision making, can range from very easy to very complex. This is essentially based on how the trajectories change. The more the trajectories vary, the more parameters must be taken into consideration.

Collision scenarios can therefore be simplified in two main types:

- I. vehicles drive in parallel in the same direction reaching a connecting node (switch)
- II. vehicles drive opposite of each other.

In the following, the second case is considered as it is relatively more complex:

1. A collision is potential if two vehicles have opposite directions and at the same time either of their coordinates are equal or fall within a certain small range.
2. A potential collision becomes critical if, in addition to 1st statement, the Euclidean distance between the two vehicles becomes smaller than a defined threshold -distance value. The threshold value is trajectory-specific and chosen rationally based on multiple experiments.
3. A critical collision becomes Incident or Inevitable, if in addition to 1st & 2nd statements, the status of the most-near switches in between two vehicles are different with respect to the vehicles and within certain timestamp do not match.

The timestamp defines the time in which a switch's status is expected to be changed from the current passing vehicle. Since the vehicles control the switches through Infrared-sensor, the timestamp can be translated into a distance which equals to 15cm. This is the maximal distance where a signal can be received via Infrared-link. (For more details, see [1]).

| If | II | 1 | 2 | 3 | ... |
|--------------|-----------|---|---|---|-----|
| Collision is | potential | | | | |
| | Critical | | | | |
| | Incident | | | | |

Table 4: Collision table for collision type I.

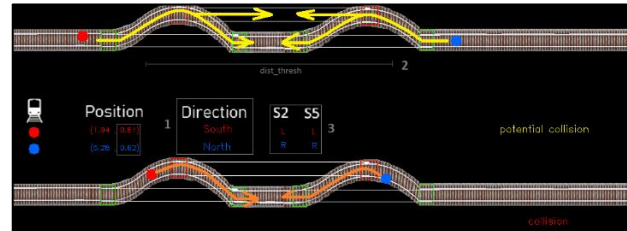


Figure 22: collision scenario. On top of the image, we can see that the system sets the decision as potential because the first condition is True. The system changes the decision to incidence because all three conditions being considered for this trajectory are True.

5. Conclusion & Future Work

In this work, a surveillance system dedicated for observing vehicles driving on the ATLAS-demonstration plant was implemented & tested. By "Observing" it is meant that the system can detect and track trajectory-driven vehicles in real time with the focus on achieving the highest possible detection speed without sacrificing much detection quality. Therefore, a Hyper-Tracking approach was used which combines two different methodologies, HBP for detecting the OOI and KCF for tracking it.

Then, the centroid of the tracked vehicle obtained from the Hyper-Tracker's output was retransformed onto the planar world using planar perspective transformation and used for determining the vehicles direction according to cardinal system. At the same time, the detected-tracked centroids were also used in corporation of other metrics for identifying the status of rail-switches relative to the vehicles.

Then, the above-mentioned information was used on a couple of a collision scenarios. In the first scenario, two vehicles arrive at the critical area at the same time and in the other one, the two vehicles arrive at different time with different velocities.

It is important to mention that the goal is to show that for limited scenarios, the system can recognize collision based on the information extracted throughout the whole process of the system. Nevertheless, the system as in its preliminary stage, cannot be generalized for detecting every possible collision on ATLAS-demonstration plant. This means, by extending the system such as adding a third camera, a new vehicle or even extending the rail-model, the systems capabilities against detection collision must also be considered accordingly.

To make the system more generalized for detecting collision and robust against any extend needed for future implementations, following aspects should be considered on two levels:

1. Computational and Performance Improvement.

- I. In the evaluation chapter, we saw that with integrating KCF, an average speed of 94 FPS was reached (on an Intel® Core™ i5-8250U CPU @ 1.6GHz). By incorporating the second camera, the speed drops down to ~50 FPS. In the current implementation, KCF uses raw pixels for updating its predicted bounding box. However, if we use additional features e.g., HOG as tracking features, we can almost double the current speed (refer to Table in Figure 18). This is important and efficient as it is expected that the speed would drop when more cameras are incorporated. However, this is not the case with, for example HOG features.
- II. Exploiting the technical aesthetics of GPU. The camera system should in the future cover up to 30m long of rail modules which are horizontally aggregated. This requires up to 6 cameras with 2.3 mega pixel. As a result, the computation speed would drastically drop. However, the system does the same exact computations for all data arriving from different cameras using single CPU-core. Instead, these computations can be executed in parallel. In other words, they can be distributed on multiple cores and thus significantly improve the performance. Since GPUs are, by large margin, a lot faster than CPUs when doing parallel programming, they can be deployed instead, and this would further improve the performance

2. Increasing System Complexity

The system should be extended to consider more metrics that contribute to more efficiently decision-making by collision detection. As an example, can be estimating the speed of each vehicle which helps knowing when a vehicle reaches a collision node.

In this work, only two collision scenarios were considered. This is not enough to generalize the system for all possible collision scenarios. In other words, the more the scenarios differ, the more the system must be extended, and probably, the more parameters must be considered and tested.

References

- [1] Christian Schultz, "ATLAS-Demonstrationsanlage für autonomes Fahren auf der Schiene ", 2018-University of Applied Science, Mittweida-Germany.
- [2] Jo˜ao F. Henriques et.al, "High-Speed Tracking with Kernelized Correlation Filters"
- [3] "Indexing via color histograms", Swain, Michael J. , Third international conference on computer vision,1990.
- [4] polesky shape descriptors.
- [5] D. S. Bolme, J. R. Beveridge, B. A. Draper, and Y. M. Lui, "Visual object tracking using adaptive correlation filters," in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on. IEEE, 2010, pp. 2544–2550.
- [6] B. Babenko, M.-H. Yang, and S. Belongie. Robust object tracking with online multiple instance learning. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(8):1619–1632, 2011.
- [7] S. Hare, S. Golodetz, A. Saffari, V. Vineet, M.-M. Cheng, S. L. Hicks, and P. H. Torr. Struck: Structured output tracking with kernels. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(10):2096–2109, 2016
- [8] Jo˜ao F. Henriques et.al, "Exploiting the Circulant Structure of Tracking-by-detection with Kernels"- Institute of Systems and Robotics, University of Coimbra.
- [9] Khaled. Jbaili, "Camera evaluation for system monitoring of ATLAS", Research Project University of Applied Science-Mittweida-Germany 2020.
- [10] A.R. Zamir, A. Dehghan, and M. Shah. GMCP-Tracker: global multi-object tracking using generalized minimum clique graphs. In ECCV, 2012.
- [11] Jo˜ao F. Henriques et.al, "High-Speed Tracking with Kernelized Correlation Filters"

***Suffizienzstrategien für
Immobilien – Darf es etwas
weniger sein?***

Wie muss man bauen, um suffizientes Wohnen zu ermöglichen?

Margarete Over (ifeu), Patrick Zimmermann (BTU), Lars-Arvid Brischke (ifeu)
ifeu – Institut für Energie- und Umweltforschung Heidelberg gGmbH, Wilckensstr. 3, 69120 Heidelberg;
BTU Cottbus-Senftenberg, Konrad-Wachsmann-Allee 8, 03046 Cottbus

Suffizientes Wohnen erfordert eine Architektur-, Gebäude- und Freiraumplanung, bei der die Befriedigung der individuellen Wohnbedürfnisse unter Einhaltung lokaler und globaler ökologischer Belastungsgrenzen im Mittelpunkt steht. Schlüsselemente bilden hierbei die absolute Reduktion der Wohnfläche pro Kopf, die bedürfnisorientierte Anpassungsfähigkeit der Wohn- und Nutzflächen sowie weiterer energie- und ressourcenverbrauchsrelevanter Eigenschaften des Gebäudes, die Bereitstellung multifunktionaler Flächen und Räume für suffiziente Alltagspraktiken und der Einbezug von Freiflächen mit Nutz- und Aufenthaltsqualität. Darüber hinaus sind Standort, Eigentumsverhältnisse und Organisation der Objekte wichtige Randbedingungen für die Umsetzung von Suffizienz. Zugleich bedarf es der Vermittlung kultureller Praktiken für einen sachgerechten, wertschätzenden und maßvollen Umgang mit Ressourcen, Infrastrukturen und der Umwelt. Diese Aspekte werden anhand eines Bauprojektes vorgestellt, welches das ifeu wissenschaftlich begleitet.

1. Einleitung

Das Konsumfeld Bauen und Wohnen hat einen entscheidenden Einfluss auf den ökologischen Fußabdruck und die Lebensqualität der Bürger:innen. So ist der Gebäudesektor in Deutschland für etwa ein Viertel der Treibhausgasemissionen [1], das Bauwesen für etwa ein Fünftel des Rohstoffkonsums der Endnachfrage [2] sowie für 32 Hektar pro Tag Netto-Flächenversiegelung durch die Zunahme von bebauten Siedlungsflächen verantwortlich [3]. Darüber hinaus hat das Bauen und Wohnen indirekt ökologische und soziale Auswirkungen auf die anderen Konsumfelder, insbesondere auf die Mobilität. Die energetische Sanierung von Gebäuden (Effizienz-Strategie) und der Einsatz von kreislauffähigen Baustoffen und erneuerbarer Energien (Konsistenz-Strategie) sind zentrale Bausteine zur Erreichung einiger vor allem ökologisch orientierter Nachhaltigkeitsziele. Dennoch gelangen diese technologischen Strategien angesichts der begrenzten lokalen Verfügbarkeit von Flächen und Ressourcen sowie Rebound-Effekten an ihre Grenzen: Obwohl es in Deutschland mit durchschnittlich 47 m² Wohnfläche pro Einwohner:in genug Wohnraum geben sollte [4], herrscht aus sozialer Sicht in zahlreichen Kommunen und Regionen Wohnraumangel, wodurch Interessenskonflikte und Druck rund um verbleibende, bisher unbebaute Flächen wachsen.

Der vorliegende Beitrag befasst sich mit Suffizienz als planerischem Ansatz, der darauf abzielt, unter Berücksichtigung lokaler und globaler ökologischer Grenzen die Bedürfnisse der Nutzer:innen nach attraktivem und bezahlbarem Wohnen zu befriedigen. Für das Bauen bedeutet das vor allem einen wertschätzenden, bedürfnisorientierten und umweltschonenden Umgang mit dem Vorhandenen, das heißt Flächen, stofflichen und natürlichen Ressourcen sowie bestehender Bausubstanz.

2. Suffizienz als Nachhaltigkeitsstrategie und planerischer Ansatz beim Bauen und Wohnen

Suffizienz als Nachhaltigkeitsstrategie besteht in Verhaltensweisen und Konsummustern, „die helfen, innerhalb der ökologischen Tragfähigkeit der Erde zu bleiben, wobei sich Nutzenaspekte des Konsums ändern“ [5]. Sie zielt darauf ab, persönliche Bedürfnisse, Bedarfe und Wünsche mit persönlichen, gesellschaftlichen und ökologischen Grenzen in Einklang zu bringen [6]. Es geht also bei Suffizienz nicht nur darum, Ressourcenverbrauch und Umweltauswirkungen auf ein lokal und global langfristig tragfähiges Maß zu reduzieren, sondern auch darum, allen Menschen die Möglichkeit zu geben, gleichermaßen an der Nutzung globaler Ressourcen teilzuhaben und ihre elementaren Bedürfnisse zu befriedigen. Mit diesen Zielen und Inhalten hat Suffizienz unter den Nachhaltigkeitsstrategien ein Alleinstellungsmerkmal und ergänzt Effizienz und Konsistenz komplementär.

Um suffizientes Wohnen zu ermöglichen, müssen die entscheidenden Weichen bereits vor Baubeginn gestellt werden. Der vorliegende Beitrag legt dar, wie Suffizienz als konsequent umgesetzte Planungsstrategie das zukünftige Verhalten der Nutzer:innen in den Mittelpunkt stellt und veranschaulicht diesen Ansatz exemplarisch an einem Heidelberger Bauprojekt, welches das Institut für Energie- und Umweltforschung Heidelberg (ifeu) wissenschaftlich begleitet [7]. Dabei wird die von Zimmermann (2018) entwickelte Suffizienz-Bewertungsmethodik für Wohngebäude genutzt [8].

3. Das Bauprojekt „Collegium Academicum“ als Reallabor

In Heidelberg arbeitet seit 2013 eine ehrenamtliche Projektgruppe junger Menschen am Aufbau des selbstverwalteten Wohnheims „Collegium Academicum“ (CA) für über 240 Studierende, Auszubildende und andere junge

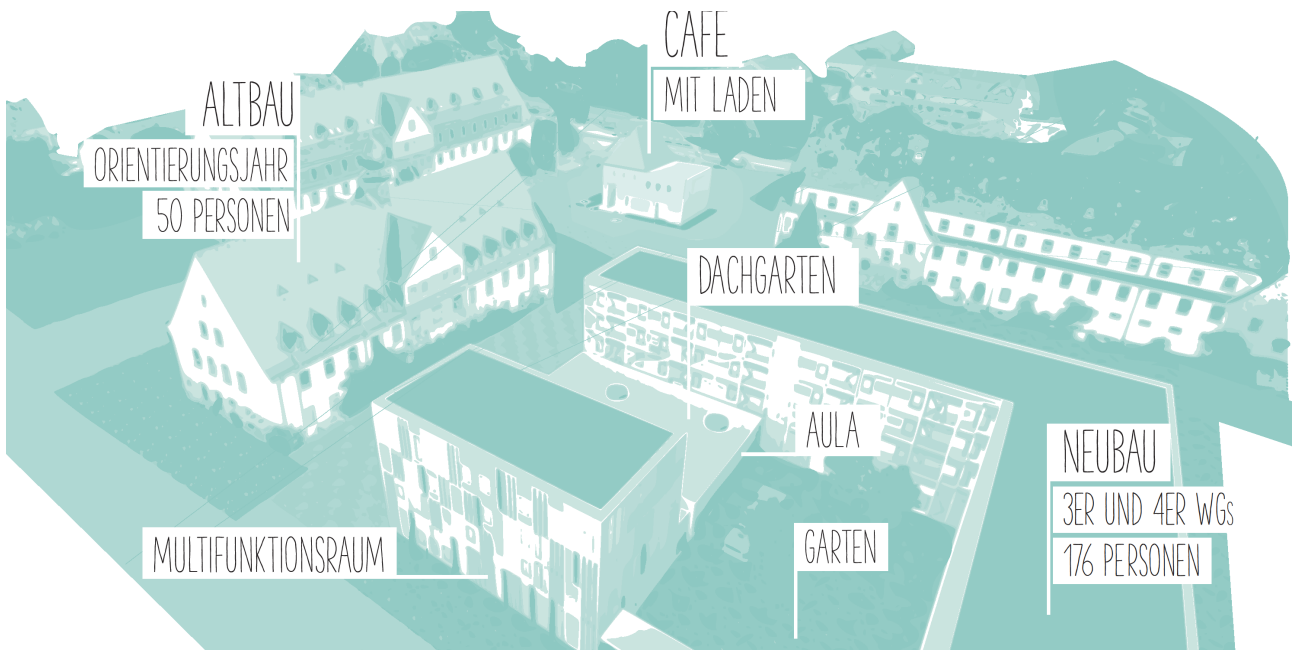


Abb. 1: Das CA im Überblick. Darstellung: Collegium Academicum

Menschen [9]. Das Wohnheim besteht aus einem Neubau mit Wohnraum für 176 Personen und zwei Bestandsgebäuden mit zusätzlich etwa 70 Wohnplätzen. Das Wohnheim soll den suffizienzorientierten Nukleus für ein neues Stadtquartier auf einer ehemals militärisch genutzten Konversionsfläche bilden. Der Neubau wird Ende 2021 bezugsfertig sein.

Ausgangspunkt der CA-Initiative war das Bestreben, einen Ort zu schaffen, der jungen Menschen Freiräume zum gemeinschaftlichen und bezahlbaren Lernen und Leben bietet und ein sozialer und kultureller Treffpunkt im Quartier wird. Daher stehen vielfältig nutzbare Gemeinschaftsflächen im Zentrum der Planung. Hierzu gehören eine Werkstatt, eine Versammlungsstätte für über 600 Personen, multifunktionale Räume mit Gemeinschaftsküche sowie ein Dachgarten. In dem kleinen Bestandsgebäude soll ein Quartierscafé entstehen (Abb.1).

Im Zuge eines ganzheitlichen Nachhaltigkeitsansatzes wird der viergeschossige Holz-Neubau nach dem Passivhausstandard (Effizienz) mit Photovoltaikanlage auf dem Dach errichtet. In der Baukonstruktion setzten die Planer:innen überwiegend auf Baustoffe aus nachwachsenden Rohstoffen (Konsistenz). Die Holzskelettbauweise kommt durch den Einsatz von Holz-Holz-Verbindungen weitestgehend ohne metallische Verbindungsmittel in der Primärkonstruktion aus. Die Planungen fanden unter der Federführung von DGJ Architektur aus Frankfurt statt. Der Neubau ist unter dem Motto „Wissen | schafft | Stadt“ Projekt der Internationalen Bauausstellung Heidelberg und eines der Modellprojekte des BBSR-Förderprogramms „Variowohnungen“ aus der Forschungsinitiative Zukunft Bau und des EFRE-Förderprogramms „Holz-Innovativ“ des Ministeriums für Ländlichen Raum Baden-Württemberg [10,11].

Schließlich hat die konsequente Integration von Suffizienz in der Planung zur hohen Vereinbarkeit von ökologischen und sozialen Zielen geführt. Hierbei ist das CA Reallabor des Forschungsprojektes „SuPraStadt – Lebensqualität, Teilhabe und Ressourcenschonung durch soziale Diffusion von Suffizienzpraktiken in Stadtquartiere“, koordiniert vom Institut für Energie- und Umweltforschung Heidelberg (ifeu) und gefördert im Rahmen des BMBF-Förderprogramms „Zukunftsstadt“ [12].

4. Suffizienz-Ziele und Kriterien für Wohngebäude am Beispiel des CA

Anhand der folgenden aus Zimmermann (2018) entnommenen Bewertungsmatrix sollen die wesentlichen Suffizienz-Ziele und -Stellschrauben erläutert und am Beispiel des CA-Neubaus veranschaulicht werden [8]. Bei der Entwicklung dieses Kriterien-Sets standen eine möglichst objektive Bewertbarkeit bzw. Quantifizierbarkeit sowie die Handhabbarkeit im Planungsprozess im Fokus. In Summe gliedert sich die Systematik hinsichtlich der Teilaspekte der Gebäudeplanung und ist aufgeteilt in sieben Themenfelder, 23 Kriterien und 58 Indikatoren, wobei an vielen Stellen auf bestehende Bewertungsmethoden aus den Zertifizierungssystemen „Deutsches Gütesiegel Nachhaltiges Bauen“ (DGNB) und

„Qualitätssiegel Nachhaltiger Wohnungsbau“ (NaWoh) zurückgegriffen wird. Im Folgenden werden beispielhaft einige zentrale Kriterien und Indikatoren aufgegriffen und die Umsetzung im CA erläutert.

4.1 Projektentwicklung und Planungsprozess

Bereits in der sogenannten „Leistungsphase 0“ werden wichtige Weichen für Nachhaltigkeit und Suffizienz von Gebäuden gestellt. Hier entscheiden sich in hohem Maße die späteren ökologischen Auswirkungen und es wird das Fundament für die Nutzer:innen-Zufriedenheit

sowie die Wohnqualität gelegt. Deshalb ist es essentiell, durch umfangreiche **Bedarfsplanung** und **Partizipation** frühzeitig darauf Einfluss zu nehmen.

Beim CA in Heidelberg waren die zukünftigen Nutzer:innen von Anfang an in die Planung eingebunden, da sie als Bauherr:innen das Projekt initiiert und die ambitionierten sozialen und ökologischen Ziele definiert haben. Dazu gehörten neben möglichst kostengünstigen Mieten und hohen ökologischen Standards insbesondere umfangreiche Gemeinschaftsflächen, mit denen Raum für Suffizienzpraktiken geschaffen wird, und die Möglichkeit, das Gebäude selbst mitzugestalten und weiterentwickeln zu können, damit die zukünftigen Generationen von Bewohner:innen es an ihre Bedürfnisse anpassen können.

Dass das Vorhaben trotz des hohen Innovations- und Modellcharakters Mieten bietet, die deutlich unter dem Durchschnitt des studentischen Wohnungsmarkts in Heidelberg liegen, liegt auch an der Non-Profit-Orientierung auf Basis des Modells des Mietshäuser Syndikats [13]. Das Mietshäuser Syndikat ist ein Verbund von gemeinschaftlich organisierten Wohnprojekten, deren Ziel die Schaffung von langfristig kostengünstigen Mieten ist. Auch aufgrund der umfangreichen ehrenamtlichen Arbeit in der Projektkoordination und -umsetzung wurde der Anteil an den Gesamtprojektkosten, der direkt in das Bauwerk investiert wurde, erhöht.

Die **Standortwahl** spielt ebenfalls eine wichtige Rolle, wobei es keinen „suffizienten Standort“ per se gibt, da die Frage nach dem rechten Maß an Urbanität oder Ländlichkeit individuell beantwortet werden muss. Dennoch haben die Verkehrsanbindung und die im Alltag zurückzulegenden Distanzen zentralen Einfluss auf die Lebensführung und den ökologischen Fußabdruck der Bewohner:innen. Deshalb wird mit der Standortwahl bereits die Grundlage für ein späteres suffizientes Mobilitätsverhalten geschaffen.

Das CA entsteht auf dem ehemals militärisch genutzten Konversionsareal „US-Hospital“. Durch den Abzug der US-amerikanischen Streitkräfte aus Heidelberg im Jahr 2013 öffnete sich ein Gelegenheitsfenster für zivilgesellschaftliche Akteure, die verschiedenen innerstädtisch gut eingebetteten Flächen im Zuge des Konversionsprozesses mitzugestalten.

4.2 Gebäudestruktur

Aus Suffizienz-Perspektive und in Anbetracht des Netto-Null-Flächenverbrauchsziels der Bundesregierung bis 2050 [14] ist die **Nutzung von bestehenden Strukturen** gegenüber Neubauten zu priorisieren. Die ökologischste Form des Bauens ist immer, nicht zu bauen bzw. nicht neu zu bauen. So ist zum Beispiel „ein fast materieloser Waldkindergarten [...] immer suffizienter als ein noch so ökologisch gebauter „Gebäudekindergarten“ [15].

An dem Standort des zukünftigen CA befand sich ein Krankenhaus sowie ein umliegender Parkplatz. Durch den flächensparenden Ansatz und die Herstellung nutzbarer Freiflächen in Form von Obstbaum- und Magerwiesen und einer Retentionsfläche rund um das Gebäude findet in Summe eine teilweise Entsiegelung der Fläche statt. Neben dem Neubau werden zwei Bestandsgebäude saniert und in das Ensemble integriert. Darin sind weitere Wohnflächen, Seminarräume, Werkstätten, ein Café sowie Verwaltungsräume geplant.

Neben der Nutzung des Bestehenden muss die Gebäudestruktur auch fit für sich zukünftig ändernde Bedürfnisse sein, das heißt **Anpassungen** ermöglichen - auch um die Akzeptanz gegenüber neuen Konzepten und Wohnformen zu Beginn des „Suffizienzzeitalters“ und die damit einhergehenden Risiken aus Investor:innen-Perspektive zu reduzieren. [15]

Der Neubau des CA ermöglicht dies, indem innerhalb der Dreier- und Vierer-Wohngemeinschaften ein flexibles Zusammenspiel aus individueller und geteilter Fläche möglich ist: Die Einzelzimmer bestehen jeweils aus einer Kernzone (7 m²) sowie einer flexiblen Zone (7 m²). Durch eine verstellbare Schiebewand können die Bewohner:innen zwischen den beiden Ausbauvarianten wählen. Langfristig können die Wohnungen zu größeren Wohneinheiten zusammengeschaltet werden. Das flexible Raumkonzept wird durch die Holzskelettbauweise erleichtert, wodurch das Tragwerk mit nur wenigen tragenden Wänden auskommt.

4.3 Architektur

Eine der zentralen Kenngrößen suffizienter Architektur ist die **Pro-Kopf-Wohn- bzw. -Bruttogrundfläche**. Sie bestimmt in hohem Maße die Performance in allen drei Nachhaltigkeitsfeldern – Ökologie, Ökonomie und Soziales. Die derzeit stetige Steigerung der Pro-Kopf-Flächen konterkariert jedoch als Rebound-Effekt bisherige Einsparungen durch Effizienz- und Konsistenzmaßnahmen [16]. Gleichzeitig korreliert die Fläche nicht direkt mit der empfundenen Wohnqualität, die seit den 80er-Jahren nicht mehr angestiegen ist. [17] Aus diesen Gründen und wegen der zahlreichen gebauten Beispiele sollte die Pro-Kopf-Wohnfläche möglichst unter 35 m² pro Person reduziert werden, was allen voran durch **gemeinschaftliche Wohnformen** und -grundrisse erzielt werden kann.

Das Zusammenleben im CA findet in kompakten Dreier- und Vierer-WGs statt, wodurch die Pro-Kopf-Wohnfläche innerhalb der WGs bei etwa 22 m² liegt. Je nach individuell gewählter Ausbauvariante der Einzelzimmer variiert die WG-interne Gemeinschaftsfläche zwischen 21 und 49 m². Unter anteiliger Hinzunahme der gesamten Gemeinschaftsflächen des Wohnheims von etwa 535 m² erhöht sich die Wohn- und Nutzfläche pro Kopf auf ca. 25 m². Durch die Erschließung der Wohneinheiten über Laubengänge wird komplett auf innenliegende Verkehrsflächen verzichtet.

Im Zentrum dieser Planung steht damit ein Umdenken im Bereich der Flächennutzung. Die Flexibilisierung und Pluralisierung der Grundrisse erhöht die Handlungs- und Nutzungsmöglichkeiten bei gleichbleibender Wohnungsgröße. Durch hochwertige gemeinschaftlich nutzbare Räume können zudem Wohnfunktionen außerhalb der rein privaten Wohnfläche erfüllt werden. Es findet dabei eine Verschiebung von privat genutzter Fläche zu zugänglichen Flächen statt. Die Gemeinschaftsräume bieten darüber hinaus die Möglichkeit, gemeinschaftliche Versorgungsangebote wie Teilen, Leihen und Tauschen vor Ort zu organisieren. Durch eine geschickte Optimierung der Wohn- und Nutzflächen entsteht auf diese Weise eine Verdichtung an Nutzungsangeboten und eine nachbarschaftliche Vernetzung, die zu einer erhöhten Wohnqualität bei geringerer Wohnfläche führen können.



Abb. 2. Zimmerprototyp des CA. Bild: Johannes Roßnagel

4.4 Bautechnik

Suffizienz auf bautechnischer Ebene zielt auf Einfachheit bzw. Komplexitätsreduktion, also der Anwendung des **Low-Tech-Prinzips** auf baukonstruktiver Ebene. Dies kann unter anderem durch monolithische Bauweisen erzielt werden. Daneben gilt es Ausbaustandards zu reduzieren. Dies erfolgt indem beispielsweise letzte Schichten weggelassen oder nicht ganz fertig gebaut werden. Diese Methode ist damit ein sehr plakatives Beispiel zur „Anwendung des Pareto-Prinzips auf das Bauen“ [18] und wirkt sich dabei kaum mindernd auf die Lebensqualität aus, reduziert aber Umweltauswirkungen und Kosten in großem Umfang.

Das CA verfügt in manchen Bereichen über einen reduzierten Ausbaustandard: Zum Beispiel werden Holzfasade und -wände weder behandelt, verputzt noch gestrichen. Ziel war einerseits eine Kostenreduktion, aber andererseits auch, die natürlichen Eigenschaften und Ästhetik der Holzkonstruktion sichtbar zu machen (Abb. 2). In den Gemeinschaftsflächen wird auf einen zusätzlichen Bodenbelag verzichtet und stattdessen eine Beschichtung des Estrichs verwendet. Bei der Elektroinstallation wird überwiegend auf Aufputzvarianten gesetzt.

4.5 Gebäudetechnik

Bei der technischen Gebäudeausrüstung steht Suffizienz für ein **minimalisiertes und robustes Technikkonzept**, das auf passive Maßnahmen, Selbstregeleffekte sowie Redundanz-Vermeidung setzt. Es werden Komforteinbußen akzeptiert, um einen wartungsarmen und unkomplizierten Betrieb zu ermöglichen. Lowtech-Technikkonzepte müssen jedoch nicht zwingend zu einem energiesuffizienten Verhalten führen. Durch die bessere Handhabbarkeit und geringere Fehleranfälligkeit ist jedoch ein suffizienter Betrieb der Bewohner:innen in einem Lowtech-Gebäude wahrscheinlicher als in einem Hightech-Gebäude [19].

Zur Wärmeversorgung ist das CA an das Fernwärmenetz angeschlossen. Das Heizen erfolgt über Fußbodenheizungen mit einzelnen Heizkreisläufen pro Zimmer sowie in den Badezimmern über Infrarotstrahler. Aufgrund des erreichten Passivhausstandards ist mit einem sehr niedrigen Heizwärmebedarf zu rechnen. Die Warmwasserbereitung findet über dezentrale Wohnungsstationen statt. Zur dezentralen Lüftung befindet sich in jedem Zimmer und dem WG-Gemeinschaftsraum je einen Pendellüfter mit manueller Bedienung. Durch den Paarbetrieb können diese vom Pendelbetrieb auf ein Durchlüften umgeschaltet werden, um im Sommer einen gewissen Kühlungseffekt zu erzeugen. Da die Lüfter pro Wohneinheit über einen eigenen Stromkreis laufen, können sie alternativ in der Nicht-Heizperiode abgeschaltet werden. Die dezentralen Lösungen ermöglichen individualisierte und bedürfnisbasierte Einstellungen in den Zimmern und Wohneinheiten. Zum Monitoring des persönlichen Stromverbrauchs steht den Bewohner:innen eine Feedback-App zur Verfügung.

4.6 Mobilitäts-Infrastruktur

Neben dem Standort hat die Mobilitäts-Infrastruktur des Gebäudes große Auswirkungen auf suffiziente Mobilitätspraktiken der Bewohner:innen. Es gilt einerseits emissions- und ressourcenintensiven Individualverkehr unattraktiv zu gestalten, nicht nur wegen der Umweltauswirkungen während der Nutzung. Bereits die Erstellung eines Tiefgaragen-PKW-Stellplatzes emittiert mehr als ein:e Deutsche:r pro Jahr [20]. Mit einem entsprechenden Mobilitätskonzept können reduzierte Stellplatzschlüssel oder sogar **PKW-freie Gebäude und Quartiere** ermöglicht werden. Andererseits muss der **Fahrradkomfort** – als suffiziente Fortbewegungsform – maximiert werden, indem Anzahl und Qualität der Fahrradstellplätze erhöht werden.

Im Sinne eines nachhaltigen Mobilitätskonzepts legt das CA den Schwerpunkt auf Fahrrad- und ÖPNV-Mobilität. Nach den Möglichkeiten der Stellplatzverordnung wurde die Zahl der PKW-Stellplätze für die gut 240 Wohnplätze auf etwa 35 reduziert, wobei davon allein 25 auf die Versammlungsstätte entfallen und nicht weiter reduziert werden konnten. Die PKW-Stellplätze sollen vorrangig in

Form von Car-Sharing genutzt und im Alltagsbetrieb temporär umgenutzt werden können.

4.7 Gebäudemanagement

Um die Suffizienz-Planungsziele weiter in die Nutzungsphase zu tragen und ein suffizientes Nutzer:innen-Verhalten anzuregen, müssen auch entsprechende Maßnahmen im Gebäudemanagement ergriffen werden. Zum einen bedeutet diese eine adäquate Belegungsdichte sicherzustellen und **Leerstands-Vermeidungs-Strategien** umzusetzen. Zum anderen kann durch entsprechende Angebote von dafür vorgesehenen Flächen, zum Beispiel Hobby-Werkstätten oder Bibliotheken, die **geteilte Nutzung von Gütern und Dienstleistungen** als Suffizienzpraktik gefördert werden. Und nicht nur in der Planungsphase, sondern auch im Betrieb des Gebäudes ermöglicht die Partizipation der Bewohner:innen eine konstante Nachjustierung des Gebäudes an die tatsächlichen Bedürfnisse. Zu guter Letzt stellen auch **Nutzer:innen-Handbücher und Nachhaltigkeitsleitfäden** für Bewohner:innen und das Facility Management die Weichen für einen suffizienten Gebäudebetrieb.

Das CA richtet sich als Wohnheim vorrangig an Studierende und Auszubildende, weshalb eine Einzelzimmervermietung für eine hohe Belegungsdichte sorgt. Die bezahlbare Miete und das Teilen von Konsumgütern ermöglichen den Bewohner:innen niedrige Lebenshaltungskosten und damit die Reduktion von Erwerbsarbeitszeit und schaffen die Voraussetzung für individuellen Zeitwohlstand und zeitlichen Spielraum für die aktive Mitgestaltung des Zusammenlebens [21].

Das Wohnheim wird in Selbstverwaltung von den Bewohner:innen getragen, so dass diese für den Betrieb gemeinschaftlich verantwortlich sind. Jede:r Bewohner:in übernimmt für die Wohndauer eine Aufgabe und stimmt sich in Arbeitsgruppen und Gremien mit den anderen Bewohner:innen ab. Hinter der Übernahme der Verantwortung für den Betrieb des Gebäudes steht zum einen ein Bildungsgedanke, zum anderen aber auch das Ziel, dass die Bewohner:innen das Gebäude stets an ihre Bedürfnisse anpassen können und sich das Gebäude mit ihnen weiterentwickelt. Die Bewohner:innen erhalten ein Benutzungshandbuch, in dem der richtige Umgang mit der Gebäudetechnik und Anregungen zum energiesparenden Verhalten enthalten sind.

Mit einer Werkstatt für Selbstbau und Reparatur, einer Gemeinschaftsküche und Nutzflächen zum Anbau von Obst und Gemüse im Außenraum stehen den Bewohner:innen hochwertige Räume zur geteilten Nutzung zur Verfügung, die ressourcenschonende Praktiken erleichtern. In der Aula und den Seminarräumen ist darüber hinaus ein vielfältiges Bildungs- und Kulturangebot geplant, das auch Interessierten über die Bewohner:innen hinaus offensteht.

5. Zusammenfassung und Ausblick

Zur Erreichung der Nachhaltigkeitsziele ist im Gebäudereich ein Umdenken in der baulichen und technischen Gestaltung von Gebäuden erforderlich. Dazu kann jede:r in der alltäglichen Planungspraxis durch kritisches Hinterfragen von Konventionen beitragen, indem unter anderem flächensparsame und ausbaureduzierte Gebäudekonzepte entwickelt werden.

Das CA in Heidelberg zeigt, wie durch ausdauerndes Engagement aus der Zivilgesellschaft heraus ein beispielgebendes Bauprojekt entstehen kann, das die Bedürfnisse der Bewohner:innen spiegelt, Raum für Suffizienzpraktiken schafft und dabei die ökologischen Grenzen respektiert. Damit solche Konzepte in den Mainstream gelangen, bedarf es jedoch richtungssicherer, an der Einhaltung der Nachhaltigkeitsziele orientierter Weichenstellungen auf politischer Ebene, die Suffizienz beim Planen, Bauen und Wohnen zum Standard machen.

Danksagung

Die Autor:innen bedanken sich bei der ehrenamtlichen Projektgruppe des CA für ihr unermüdliches Engagement im Aufbau des Wohnheims sowie die Mitwirkung am Forschungsprojekt „SuPraStadt“. Ein weiterer Dank geht an das Bundesministerium für Bildung und Forschung (BMBF), das das Projekt „SuPraStadt“ und damit den Transfer der gewonnenen Erfahrungen in Praxis, Wissenschaft und Politik durch die Förderung der sozial-ökologischen Forschung im Rahmen der FONA-Strategie ermöglicht.

Literaturverzeichnis

- [1] Bundesregierung (2020): *Klimaschutzbericht 2019 zum Aktionsprogramm Klimaschutz 2020 der Bundesregierung*. https://www.bmu.de/fileadmin/Daten_BMU/Download_PDF/Klimaschutz/klimaschutzbericht_2019_kabinettsfassung_bf.pdf
- [2] Umweltbundesamt (2018): *Die Nutzung natürlicher Ressourcen. Bericht für Deutschland 2018*. https://www.umweltbundesamt.de/sites/default/files/medien/3521/publikationen/deuress18_de_bericht_web_f.pdf
- [3] Destatis (2021a): *Online-Plattform der Deutschen Nachhaltigkeitsstrategie (DNS), Flächeninanspruchnahme - Indikator 11.1.a: Anstieg der Siedlungs- und Verkehrsfläche*. <https://sustainabledevelopment-deutschland.github.io/11-1-a/>
- [4] Destatis (2021b): *Gesellschaft und Umwelt – Wohnen*. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Wohnen/_inhalt.html
- [5] Fischer, C., & Grießhammer, R. (2013): *Mehr als nur weniger. Suffizienz: Begriff, Begründung und Potentiale*. Öko-Institut Working Paper 2/2013, S. 10.
- [6] Brischke, L.-A. (2020): *Inhalte und politische Verankerung einer Suffizienzstrategie im Energiebereich. Briefe. Zur Orientierung im Konflikt Mensch - Erde*. 41 (137), S. 11-18, Evangelische Akademie Sachsen-

- Anhalt, Wittenberg. https://ev-akademie-wittenberg.de/sites/default/files/publikationen/briefe_2020-4.pdf
- [7] Over, M., Brischke, L.-A., & Leuser, L. (2020): Das selbstverwaltete Studierendenwohnheim Collegium Academicum Heidelberg: Suffizienz lernen, leben und verbreiten. In A. Brokow-Loga und F. Eckardt (Hrsg.), *Postwachstumsstadt. Konturen einer solidarischen Stadtpolitik* (S. 306-319). oekom Verlag München
- [8] Zimmermann, P. (2018): *Bewertbarkeit und ökobilanzieller Einfluss von Suffizienz im Gebäudebereich*. Masterarbeit, Technische Universität München
- [9] Weitere Informationen zum Collegium Academicum: <https://collegiumacademicum.de>
- [10] Weitere Informationen zum Förderprogramm „Variowohnungen“ des Bundesinstituts für Bau-, Stadt- und Raumforschung: <https://www.zukunftbau.de/variowohnungen/>
- [11] Weitere Informationen zum EFRE Förderprogramm Holz-Innovativ des Ministeriums für ländlichen Raum und Verbraucherschutz Baden-Württemberg: <https://efre-bw.de/foerderaufruf/aufruf-zum-foerderprogramm-holz-innovativ/>
- [12] Weitere Informationen zum BMBF-geförderten Forschungsprojekt „SuPraStadt – Lebensqualität, Teilhabe und Ressourcenschonung durch soziale Diffusion von Suffizienzpraktiken in Stadtquartiere“: <https://www.ifeu.de/projekt/suprastadt/>
- [13] Weitere Informationen zum Mietshäuser Syndikat: <https://www.syndikat.org/de/>
- [14] BMU (2016): *Klimaschutzplan 2050*.
- [15] Steffen, A., & Fuchs, M. (2015): Weniger ist weniger - und anders. *db deutsche bauzeitung* 06.2015, S. 26-30.
- [16] Kopatz, M. (2016): *Kommunale Suffizienzpolitik – Strategische Perspektiven für Städte, Länder und Bund*. Kurzstudie des Wuppertal Instituts für Klima, Umwelt, Energie.
- [17] DIW SOEP (2015): *SOEP 2013 - SOEPmonitor Household 1984-2013 (SOEP v30)*, S. 33.
- [18] Dömer, K., Drexler, H., & Schultz-Granberg, J. (2016): *Bezahlbar. Gut. Wohnen. Strategien für einen erschwinglichen Wohnraum*. Berlin: Jovis Verlag.
- [19] Preisler, A., Berger, M., & Gasser, B. (2016): *ANNEX III: Low Tech - Suffizienz in der Haustechnik*.
- [20] Lang, W., & Schneider, P. (2017): *Gemeinschaftlich nachhaltig bauen - Forschungsbericht der ökologischen Untersuchung des genossenschaftlichen Wohnungsbauprojektes wagnisART*.
- [21] Brischke, L.-A., & Over, M. (2020): *Zeitwohlstand durch Suffizienz in einem Reallabor: Wie bauliche und organisatorische Strukturen die Zeitnutzung verändern*. In: *Ökologisches Wirtschaften* 4/2020 (35), Schwerpunktthema: *Sozial-ökologische Infrastrukturen*, S. 27-29. IÖW/VÖW Berlin (Hrsg.), oekom Verlag München.

Wohnraum: Wird die Qualität in Quadratmetern gemessen?

Lena-Katharina Peter, Anja Bierwirth

Wuppertal Institut für Klima, Umwelt, Energie gGmbH, Döppersberg 19, 42103 Wuppertal

Das Projekt „Flächennutzung optimieren – Neubaudruck mindern (OptiWohn)“ untersucht kommunale Handlungsmöglichkeiten, um in wachsenden Städten dem Wohnraumbedarf nicht (nur) mit Neubau, sondern vor allem im Bestand zu begegnen. Denn auch in den beteiligten Projektstädten Göttingen, Köln und Tübingen finden sich leerstehende oder mindergenutzte Gebäude und Wohnungen. Im Rahmen einer Online-Umfrage haben rund 2.500 Teilnehmende auf Fragen zu ihrer Wohnsituation und der wahrgenommenen Wohnqualität geantwortet. Trotz recht überwiegender Zufriedenheit zeigt die Auswertung eine weit verbreitete Bereitschaft auf Seiten der Wohnenden, weniger Wohnraum in Anspruch zu nehmen. Der Beitrag stellt ausgewählte Ergebnisse und mögliche Handlungsoptionen vor, um das Potenzial einer optimierten Wohnflächennutzung zu erschließen.

The project „Optimizing the use of living space – reducing the pressure to build new buildings (OptiWohn)“ examines municipal options for action to meet the demand for housing in growing cities not (only) with new construction, but with an optimized use of the existing building stock. Vacant or under-used buildings and apartments can also be found in the participating project cities of Göttingen, Cologne and Tübingen. In a survey, about 2,500 participants answered questions about their housing situation and perceived quality of living. Despite a rather high satisfaction the evaluation shows a broad willingness of the residents to take up less living space. The article outlines some of the most important findings and points out possible courses of action to unlock this potential for an optimized use of living space.

1. Hintergrund

Wachsende Städte in Deutschland stehen heutzutage vor einer großen Herausforderung: Das urbane Wohnraumangebot kann mit der steigenden Nachfrage kaum noch mithalten. Dies hat sowohl soziale, als auch ökologische Folgen. Durch den immer schnelleren Anstieg der Miet- und Immobilienpreise haben vor allem Familien mit Kindern sowie einkommensschwächere Haushalte zunehmend Probleme, eine adäquate und bezahlbare Wohnung zu finden. Gleichzeitig verbrauchen Neubaugenossenschaften auf der grünen Wiese knappe Naturräume und kostbare Ressourcen (Böcker et al. 2020: 52 f.). Die Entwicklung der individuellen Wohnpräferenzen verschärft das Problem noch weiter: Die Anzahl der Personen in einem Haushalt hat sich in den letzten Jahren verkleinert, während die in Anspruch genommene Wohnfläche immer größer wurde. Ein- und Zwei-Personenhaushalte sind in Deutschland mittlerweile die häufigsten Haushaltstypen, die durchschnittliche Wohnfläche pro Kopf beträgt dabei 46,5qm (Flade 2020: 133, 193 f., Kaltenbrunner & Waltersbacher 2019: 103, Kohl et al. 2019: 20, Voigtländer 2017: 53 f.). Bis zum Jahre 2030 prognostizieren Forscherinnen und Forscher sogar einen Anstieg der Pro-Kopf-Wohnfläche auf 51,5qm (Deschermeier & Henger 2015: 23, von Einem 2016: 146 f.).

Das Prinzip des suffizienten Wohnens kann eine Antwort auf diese Herausforderungen sein. Dabei geht es in erster Linie darum, den Wohnraum optimal an die aktuelle Lebenssituation anzupassen und so einen Überkonsum von Wohnfläche zu vermeiden. Dem städtischen Wohnraumangel kann demnach nicht (nur) mit Neubau,

sondern vor allem mit einer besseren Nutzung bereits bestehender Gebäude begegnet werden (Böcker et al. 2020: 52 f., Hasse 2019: 93, Kaltenbrunner & Waltersbacher 2019: 102).

Diesen Ansatz verfolgt auch das Projekt „Flächennutzung optimieren – Neubaudruck mindern (OptiWohn)“ in Zusammenarbeit mit den Projektstädten Göttingen, Köln und Tübingen. Bevor jedoch kommunale Strategien zu einer optimierten Wohnraumnutzung entwickelt werden können, gilt es, die Bedarfe der Bevölkerung vor Ort zu ermitteln. Aus diesem Grund wurde im September 2020 eine Online-Umfrage durchgeführt, in deren Rahmen rund 2.500 Teilnehmende auf Fragen zur aktuellen Wohnsituation, der empfundenen Wohnqualität und der individuellen Zufriedenheit mit dem derzeitigen Wohnraum geantwortet haben. Dabei interessierte das Projektteam vor allem, inwiefern die tatsächliche Wohnfläche einen Einfluss auf die erlebte Wohnqualität ausübt und ob die befragten Personen im Sinne des suffizienten Wohnens zu einer Anpassung ihrer aktuellen Wohnfläche bereit wären.

Der aktuelle Forschungsstand betrachtet die Wohnungsgröße als einen wichtigen Einflussfaktor im Zusammenhang mit der empfundenen Wohnqualität, verweist aber zugleich auf ein Paradox: Objektive und subjektive Qualitätsmerkmale des Wohnraums sowie der Wohnumgebung können durchaus völlig konträr zueinander verlaufen (Flade 2020: 88, 138, 145, Schreiner 2014: 80). So kann eine Person mit ihrer Wohnumwelt zum Beispiel sehr zufrieden sein, obwohl die objektive Wohnqualität eher als schlecht zu beurteilen ist und umgekehrt. Die

erlebte Wohnqualität ist demnach eine subjektive Bewertung, die von persönlichen Merkmalen, individuellen Wohnbiografien sowie dem jeweiligen Lebensstil und Qualitätsmaßstab einer Person beeinflusst wird (Flade 2020: 128, Schreiner 2014: 75). Hinzu kommt, dass sich diese Wahrnehmung des eigenen Wohnraums mit dem Lebenslauf eines Menschen verändert: Alleinlebende Berufseinsteiger stellen zum Beispiel andere Ansprüche an ihre Wohnung als Familien mit Kindern oder ältere Menschen (Hasse 2019: 92 f.).

Die tatsächliche Wohnsituation passt sich den veränderten Ansprüchen allerdings nicht immer an. So leben Eltern nach dem Auszug der Kinder zum Beispiel häufig auf großer Wohnfläche weiter, während Familien mit Kindern Schwierigkeiten haben, eine größere und bezahlbare Wohnung in ähnlicher Wohnlage zu finden. An diesen Zielgruppen könnte eine kommunale Strategie der Wohnraumoptimierung ansetzen, denn die Umfrageergebnisse zeigen durchaus eine Bereitschaft seitens der Bewohnerinnen und Bewohner, ihre Wohnfläche besser an die aktuelle Lebenssituation anzupassen.

2. Methodisches Vorgehen

Die Umfrage im Rahmen des OptiWohn-Projektes zeichnet sich besonders durch ihren explorativen Charakter aus, der einen ersten Zugang zu den Themen Wohnqualität, Wohnzufriedenheit und Wohnsuffizienz sichern soll. Vor allem letzteres ist bislang kaum in größeren Befragungen erfasst worden. Mit Hilfe eines Online-Fragebogens, der von dem OptiWohn-Projektteam gemeinsam mit zwei Masterstudierenden entwickelt und in einem Pretest geprüft wurde, sollen erste Hypothesen über die Umsetzungsmöglichkeiten einer optimierten Wohnflächennutzung formuliert werden. Der verwendete Fragebogen umfasst daher Fragen zu individuellen Einstellungen und Bewertungen, die ein breites Spektrum an Merkmalen der Wohnumgebung und des aktuellen Wohnraums abdecken. Darüber hinaus wurde auch nach vergangenen sowie zukünftigen Wohnvorstellungen gefragt. Mittels des Schneeball-Systems wurde die Umfrage über persönliche Netzwerke und Kontakte des Projektteams sowie über thematisch relevante E-Mail-Verteiler und Seiten in den sozialen Medien verbreitet. Da keine repräsentativen Ergebnisse, sondern vielmehr eine Exploration des Themenfeldes angestrebt wird, wurde auf eine zufällige Auswahl der Umfrageteilnehmerinnen und -teilnehmer verzichtet.

Für die Auswertung der Online-Umfrage wurde der Datensatz zunächst bereinigt. So ergab sich eine Gesamtzahl von 2.434 vollständig ausgefüllten Fragebögen, deren Antworten in einem ersten Schritt mit Hilfe von deskriptiven Untersuchungen ausgewertet wurden. Um mögliche statistische Zusammenhänge zwischen interessierenden Variablen aufzudecken, wurden die Daten anschließend mit dem Statistikprogramm R weiterverarbeitet. Zur Berechnung der Zusammenhänge wurde der Rangkorrelationskoeffizient Kendalls Tau verwendet, da die hier vorliegenden Daten ordinal skaliert sind und es

bei der relativ kleinen Anzahl von Antwortkategorien recht wahrscheinlich ist, dass statistische Bindungen vorliegen. Kendalls Tau prüft, ob ein monotoner Zusammenhang zwischen zwei Rangreihen vorliegt und ist dabei weniger anfällig gegenüber Ausreißern als der Rangkorrelationskoeffizient Spearman's Rho (Bortz & Lienert 2008: 290 ff.).

3. Ergebnisse der Online-Umfrage

3.1. Wohnsituation

In Bezug auf die Wohnsituation der Umfrageteilnehmerinnen und -teilnehmer interessierten das OptiWohn-Projektteam vor allem folgende Fragen: Wie wohnen die befragten Personen aktuell? Leben sie alleine, mit Verwandten in einem Haushalt oder in einer gemeinschaftlichen Wohnform? Wie groß ist die durchschnittliche Wohnfläche pro Kopf? Die Ergebnisse zeigen, dass die Mehrheit der befragten Personen in einer städtischen Region (72%) lebt und sich zum Zeitpunkt der Umfrage in einem Mietverhältnis (56%) befindet. Die Hälfte (50%) wohnt in einem Mehrfamilienhaus mit bis zu 12 Wohnungen, die durchschnittliche monatliche Kaltmiete liegt bei rund 745,00€.

Zudem wohnt die Mehrheit der befragten Personen (38%) in einem 2-Personen-Haushalt, meist zusammen mit der Partnerin oder dem Partner (47%). Der Anteil der alleinlebenden Personen ist mit 19% etwas geringer als der Anteil der Personen, die mit einem oder mehreren Kindern zusammen im Haushalt wohnen (24%). Diese Beobachtungen decken sich auch mit den häufigsten Wohnformen in ganz Deutschland (1- und 2-Personenhaushalte). Für die Pro-Kopf-Wohnfläche ergibt sich ein Mittelwert von 50qm. Da der Wert der durchschnittlichen Pro-Kopf-Wohnfläche jedoch von einzelnen sehr großen oder extrem kleinen Wohnflächengrößen beeinflusst werden kann, sollte an dieser Stelle ergänzend festgehalten werden, dass die meisten befragten Personen (23%) eine Wohnfläche zwischen 31 und 40 Quadratmetern pro Kopf aufweisen. Die Pro-Kopf-Wohnfläche der Umfrageteilnehmerinnen und -teilnehmer liegt daher leicht unter dem deutschlandweiten Durchschnitt von 46,5qm.

Für das Projekt OptiWohn spielen flächensparende Wohnformen eine besondere Rolle. Zum Zeitpunkt der Umfrage leben rund 8% der teilnehmenden Personen in einer gemeinschaftlichen Wohnform, die Mehrheit davon (66%) in einer Wohngemeinschaft innerhalb einer Wohnung.

3.2. Bewertung der Wohnqualität

Vor dem Hintergrund der aufgezeigten Wohnsituationen der Umfrageteilnehmerinnen und -teilnehmer wurde deren Bewertung der individuellen Wohnqualität näher betrachtet: Wie zufrieden sind die befragten Personen mit ihrem Wohnraum? Welche Faktoren spielen für diese Zufriedenheit eine besondere Rolle? Und welche zusätzlichen Variablen, wie zum Beispiel das Alter, die

Wohnfläche oder die Haushaltsgröße, können die empfundene Wohnqualität beeinflussen?

Insgesamt ist die Mehrheit aller befragten Personen „sehr zufrieden“ (43%) oder „eher zufrieden“ (35%) mit ihrer aktuellen Wohnsituation. Die individuelle Qualität der Wohnumgebung wird von den meisten Befragten als „gut“ (34%) oder „sehr gut“ (32%) bewertet. Die wichtigsten Faktoren für die empfundene Qualität der Wohnsituation sind dabei die Kosten des Wohnens (22%), eine ruhige Wohnlage (14%) sowie eine gute Verkehrsanbindung (12%). Daneben wird die tatsächliche Wohnzufriedenheit aller befragten Personen jedoch durch weitere Faktoren beeinflusst, die bewusst oder unbewusst auf die persönliche Bewertung der Wohnsituation einwirken. So weisen z.B. folgende Merkmale eine leichte statistische Korrelation mit der individuell wahrgenommenen Wohnzufriedenheit auf:

- Privater Raum im Freien (Kendalls Tau = 0,17)
- Wohnverhältnis (Kendalls Tau = 0,28)
- Wohnfläche (Kendalls Tau = 0,24)

Demnach steigt die individuelle Wohnzufriedenheit der Umfrageteilnehmerinnen und -teilnehmer mit zunehmender Wohnungsgröße an (siehe Abbildung 1). Ein vergleichbarer Anstieg der empfundenen Wohnzufriedenheit lässt sich auch bei einer Zunahme der Pro-Kopf-Wohnfläche beobachten. Keinen erkennbaren Einfluss haben dagegen die Anzahl der Personen im Haushalt, die Stadt-Land-Verteilung der befragten Personen sowie die möglichen Veränderungen im Zuge der Corona-Pandemie.

Neben der empfundenen Wohnqualität und Wohnzufriedenheit, wurde in der Online-Umfrage auch nach einer persönlichen Bewertung der aktuellen Wohnfläche gefragt. Diese nimmt die Mehrheit der befragten Personen als „genau richtig“ (53%) wahr. Dabei besteht ein statistischer Zusammenhang zu der tatsächlichen Größe der Wohnfläche (Kendalls Tau = 0,32); großer Wohnraum wird dementsprechend auch als „etwas zu groß“ oder „viel zu groß“ empfunden.

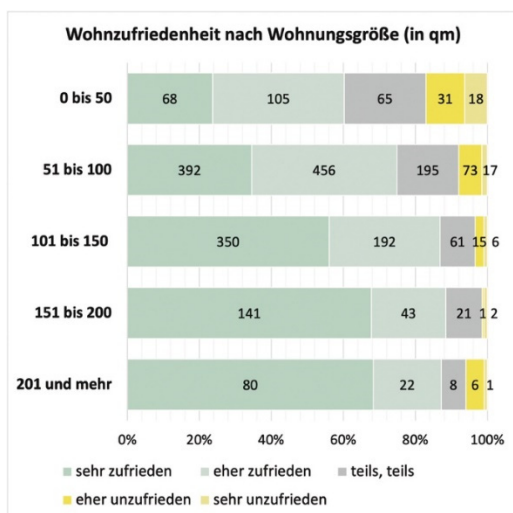


Abbildung 1: Wohnzufriedenheit nach Wohnungsgröße

3.3. Wohnberatung und Veränderungsbereitschaft

Im letzten Abschnitt der Online-Umfrage waren mögliche zukünftige Veränderungen und Pläne der befragten Personen in Bezug auf deren Wohnsituation von Interesse.

Zunächst lässt sich festhalten, dass die Mehrheit der Umfrageteilnehmerinnen und -teilnehmer in der nächsten Zeit eher zum Verbleib in der aktuellen Wohnsituation tendiert (46%). Setzt man die Frage nach der Veränderungstendenz jedoch in einen Zusammenhang mit einer möglichen Änderung der Personenzahl im Haushalt, ergibt sich ein anderes Bild: 42% der befragten Haushalte, die in den nächsten 10 Jahren mit einer Verringerung der Personenzahl rechnen, können sich vorstellen, in eine kleinere Wohnung umzuziehen und so ihren Wohnraum der aktuellen Lebenssituation anzupassen. Bei den Haushalten, die zukünftig von ihrer Auflösung ausgehen, z.B. durch einen Umzug, liegt die Verkleinerungsbereitschaft sogar bei 45% (siehe Abbildung 2).

An dieser Stelle ergibt sich ein wichtiger Ansatzpunkt für eine Wohnraumberatung im Sinne von OptiWohn: So gilt es, den Haushalten, die sich z.B. durch den Auszug der Kinder verkleinern, adäquate Umzugsmöglichkeiten anzubieten oder Wege aufzuzeigen, wie sie in ihrer aktuellen Wohnsituation bleiben können, aber dennoch den frei gewordenen Wohnraum besser nutzen können (z.B. durch Untervermietung, Umbau oder gemeinschaftliche Wohnformen).

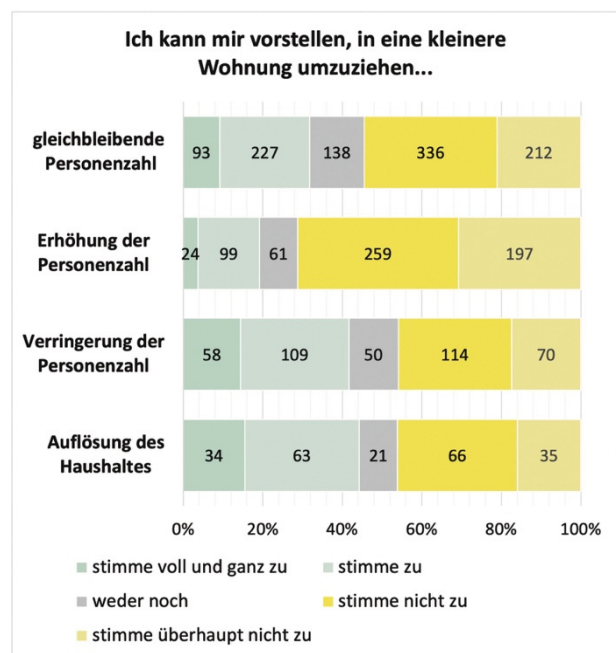


Abbildung 2: Verkleinerungsbereitschaft nach geschätzter Veränderung der Haushaltsgröße in den nächsten 10 Jahren

Das Potenzial für Wohnraumveränderungen lässt sich in den Ergebnissen der Online-Umfrage deutlich erkennen:

- 31% der befragten Personen können sich vorstellen, in eine kleinere Wohnung umzuziehen.
- 26% können sich vorstellen, ihr Haus umzubauen, um den Einzug weiterer Personen zu ermöglichen.

- 51% aller befragten Personen können sich gemeinschaftliches Wohnen vorstellen.

Dabei gibt es einige individuelle Merkmale, die die persönliche Bereitschaft für eine Wohnraumveränderung beeinflussen können. So lässt sich z.B. ein leichter statistischer Zusammenhang zwischen der Bereitschaft, sich wohnlich zu verkleinern und dem Alter der befragten Personen erkennen (Kendalls Tau = 0,14): Ältere Personen können sich demnach eher eine Verkleinerung des Wohnraums vorstellen als jüngere Befragte.

Auch die Bewertung der eigenen Wohnfläche hat einen Einfluss auf die individuelle Verkleinerungsbereitschaft (Kendalls Tau = 0,46): Wer seine Wohnfläche als etwas oder viel zu groß empfindet, ist auch eher bereit, sich zu verkleinern. Andere Faktoren wie z.B. das Haushaltseinkommen, die aktuelle Personenzahl im Haushalt, die Wohnsituation im Eigentum oder in einem Mietverhältnis, das Geschlecht der befragten Personen oder die monatliche Kaltmiete haben dagegen keinen nennenswerten Einfluss auf die Bereitschaft, in eine kleinere Wohnung zu ziehen.

Das größte Hemmnis in Bezug auf eine tatsächliche Veränderung der Wohnsituation, wie z.B. einen Umzug, stellt für die Umfrageteilnehmerinnen und -teilnehmer die Schwierigkeit dar, eine „geeignete und preisgünstige Wohnung zu finden“ (87%). Zwei weitere wichtige Hinderungsgründe liegen in dem Wunsch der befragten Personen, sowohl in ihrem aktuellen Ortsteil (63%), als auch in ihrer aktuellen Wohnung bzw. Haus wohnen zu bleiben (62%).

Um diese Hemmnisse abzubauen, könnten sich kommunale Beratungs- und Unterstützungsangebote zum Thema Wohnen als hilfreich erweisen. Das Interesse an solchen Angeboten ist unter den befragten Personen jedenfalls recht groß: Die Mehrheit (36%) findet die Beratungs- und Unterstützungsangebote „interessant“ oder „sehr interessant“. Dabei stoßen vor allem die Angebote einer „Wohnberatung an einer zentralen Stelle“, „Beratung zum Wohnraumtausch“ sowie „Unterstützung bei Bauanträgen, etc.“ auf großes Interesse unter den Umfrageteilnehmerinnen und -teilnehmern. Bisher sind derlei Angebote jedoch nur 11% der befragten Personen bekannt.

Zuletzt hat die Online-Umfrage auch nach den Hauptmotiven bzw. übergeordneten Zielen gefragt, die Personen dazu bewegen, eine Verkleinerung ihres Wohnraumes vorzunehmen oder in einem gemeinschaftlichen Wohnprojekt zu leben. So kann sich die Mehrheit der befragten Personen vor allem dann eine Verkleinerung der Wohnfläche vorstellen, wenn dies zur Schonung von Ressourcen (17%) oder zur Klimaneutralität (16%) beiträgt. Für das Leben in einer gemeinschaftlichen Wohnform spricht für die befragten Personen in erster Linie eine mögliche Kostenersparnis bei Wohnen und Haushalt (23%). Dieses Motiv verliert jedoch mit zunehmendem Alter der Umfrageteilnehmerinnen und -teilnehmer an Wichtigkeit (Kendalls Tau = -0,12), sodass mehr Raum

für soziale Anreize entsteht. Ebenso steigt auch die generelle Bereitschaft, gemeinschaftlich zu wohnen, leicht an, je älter die befragten Personen sind (Kendalls Tau = 0,11). Somit könnten vor allem ältere Menschen, deren Kinder bereits ausgezogen sind, eine wichtige Zielgruppe für gemeinschaftliche Wohnprojekte darstellen.

4. Zusammenfassung

Im Zuge der Umfrageauswertung lassen sich folgende Hypothesen aufstellen, die mit Blick auf das Thema des suffizienten Wohnens von großer Relevanz sein könnten:

- Eine Bereitschaft für Wohnraumveränderungen, wie zum Beispiel Verkleinerungen oder der Umzug in eine gemeinschaftliche Wohnform, ist deutlich erkennbar; Das Interesse an Beratungs- und Unterstützungsangeboten zum Thema Wohnraum ist groß.
- Vor allem die Zielgruppe der älteren Menschen sowie der Personen, deren Haushalt sich z.B. durch den Auszug der Kinder verkleinert hat, steht Wohnraumveränderungen offen gegenüber und könnte sich von adäquaten Angeboten der Wohnraumberatungen angesprochen fühlen.
- Die größten Herausforderungen für die Umsetzung einer kommunalen Strategie zur Förderung von suffizientem Wohnen liegen in der starken kleinräumigen Verwurzelung der Bewohnerinnen und Bewohner sowie in deren mangelndem Wissen über Angebote und Alternativen zum Thema Wohnraum und Wohnraumberatung.

Nun gilt es, diese Hypothesen in die Entwicklung der kommunalen Strategien in den OptiWohn-Projektstädten Göttingen, Köln und Tübingen mit einfließen zu lassen und so das flächensparende Wohnen vor Ort zu fördern. Außerdem können die vorgestellten Ergebnisse als Basis für weitere Fragestellungen und Untersuchungen dienen.

Literaturverzeichnis

- [1] Böcker, M., Brüggemann, H., Christ, M., Knak, A., Lage, J. & Sommer, B. (2020): Wie wird weniger genug? Suffizienz als Strategie für eine nachhaltige Stadtentwicklung. München: oekom.
- [2] Bortz, J. & Lienert, G. A. (2008): Kurzgefasste Statistik für die klinische Forschung: Leitfaden für die verteilungsfreie Analyse kleiner Stichproben. Heidelberg: Springer Medizin Verlag.
- [3] Deschermeier, P.; Henger, R. (2015): Die Bedeutung des zukünftigen Kohorteneffekts auf den Wohnflächenkonsum. In: IW-Trends – Vierteljahresschrift zur empirischen Wirtschaftsforschung 42, 3, 23-39.
- [4] Flade, A. (2020): Wohnen in der individualisierten Gesellschaft – Psychologisch kommentiert. Hamburg: Springer.
- [5] Hasse, J. (2019): Wohnen – eine existenzielle Herausforderung. In: Bürger & Staat 2/3 – 2019, 88-93.

- [6] Kaltenbrunner, R. & Waltersbacher, M. (2019): Wohnungsbau: Zwischen Ökonomie und Gesellschaft. In: Bürger & Staat 2/3 – 2019, 101-107.
- [7] Kohl, S., Sagner, P. & Voigtländer, M. (2019): Mangelware Wohnraum – Ökonomische Folgen des Mietpreisbooms in deutschen Großstädten. Düsseldorf: Forschungsinstitut für gesellschaftliche Weiterentwicklung (Hrsg.).
- [8] Schreiner, K. (2014): Zur Bedeutung von Umweltmerkmalen und -aneignungen im Wohnen – Konstruktion eines Fragebogens zum Wohlfühlen/Zuhausefühlen in Wohnungen. Online: <https://unipub.uni-graz.at/obvugrhs/content/titelinfo/762451/full.pdf> (Stand: 26.01.2021).
- [9] Voigtländer, M. (2017): Luxusgut Wohnen – Warum unsere Städte immer teurer werden und was jetzt zu tun ist. Köln: Springer.
- [10] Von Einem, E. (2016): Wohnen. Markt in Schiefelage – Politik in Not. Berlin: Springer.

Wohnflächeneffizienz möchte erhöht werden

Arne Steffen

werk.um architekten, Rheinstraße 99, 64295 Darmstadt

Das Themenfeld Bauen trägt wesentlich zum Klimawandel bei. Bisher wird versucht, mit Effizienz- und Konsistenzmaßnahmen in Erstellung und Betrieb gegen eine weitere Aufheizung zu wirken. Die Nachhaltigkeitsstrategie der Suffizienz wird bisher kaum verfolgt, ein vor kurzem begonnenes Fallbeispiel zur Erhöhung von Wohnflächeneffizienz wird nachfolgend dargelegt.

Einleitung

Der Untersuchungsgegenstand: Ein Wohnquartier mit 1091 Wohneinheiten im Eigentum einer kommunalen Wohnungsgesellschaft. Dort wird aktuell untersucht, inwieweit die Wohnflächeneffizienz in dem Quartier erhöht werden kann. Anhand von diesem Beispiel soll ein Muster einer Untersuchungsstruktur entwickelt werden, welche an zwei weiteren Fallbeispielen überprüft werden soll.

Ausgangssituation Themenfeld Wohnen

Obwohl man sich seit Jahrzehnten von der Gesetzgebung ja bemüht, die Rahmenbedingungen für das Themenfeld Bauen so zu setzen, dass die Energieverbräuche sinken müssten, ist das Verbrauchsniveau in Deutschland stabil geblieben.

Schon 1977 wurde unter Bundeskanzler Helmut Schmidt eine erste Wärmeschutzverordnung verabschiedet, in welcher der Wärmedurchgang in Bauteilen und Wärmeverluste durch Undichtigkeiten begrenzt wurde (Wikipedia, 2019). Aus der Wärmeschutzverordnung wurde eine Energieeinsparverordnung (EnEV), welcher 2020 das Gebäude-Energiegesetz folgte. Die Anforderungen an Gebäude in Hinblick auf Energieverbrauch stieg mit jeder Novellierung und wird in wenigen Jahren für Neubauten den Standard des Nullenergiehauses für den Betrieb erreicht haben.

So konnte seitdem der durchschnittliche theoretische Verbrauch von Wärmeenergie pro Quadratmeter Wohnfläche in Deutschland von 250 kWh/m²a auf ca. 100 kWh/m²a gesenkt werden. Da jedoch im annähernd gleichen Zeitraum die durchschnittliche Wohnfläche pro Kopf sich von ca. 28m² (1972 waren es 26,4 m² (DER SPIEGEL, 2016)) auf 47m² in 2020 erhöht hat, ist der Wärmeenergieverbrauch pro Kopf nicht gesunken. Damit hat die erste Wärmeschutzverordnung es möglich gemacht, dass wir heute auf fast doppelt so viel Wohnfläche leben, ohne mehr zu verbrauchen als damals, aber das eigentliche Ziel der Energieeinsparung krachend verfehlt.

In den Verordnungen zum Wärme- und Energieschutz von Gebäuden werden ausschließlich die Nachhaltigkeitsstrategien von Effizienz und (in den vergangenen Novellierungen) Konsistenz zugrunde gelegt. Die Suffizienz als eine den Verbrauch tatsächlich begrenzende Strategie wird an keiner Stelle verfolgt.

Um nun die immer dringender werdenden Einsparungen auch im Themenfeld Wohnen zu erreichen, müssen unbe-

dingt – so die Auffassung des Verfassers – Suffizienzstrategien verfolgt werden. Damit diese allerdings auch erfolgreich sein können, müssen die Gründe für die nichtsuffiziente Entwicklung verstanden werden, um Potenziale untersuchen und quantifizieren zu können. Auf der Grundlage können hoffentlich attraktive Angebote an die Bewohner*innen gemacht werden, suffizienter = flächensparender zu wohnen. Verordnungen zur Suffizienz sind zurzeit nicht vorstellbar.

Verständnis zur Entwicklung

Welche Umstände haben zu der nahezu ununterbrochenen Zunahme von Wohnfläche pro Kopf in Deutschland geführt?

Insbesondere zwei Treiber haben den Trend befördert

- Kleinere Haushalte
In Deutschland besteht ein Trend zu Haushalten mit weniger Personen. Laut dem Umweltbundesamt lag der Anteil von Ein-Personenhaushalten 2019 bei 39,7%, im Durchschnitt leben nur noch zwei Personen pro Haushalt. Die durchschnittliche Wohnfläche von Ein-Personen-Haushalten lag 2018 schon bei 68m²; damit lag der Flächenverbrauch der alleinlebenden Menschen um 44% über dem Durchschnitt der Bevölkerung in Deutschland. Bereits Zwei-Personen-Haushalte benötigen im Schnitt vergleichsweise wenige 49m² pro Person (Umweltbundesamt, 2020).
- Alte Menschen in großen Wohnungen
Die Daten des Bundesinstituts für Bau-, Stadt- und Raumforschung (BBSR) und des Statistischen Bundesamtes zeigen, dass viele ältere Menschen in den Wohnungen bleiben, in welchen sie früher viele Jahre mit mehr Bewohner*innen gelebt haben. Diese Entwicklung wird das empty-nest-Phänomen genannt: Eltern ziehen nach dem Wegzug der Kinder oder auch Auszug/ Tod von Partnern selten in eine kleinere Bleibe. Besonders Eigentümer bleiben gerne in der großen Familienwohnung. Dazu kommt, dass erstens der Anteil älterer Wohneigentumsbesitzer in den letzten Jahren noch zugenommen hat auf 56% der über 65-jährigen. Und zweitens, dass die älteren Menschen nochmal überproportional groß wohnen: laut BBSR lebten die über 75-jährigen in Ein-Personen-Haushalten 2010 schon auf üppigen 78 m² (Umweltbundesamt, 2020).

Die Entwicklung macht deutlich, dass ein Gegensteuern dringend geboten ist. Denn Wohnfläche in Häusern muss nicht nur beheizt, mit Strom versorgt, renoviert und möbliert, sondern auch erschlossen werden. Das führt unverändert zu hohem Ressourcenverbrauch und der entsprechenden Flächenversiegelung (Umweltbundesamt, 2020).

Es stellt sich allerdings die Frage, ob die Menschen freiwillig in den großen oder zu groß gewordenen Wohnungen leben oder ob es kein Angebot beziehungsweise keine Anregung und Unterstützung beim Wechsel in eine besser passende Wohnung gab?

Die weiteren, grundlegenden und inneren fünf Barrieren von Konsument*innen hat Oliver Stengel 2011 näher beschrieben:

- Barriere 1: Der Preis für den Verlust von empfundenen Selbstverständlichkeiten blockiert Lebensstilveränderungen.
- Barriere 2: Das Konsumniveau entspricht in der Regel dem Status von Menschen. Ein hohes Konsumniveau eines ressourcenlastigen Lebens „beweist“ Erfolg und Wohlstand.
- Barriere 3: Solange „Alle“ so leben, gibt es keine Veranlassung, einen eigenen Lebensstil zu hinterfragen und korrigieren.
- Barriere 4: Die Verantwortung für den klimaschädigenden Lebensstil wird zwischen Politik und Konsument*innen hin und her geschoben. Beide Seiten erwarten Signale, dass man handeln sollte.
- Barriere 5: Der Kapitalismus lebt vom Mehr und muss für das Überleben zum Konsum verführen.

Um die Barrieren überwinden zu wollen, um bereit zu sein, die von Oliver Stengel so genannten persönlichen Kosten zu „bezahlen“, braucht es Nutzen für die Menschen. Naheliegende Vorteile von einem Wohnen auf weniger Fläche sind:

- Reduktion der Wohnkosten,
- Näherrücken an andere Menschen,
- Verringerung von Einsamkeit,
- Entlastung bei der Wohnungspflege.

Diese Nutzen sind in der Kommunikation herauszustellen.

Wege zur höheren Wohnflächeneffizienz

Einerseits ist viel Platz natürlich eine Qualität, an verschiedenen Orten in einer Wohnung sein zu können. Vielleicht auch ein extra Arbeitszimmer, ein Gäste- oder sogar ein Fitnesszimmer nutzen zu können. Doch führt eine höhere Wohnfläche pro Kopf logischerweise dazu, dass die Räume seltener genutzt werden – auch wenn diese Flächen durchgehend bezahlt, beheizt, instandgehalten werden müssen.

Folgende Angebote könnten Anreize schaffen, sich in der Fläche zu verkleinern ohne sich räumlich einzuschränken:

- Angebote an pfiffigen kleineren Wohnungen, die tatsächlich günstiger sind als das große bisherige Zuhause,
- Angebot an zubuchbaren Räumen wie Gästezimmer, Arbeitszimmer,
- Angebot an gemeinschaftlich zu nutzenden Räumen wie Fitness- oder Eventräume,
- Hilfe zur Veränderung: von der Suche nach einer situationsgerechten kleineren Wohnung bis zum Umzug.

Die oben aufgeführten Raumkonzepte werden im Rahmen von sogenannten Gemeinschaftlichen Wohnprojekten, die oft von privaten Wohngruppen entwickelt werden, angeboten. Doch bedarf es in solchen Initiativen ein hohes Niveau von Gemeinschaftsenergie, und es stellt sich beim Auszug der Erstbewohner*innen heraus, dass der Gemeinschaftsgeist in den Projekten oft von der motivierten Gründergeneration gelebt hat. Oftmals verweisen die Gemeinschaftsvorteile solcher Häuser im Laufe der Zeit.

Fallstudie Berliner Viertel in Rüsselsheim

Beteiligte: gewobau Rüsselsheim, Wuppertal Institut für Klima, Umwelt und Energie gGmbH, werk.um architekten als Forschungspartner.

Anlass: Die gewobau hat im Rahmen einer internen Zukunftswerkstatt nach Lösungsansätzen gesucht, um mehr Anfragen nach Wohnungen befriedigen zu können (als die in der Gesellschaft bewältigbaren 50 Neubauwohneinheiten pro Jahr) und eine Senkung der Quote von Neumieter*innen, die nach weniger einem Jahr Mietzeit wieder ausziehen. Am Beispiel vom Berliner Viertel soll ein neues Modell von Wohnraumbewirtschaftung durchgeplant und gegebenenfalls umgesetzt werden.



Berliner Viertel in Rüsselsheim

Frage: Ist es möglich, die Prinzipien zum Gemeinschaftlichen Wohnen auf eine Quartiersebene zu übertragen, in der Hoffnung, für alle Mieter*innen in Quantität und Qualität die optimalen Flächen und Wohnformen und Wohnräume zur Verfügung zu stellen?

Gliederung und Struktur der Bearbeitung

In Kooperation mit der gewobau wurde die nachfolgende prinzipielle Struktur erarbeitet:

- Identifikation von Quartieren/ Siedlungen
Kriterien, die helfen, Potenziale einzuschätzen
- Durchführung einer Mieterbefragung
Generelle Bedarfserkundung und Interpretation von Ergebnissen
- Untersuchung von Grundrissvarianten und Umzugsszenarien
Welche Bestandsgebäude eignen sich für welche Umbauten?
- Analyse von Kosten und Nutzen
Aspekte einer Wirtschaftlichkeitsberechnung mit Kostenprognosen und Erlöschancen
- Durchführung einer zweiten Mieterbefragung
Konkrete Bedarfsanalyse auf Grundlage eines konkreten Angebots
- Herbeiführen von Entscheidungen der Beteiligten
Aufführung von Notwendigkeiten, Abhängigkeiten, zu beachtenden Aspekten
- Umsetzung = Change-Prozess
Auflisten von Aktionen zur Begleitung des Prozesses: Presseinformationen, Umzugsunterstützung, Umzugsfeste, ...

In der anschließenden Auswertung der Ergebnisse in Hinblick auf die Nachhaltigkeit wird untersucht:

- Energetische Bilanz für Erstellung und Betrieb
- Bilanz des Ressourcenverbrauchs
- Bilanz vom Flächenverbrauch
- Betriebswirtschaftliche Auswertung

Das abschließende Fazit zeigt auf, ob und - wenn nicht - unter welchen voraussichtlichen Voraussetzungen, das Ziel der erhöhten Wohnflächeneffizienz erreicht werden kann.

Erste Ergebnisse Berliner Viertel

In diesem Fall wurde das Berliner Viertel nicht aufgrund im Vorfeld untersuchter Kriterien aus verschiedenen Quartieren ausgewählt, sondern war der „Pilotfall“, an dem eine vorher ausgewählte Konzeption versucht wird. Möglicherweise werden die Erfahrungen, die im Berliner Viertel gemacht werden, dazu dienen, die Kriterienliste anzupassen.

1. Durchführung einer Mieterbefragung

Der Fragebogen wurde durch die gewobau vorbereitet, vom Wuppertal Institut in Hinsicht auf Informationen um Wohnflächeneffizienz ergänzt wurde.

Im Findungsprozess wurde deutlich, dass

- es Sinn macht, das Thema Fragebogen für nächste Quartiere zu vertiefen,
- eine jeweils anzupassende Standard-Vorlage weiteren Projekten sehr dienlich sein könnte,
- die Ergebnisse unbedingt auch in Hinblick auf die Verbesserung des Fragebogens ausgewertet werden müssen.

Es ist parallel in zwei Quartieren die Bewohnerschaft befragt worden. Der Rücklauf im Berliner Viertel betrug 86 verwertbare Fragebögen (von 1091 WE im Berliner Viertel) = 7,9%. Damit kann qualitativ weitergearbeitet werden, nur bedingt quantitativ.

Die wichtigsten Ergebnisse für den zu prüfenden Umbau des Quartiers:

Wohnungstauschbereitschaft

Es besteht eine höhere Wohnungstauschbereitschaft als erwartet. Auch wenn diese nicht wie erhofft von Groß nach Klein besteht, sondern überwiegend in der gegenteiligen Richtung: 50% der Befragten ziehen einen Wohnungstausch in eine größere Wohnung in Betracht und nur 20% wollen sich verkleinern.

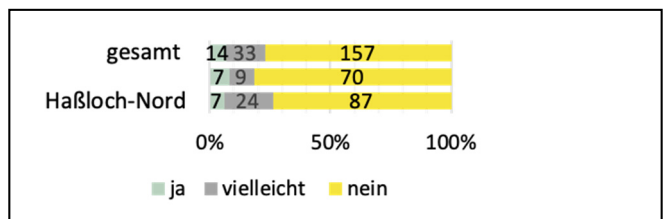


Abbildung 1: Wohnungstauschbereitschaft gegen eine kleinere Wohnung

Das Ergebnis motiviert jedoch die Forschenden und Architekt*innen, das Bedürfnis derer, die in eine größere Wohnung wollen, zu relativieren mit neuen Angeboten im Quartier.

Barrierefreiheit

Die Barrierefreiheit hat in dem Quartier einen höheren Stellenwert als vermutet. Für 24% der Befragten wäre Barrierefreiheit ein Umzugsanlass.

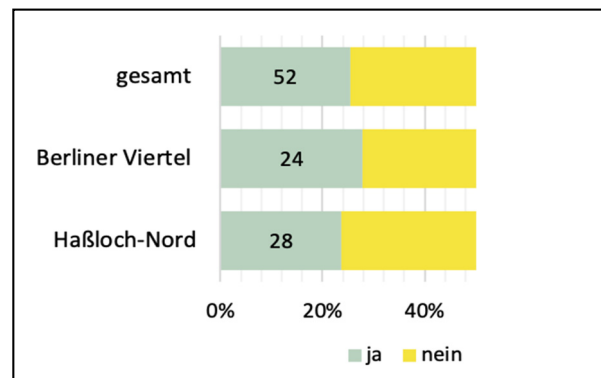


Abbildung 2: Barrierefreiheit als Grund für Wohnungstausch

Gemeinschaftliches Wohnen

Eine bisher in entsprechenden Quartieren sicherlich selten gestellte Frage, war, ob die Bewohner*innen Interesse an gemeinschaftlichen Wohnformen haben? Das Ergebnis hat sehr überrascht: 40% der Befragten sind an einer der abgefragten Formen des gemeinschaftlichen Wohnens interessiert.

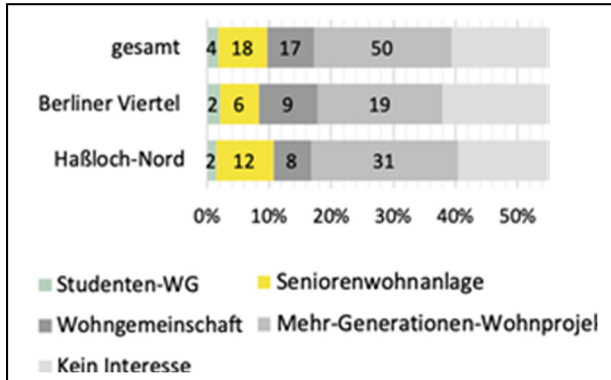


Abbildung 3: Interesse an gemeinschaftlichen Wohnformen

Zubuchbare Räume

Ein Angebot, das es schon vereinzelt in Wohnquartieren gegeben hat, das aber vermutlich noch nicht in der Konsequenz abgefragt wurde. Das Ergebnis hat die Fragenden jedenfalls gefreut: über 40% haben Bedarf.

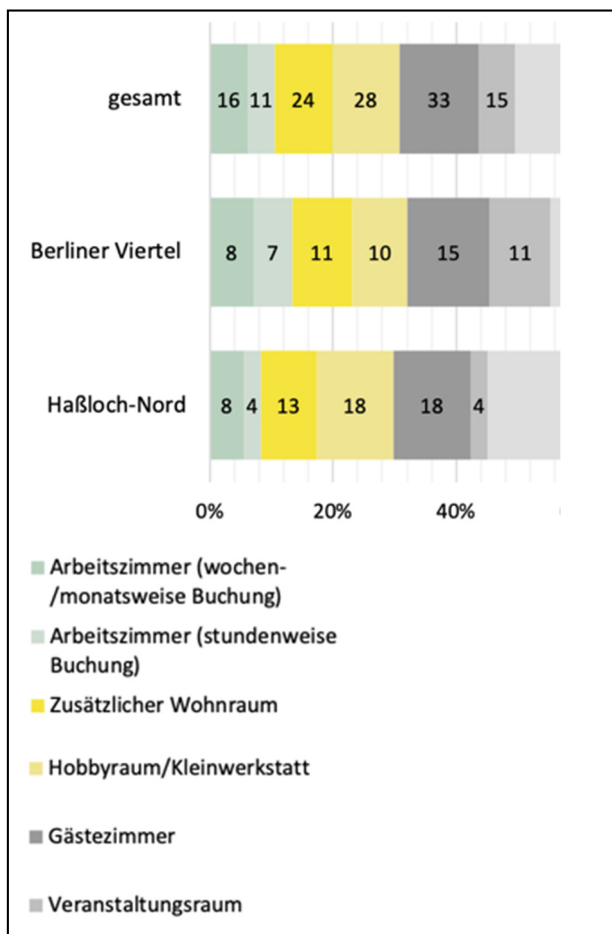


Abbildung 4: Interesse an zubuchbaren Räumen

Auf Basis dieser Ergebnisse werden nun verschiedene Szenarien skizziert, um festzulegen,

- wie viele barrierefreie Wohnungen,
- welche Formen des gemeinschaftlichen Wohnens in welcher Menge,
- welche Art und in welcher Anzahl zubuchbarer Räume

im Quartier angeboten werden sollten.

2. Untersuchung von Grundrissvarianten und Umzugsszenarien

Parallel zur Auswertung der Umfrageergebnisse wird der Bestand untersucht auf die wirtschaftliche Umbaubarkeit. Dabei wird geprüft,

- welche Wohnungen sich eignen, um barrierefrei oder mindestens barrierearm werden zu können,
- wie Wohnungen zusammengelegt werden können für Großfamilien oder Clusterwohngemeinschaften,
- wo günstigerweise welche Angebote an zubuchbaren Räumen platziert werden können.

Auf der Grundlage wird eine erste räumliche Projektskizze erstellt in nachfolgender qualitativer Art:

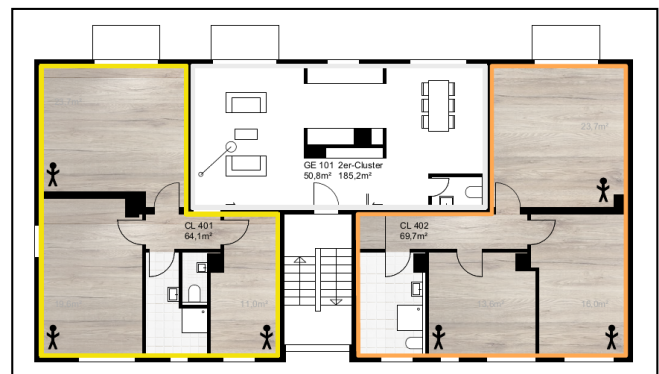


Abbildung 5: Umwandlung Zweispänner in Clusterwohnung

3. Analyse von Kosten und Nutzen

Um für alle Beteiligten nicht nur ökologische Nutzen zu generieren, sondern auch ökonomische Vorteile zu schaffen, muss frühzeitig das Verhältnis von Aufwand und möglichen Gewinnen geprüft werden. Dazu werden die Baukosten für die Projektskizze geschätzt, und in Zusammenarbeit mit dem Institut für Nachhaltigkeits- und Immobilienmanagement (INIM) und den Studierenden der HS Mittweida soll ein passendes Geschäfts- und Betriebsmodell entwickelt werden.

4. Durchführung einer zweiten Mieterbefragung

Sobald eine räumliche Plan- und Kostengrundlage vorliegt, wird die Bewohnerschaft ein weiteres Mal befragt: nun mit konkreten Plänen, mit Informationen zu Kosten, zu Ausstattung der neuen Raumangebote, mit Erläuterung der Betriebskonzepte, um nach dem anfangs abstrakten Bedarf dieses Mal einen tatsächlichen Bedarf abzufragen.

5. Herbeiführen von Entscheidungen der Beteiligten

Mit dem Ergebnis der zweiten Umfrage in der Mieterschaft kann dann die Planung vertieft werden. Dabei sollten folgende Grundsätze beachten werden:

- In 2-3 Abschnitten die Planung entwickeln.
- Eine kritische Masse für die neuen Nutzungsmodelle nicht unterschreiten.
- Fall-Back einplanen: Umbauten möglichst so planen, dass ohne viel Aufwand wieder zurückgebaut werden kann.
- Die Bewohnerschaft partizipativ einbinden in den Change-Prozess.

6. Umsetzung = Change-Prozess

Change-Prozesse sind immer aufregend. Besonders vermutlich für die betroffene Bewohnerschaft.

Um Sorgen über die Veränderungen gering zu halten, gilt es Vertrauen zu schaffen, dass alle auf den Weg „mitgenommen werden“, dass viel über Chancen und Nutzen informiert wird, dass Unterstützung von Erläuterungen bis zum Umzug angeboten wird.

Literaturverzeichnis

- [1] Wikipedia 2019. Internetrecherche 14.02.2021: <https://de.wikipedia.org/wiki/Wärmeschutzverordnung>
- [2] Umweltbundesamt 2020. Internetrecherche am 14.02.2021: <https://www.umweltbundesamt.de/daten/private-haushalte-konsum/wohnen/wohnflaeche#wohnflaeche-pro-kopf-gestiegen>
- [3] DER SPIEGEL 2016. Internetrecherche am 14.02.2021: <https://www.spiegel.de/spiegel/print/d-144989304.html>
- [4] O. Stengel, GAI/20/1(2011), 26–30, einsehbar auf: https://epub.wupperinst.org/frontdoor/deliver/index/docId/3754/file/3754_Stengel.pdf

Ökologischer Wandel und soziale Teilhabe

Ökologischer Wandel und Transformation: Neue Herausforderungen für die Hochschulkommunikation

Markus Drowatzky

Hochschule Mittweida, University of Applied Sciences, Technikumplatz 17, 09648 Mittweida

Die fortschreitende Digitalisierung, neue Wettbewerbssituationen und die Transformation zu nachhaltigen Organisationen stellen neue Herausforderungen für sämtliche Organisationsformen dar. Dieser, sich stetig wandelnde Zustand ist längst Alltag für Unternehmen, die ihre Kommunikation mit integrierten Konzepten darauf angepasst haben. Hochschulen dagegen fehlt es an einem integrierten Kommunikationskonzept, das der neuen Situation gerecht wird und dabei hilft, sich im zunehmenden Wettbewerbsumfeld nachhaltig zu platzieren. Der Beitrag diskutiert daher zunächst den Status quo der Hochschulkommunikation und untersucht, inwiefern Anknüpfungspunkte und Elemente zur Einführung eines integrierten Kommunikationskonzepts vorliegen. Die Untersuchung erfolgt durch eine Befragung der Verantwortlichen für Hochschulkommunikation mit dem Ziel, Anknüpfungspunkte zu diskutieren, wie und mit welchem Konzept sich Hochschulen kommunikativ auf die stetig wachsenden Herausforderungen und die Transformation des Umfeldes einstellen können.

1. Einleitung

Der Wettbewerb um Absatzmärkte, Aufmerksamkeit und die besten Arbeitskräfte ist für Unternehmen längst Alltag; der Umgang mit der Konkurrenz ist Tagesgeschäft. Dazu kommen die sich stetig wandelnden Anforderungen durch die Transformationsprozesse im Bereich der Digitalisierung und des ökologischen Wandels. Um diese Herausforderungen zu bewältigen, benötigen Unternehmen unter anderem etablierte und umfassende strategische Konzepte zur Kommunikation, die längst das Denken in klassischen Marketingdimensionen abgelöst haben. Unternehmen sind heute sogenannte „brands“ (Marken) geworden, deren Werte und Leistungen kommuniziert werden müssen (vgl. Kirchner 2001, 17), die sich im sich schnell wandelnden Markt etablieren und die eigene Leistungsfähigkeit herausstellen müssen.

Diesen kommunikativen Herausforderungen müssen sich auch deutsche Hochschulen zunehmend stellen. Der Wettbewerbsdruck hat sich durch private Anbieter und die Internationalisierung verschärft, die digitale Transformation bringt zudem neue Herausforderungen mit sich. Nationale und internationale Studierende haben die freie Standortwahl und entscheiden sich anhand der individuellen Kriterien und Merkmale für oder gegen eine Hochschule. Die konsistente Außendarstellung und die Verwendung von einprägsamen Botschaften zur Erreichung der avisierten Zielgruppen ist einer der Schlüssel für den Fortbestand. Jedoch haben Hochschulen in Deutschland in diesem Punkt noch Nachholbedarf. Orientierung können daher Unternehmen geben, die bereits etablierte Konzepte, wie die integrierte Kommunikation, einsetzen.

Der grundlegende Ansatz der Organisationskommunikation liegt darin, ein möglichst widerspruchsfreies, glaubhaftes und klares Bild gegenüber verschiedenen

Zielgruppen zu vermitteln, dass sowohl leicht wiedererkannt und gelernt wird und sich im besten Fall an den individuellen Kommunikationsbedürfnissen orientiert (vgl. Kirchner 2001: 33; Bruhn 2009: 1). Die Konzentration von Kommunikationsbemühungen führen zu höheren Erfolgsraten und einem konsistenten Bild (vgl. Bruhn 2009: 4). Dies ist das Ziel der integrierten Kommunikation, die gleichzeitig eine bessere Ausnutzung der Potentiale und Synergien der Kommunikationsaktivitäten verspricht (vgl. Kirchner 2001: 33f).

Es erscheint daher zweckmäßig, die Kommunikationsbemühungen und Strukturen der deutschen Hochschulen zunächst zu erfassen und anschließend mit der wissenschaftlich gut beschriebenen integrierten Kommunikation zu vergleichen, um zusätzliche Potentiale für die Kommunikationsarbeit zu beschreiben. Basis und Ausgangspunkt dieser Untersuchungen bildet die integrierte Kommunikation nach Bruhn, die nachfolgend kurz vorgestellt wird.

Integrierte Kommunikation

Das Konzept der integrierten Kommunikation basiert auf einem holistischen Ansatz der versucht, mit einer interdisziplinären Lösung die Unterteilung in Disziplinen zu überwinden. Laut Schultz geht es um den Perspektivwechsel hin zu einer Kommunikation, so wie sie die Kunden in ihrem Alltag wahrnehmen (Schultz et al. 1992: xvii). In einem vorangegangenen Paper hat der Autor bereits verschiedene Ansätze der integrierten Kommunikation untersucht und ist zum Schluss gekommen, dass die integrierte Kommunikation nach Bruhn für Hochschulen als Konzept die größte Praxisrelevanz aufweist (vgl. Drowatzky 2020). Für die nachfolgende Untersuchung wurde daher dieser Ansatz als Grundlage gewählt.

Nach Bruhn ist die integrierte Kommunikation ein strategischer und operativer Prozess der Analyse, Planung, Organisation, Durchführung und Kontrolle, der darauf

ausgerichtet ist, aus den differenzierten Quellen der internen und externen Kommunikation von Unternehmen eine Einheit herzustellen, um ein konsistentes Erscheinungsbild zu vermitteln (vgl. Bruhn 2009: 22). Es stellt sich daher die Frage, inwiefern Hochschulen bereits für ein solches Konzept empfänglich sind. Diese Frage führt über die Erhebung des Status quo der Hochschulkommunikation unter Berücksichtigung der Elemente der integrierten Kommunikation.

Aus dieser Erkenntnis ergibt sich folgende Forschungsfrage:

„Wenden Hochschulen in Deutschland bereits eine strategische und geplante Kommunikation an und inwiefern kann diese bereits als integrierte Kommunikation nach Bruhn beschrieben werden?“

2. Erhebungsverfahren und Analyse

Als Erhebungsverfahren zur Beantwortung der Forschungsfragen hat sich eine quantitative Befragung der Kommunikationsverantwortlichen der Hochschulen angeboten, mit dem Ziel, den Ist-Stand der Kommunikation von Hochschulen in Deutschland zu erfassen. Für die Befragung der Hochschulen wurde die Form eines standardisierten Online-Fragebogens gewählt, da sich dieser inzwischen zu einer etablierten empirischen Datenerhebungsmethode entwickelt hat (vgl. Zerback et al 2009: 15; Welker 2007:25f). Da quantitative Online-Befragungen einen hohen Grad an Präzisierung erfordern, wurden hauptsächlich geschlossene Fragen gestellt. Dabei wurde vor allem auf eine vierstufige Ratingskala (vgl. Möhring/Schlütz 2010: 88ff; Brosius 2016: 108) mit verbalisierten Skalenausprägungen zurückgegriffen, da diese gegenüber verbal polaren Skalen die Befragungsteilnehmer bei der Wahl ihrer Antworten unterstützt. Als Zeitraum der Befragung wurden vier Wochen im März 2020 gewählt und die Umfrage an 1203 Kommunikationsverantwortliche verteilt, die dem Bundesverband für Hochschulkommunikation angehören. Damit konnten rund 85 Prozent aller Hochschulkommunikationseinheiten erreicht werden.

Rücklauf und Datenauswertung

Von 1203 kontaktierten Kommunikationsverantwortlichen der Hochschulen nahmen insgesamt 114 Teilnehmer an der Umfrage teil. Das entspricht einer Rücklaufquote von rund 9,5 Prozent. Die statistische Auswertung der Daten erfolgte überwiegend deskriptiv mit einer Konzentration auf Summenstatistiken, Häufigkeitsverteilungen und Mittelwerte.

Auswertung der Ergebnisse

Die Untersuchung wurde mit dem Ziel durchgeführt, die folgende Fragestellung zu beantworten: „Wenden Hochschulen in Deutschland bereits eine strategische und geplante Kommunikation an und inwiefern kann diese bereits als integrierte Kommunikation nach Bruhn beschrieben werden?“

An der Umfrage unter den Mitgliedern des Bundesverbands Hochschulkommunikation haben insgesamt 114 Kommunikationsverantwortliche teilgenommen, davon stammten 54 Teilnehmer aus einer Universität und 45 von einer Fachhochschule, die im weiteren Verlauf der Untersuchung die relevanten Gruppen gebildet haben. Die Hochschulen selbst waren vergleichsweise groß (mehr als die Hälfte der Teilnehmer gab an, dass die Hochschule mehr als 5.000 Studierende hat) und entsprechen ungefähr dem Durchschnitt von 7.360 Studierenden pro Hochschule in Deutschland (vgl. HRK 2019). Dieser Vergleichswert ist auch in der Anzahl der Professoren abgebildet, sodass die Befragung in jedem Fall zur Lagebeschreibung der deutschen Universitäten und Fachhochschulen dient und dabei den Durchschnitt der Organisationen gut abbildet.

Zunächst wurden die Hochschulen dahingehend befragt, wie viele Abteilungen überhaupt mit den Kommunikationsaufgaben betraut sind. In der Erhebung des Status Quo hat sich gezeigt, dass 75 Prozent der befragten Hochschulen nur mit einer Kommunikationsabteilung ausgestattet sind, was nach Bruhns Einschätzung relevant für konzentrierte Kommunikationsaufgaben ist.

Im Verlauf der Befragung wurde zudem erfragt, wie denn die Wichtigkeit der Kommunikation eingeschätzt wird. Dabei bestand über die herausragende Bedeutung der externen Kommunikation innerhalb der Kommunikationsverantwortlichen Konsens. Diese Aussage trifft grundsätzlich auch auf die interne Kommunikation zu, die jedoch prozentual in der Wichtigkeit schlechter abgeschnitten hat. Für die interne und externe Kommunikation besteht innerhalb der integrierten Kommunikation der Auftrag, zu einem einheitlichen Gesamtbild über das Unternehmen oder die Hochschule beizutragen. Um abzuschätzen, inwiefern diese Absicht auch innerhalb der Kommunikationsabteilungen vorherrscht, wurde zunächst nach der Einstellung zur Einheitlichkeit gefragt. Insgesamt messen die Hochschulen dieser Einheitlichkeit eine große und sogar sehr große Bedeutung bei. Allerdings bestehen Unterschiede zwischen Universitäten und Fachhochschulen, denn letztere sind im prozentualen Vergleich in der Bewertung der Wichtigkeit der Einheitlichkeit der Darstellung verhaltener. Die Frage nach der operativen und strategischen Umsetzung dieser Thematik in der eigenen Hochschule zeigte jedoch erste Lücken für die integrierte Kommunikation auf. So scheinen zwischen der persönlichen Wahrnehmung der Wichtigkeit und der tatsächlichen Umsetzung Probleme aufzutreten. Ein Großteil der Befragten aus dem Kreis der Universitäten schätzt die Umsetzung dieser eigenen Einheitlichkeit nur eher weniger erfolgreich ein, was insbesondere auf die großen Universitäten (mehr als 10.000 Studierende) zutrifft. Die Teilnehmer der Fachhochschulen hingegen werten die Einheitlichkeit der Kommunikation als eher erfolgreich, haben aber gleichzeitig auch einen geringeren Anteil an Studierenden, so-

dass ein Zusammenhang zwischen der Größe der Hochschule anhand der Studierendenzahl und der Einheitlichkeit der Kommunikationsaktivitäten vermutet wird.

Nach Bruhns Verständnis folgt die integrierte Kommunikation vorher festgelegten Kommunikationszielen. Deshalb wurden die Hochschulen auch nach dem Vorhandensein dieser Kommunikationsziele gefragt. Universitäten verfügen mehrheitlich über Kommunikationsziele, Fachhochschulen dagegen mehrheitlich nicht. Zudem gibt es eine Ausprägung in den Größenverhältnissen nach Studierendenzahlen, sodass größeren Hochschulen insgesamt mehrheitlich Kommunikationsziele vorliegen.

Für Bruhn steht in seinem Konzept auch die Anbindung und der Auftrag der Kommunikationsabteilungen im Fokus. So hält er eine Zentralisierung der Kommunikationsaktivitäten in einer Abteilung, deren Anbindung an das Management (hier gleichbedeutend mit der Hochschulleitung) und die Zuständigkeit in der gleichen Abteilung für Erfolgskriterien, ebenso wie einen verschriftlichten Auftrag in Form eines Kommunikationskonzepts (vgl. Bruhn 2014).

Diejenigen Hochschulen, die über ein verschriftlichtes Kommunikationskonzept verfügen, bündeln auch überwiegend ihre Kommunikationsaktivitäten in einer Abteilung. Für die Untersuchung der Zuständigkeit für die Kommunikationsaktivitäten hat sich abgezeichnet, dass sowohl für die Planung als auch für die Umsetzung eine Kommunikations- bzw. Pressestelle verantwortlich ist, wobei dieser Begriff nicht trennscharf zu benutzen ist. In vielen Fällen ist zudem die Hochschulleitung in die Planung der Kommunikation eingebunden, in die Umsetzung jedoch in den seltensten Fällen.

Für eine strategische Kommunikationsarbeit ist eine Evaluation der Maßnahmen unerlässlich. Etwa die Hälfte der Befragten hat dieser Aussage zugestimmt, wobei die Universitäten ihre Kommunikationsmaßnahmen ausgeprägter evaluieren, als Fachhochschulen.

Relevant für die praktische Einführung der integrierten Kommunikation ist ein Verständnis dafür, dass die integrierte Kommunikation die Leistungsfähigkeit der Kommunikation verbessert. Im Fragebogen stimmt die deutliche Mehrheit der Kommunikationsverantwortlichen der Hochschulen diesem Verständnis zu. Dies unterstreicht die Relevanz des Konzepts sehr deutlich.

Kern der integrierten Kommunikation ist die inhaltlich, zeitliche und formale Integration der Maßnahmen und Instrumente. Auf dem Niveau von Einzelmaßnahmen, wähen sich die Hochschulen hier bereits gut aufgestellt. Dabei ist die formale Integration grundsätzlich am ausgeprägtesten. Dadurch verfestigt sich der Eindruck, dass die Hochschulen in der operativen Arbeit bereits einige Elemente der integrierten Kommunikation einsetzen, jedoch die strategische Anbindung fehlt.

Bei der Abfrage der Selbsteinschätzung der jeweiligen Hochschulen, in welchem Maß sie bereits integrierte Kommunikation betreiben, zeigt sich deutlicher Nachholbedarf, die Hochschulen wähen sich noch ganz am Anfang. Die Umfrage hat insgesamt gezeigt, dass die integrierte Kommunikation innerhalb der Hochschullandschaft noch nicht stark verbreitet ist, zum anderen aber ein Bedarf an diesem Kommunikationskonzept besteht.

Beantwortung der Forschungsfrage und Ausblick

Die Ergebnisse zum Status Quo der Hochschulkommunikation und hinsichtlich eines integrierten Kommunikationskonzepts fallen sehr uneinheitlich aus. Grundsätzlich braucht es ein Grundverständnis für die Kommunikation hinsichtlich interner und externer Kommunikation. Dies ist bei den Hochschulen weitestgehend vorhanden. Diese Aussage trifft auch auf das Ziel der Einheitlichkeit der Kommunikation zu, welches gleichzeitig eines der Kernziele der integrierten Kommunikation ist. Der Erfolg der Einheitlichkeit durch die aktuell umgesetzten Kommunikationsmaßnahmen ist jedoch innerhalb der Teilnehmer im Ergebnis heterogener, weshalb bereits hier ein Bedarf erkennbar ist. Im Rahmen der strategischen und planvollen Verankerung gibt es jedoch große Lücken und ein insgesamt heterogenes Bild. Das Kommunikationskonzept, festgelegte Kommunikationsziele und das Erreichen der Kommunikationsziele sind je nach Hochschulart und Größe der Hochschule anhand der Studierendenzahl stärker oder weniger stark ausgeprägt. Zudem zeigt sich ein Zusammenhang zwischen Hochschulgröße anhand der Studierendenzahl und dem Vorliegen dieser strategischen Aspekte. In der organisatorischen Verankerung liegen auch die Zentralisierung der personellen Ressourcen und der Kommunikationsaktivitäten. Hier zeigt sich, dass die Hochschulen mehrheitlich versuchen, in einer Abteilung zentral zu agieren, was dem Verständnis von Bruhn entspricht. Ebenso wird die Planung der Kommunikation mehrheitlich mit der Hochschulleitung abgestimmt, was ebenso dem Verständnis nach Bruhn entspricht.

Bei den Hochschulen ist insgesamt der Wunsch nach einer integrierten Kommunikation wahrnehmbar. Gleichzeitig schätzen die Hochschulen ein, dass sie selbst ihre Kommunikation noch nicht gut integrieren. Es ergibt sich daher auch ein differenziertes Bild zur Beantwortung der Forschungsfrage. Es besteht ein unmittelbarer Zusammenhang zwischen einem hohen Umsetzungsgrad der Integration der Hochschule und den positiven Antworten auf insbesondere strategische Kommunikationselemente aus dem Fragebogen, was bedeutet, dass hier ein Zusammenhang besteht. Gleichzeitig lässt sich kein Zusammenhang zwischen einer hohen Zustimmung zur individuellen Umsetzung eines integrierten Kommunikationskonzepts und der Größe der Hochschule ablesen. Dies führt zu zwei Erkenntnissen, denn die Umsetzung der integrierten Kommunikation scheint unabhängig von der Größe der Hochschule zu sein, sondern hängt

überwiegend davon ab, wie strategisch und geplant die Kommunikationsarbeit betrieben wird.

Der überwiegende Anteil der Hochschulen erfüllt Teile der festgelegten Kriterien zur integrierten Kommunikation. Teilweise kann von daher einer geplanten und auch strategischen Kommunikation gesprochen werden. Diese kann jedoch noch nicht vollständig als integrierte Kommunikation nach Bruhn beschrieben werden, da es insbesondere in den strategischen Elementen Nachholbedarf für alle Hochschulen gibt. Die integrierte Kommunikation ist dennoch nach Selbsteinschätzung der Kommunikationsverantwortlichen an Hochschulen ein wichtiger und für den Kommunikationserfolg wahrgenommenes Erfolgskriterium.

Schlussendlich verfestigt sich der Eindruck, dass die Hochschulen, anders als die Unternehmen im Bereich der strategischen Kommunikation dennoch nicht gut aufgestellt sind. In der Ableitung bleibt daher der Eindruck, dass die Hochschulen vor großen kommunikativen Herausforderungen stehen, wenn die Themenvielfalt durch den ökologischen Wandel und die digitale Transformation noch weiter zu nehmen. Es gilt, die in weiten Teilen klaffende Lücke zwischen Unternehmen und Hochschulen zu schließen, um nicht bald gänzlich den kommunikativen Anschluss zu verlieren und für die kommenden Herausforderungen nicht mehr gewappnet zu sein.

Einordnung der Befragung und Ausblick

Die Befragung der Kommunikationsverantwortlichen an Hochschulen fand vor Beginn der Corona-Pandemie statt. Durch die pandemiebedingte voranschreitende Digitalisierung vieler Prozesse ist zu vermuten, dass auch die Hochschulen im Bereich der Kommunikation notgedrungen vorangeschritten sind. Daher wäre es zur weiteren Erschließung des Themenfeldes relevant, die Umfrage nach der Corona-Pandemie noch einmal durchzuführen.

Literaturverzeichnis

- [1] Brosius, H.; Haas, A; Koschel, F (2016): Methoden der empirischen Kommunikationsforschung. Eine Einführung. 7., überarbeitete und aktualisierte Auflage. Wiesbaden: Springer VS (Studienbücher zur Kommunikations- und Medienwissenschaft). Online verfügbar unter <http://dx.doi.org/10.1007/978-3-531-19996-2>.

- [2] Bruhn, M. (2009): Integrierte Unternehmens- und Markenkommunikation. Strategische Planung und operative Umsetzung. 5., überarb. und aktualisierte Aufl. Stuttgart: Schäffer-Poeschel.
- [3] Bruhn, M. (2014): Integrierte Unternehmens- und Markenkommunikation: Strategische Planung und operative Umsetzung (6. Aufl.). Stuttgart: Schäffer-Poeschel.
- [4] Drowatzky, M. (2020). Integrated communication - theoretical construct or practical approach in Journal of Media Research, Vol. 13, Issue 2(37), 2020.
- [5] HRK (2019): Die Hochschulen als zentrale Akteure in Wissenschaft und Gesellschaft. https://www.hrk.de/fileadmin/redaktion/hrk/02-Dokumente/02-01-Beschluesse/HRK_-_Eckpunkte_Hochschulsystem_2019.pdf. Abgerufen am 23.02.2019.
- [6] Kirchner, K. (2001): Integrierte Unternehmenskommunikation. Theoretische und empirische Bestandsaufnahme und eine Analyse amerikanischer Großunternehmen. Wiesbaden: Westdeutscher Verlag.
- [7] Schultz, D. E.; Kitchen, P.(1992): "Integrated Marketing Communications in US Advertising Agencies: An Exploratory Study," Journal of Advertising Research, Vol. 37, no, 5, 1997, pp.7-17.
- [8] Möhring, W.; Schlütz, D. (2010): Die Befragung in der Medien- und Kommunikationswissenschaft.
- [9] Welker, M.; Sattler, S. (2007): Online-Befragung von Journalisten in Deutschland: Ein Modell zur Abschätzung von Coverage- und Responsefehlern bei einer amorphen und dispersen Großgruppe. In: Welker, M./Wenzel, O. (Hrsg.): Online-Forschung 2007. Grundlagen und Fallstudien. Köln: 333369 Wiesbaden: Westdeutscher Verlag.
- [10] Zerback, T. (2009): Zehn Jahre Sozialforschung im Internet eine Analyse zur Nutzung von Online-Befragungen in den Sozialwissenschaften. In: Jakob, M./Schoen, H./ Zerback, T. (Hrsg.): Sozialforschung im Internet. Methodologie und Praxis der Online-Befragung. Wiesbaden.

StakeholderInnen- und Kommunikationsmanagement sächsischer Energie- und Rohstoffunternehmen

Stefanie Walter

Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

Um den ökologischen Wandel zu ermöglichen, ist eine Energie- und Ressourcenwende nötig. Der bürgerschaftliche Protest gegen den Bau von erneuerbaren Energieumwandlungsanlagen sowie heimische Rohstoffgewinnung hat allerdings zugenommen. Teile der Bevölkerung sind sowohl mit neuen Vorhaben per se, als auch den Informations- und Beteiligungsmöglichkeiten unzufrieden. Unternehmen können auf Basis von StakeholderInnenanalysen sowie angepassten Kommunikationsmaßnahmen die projektbezogene Akzeptanz erhöhen. Die Befragung sächsischer UnternehmensvertreterInnen zeigt jedoch, dass die UnternehmerInnen sich der Vielzahl der StakeholderInnengruppen nicht bewusst sind und sie das Akzeptanzniveau der sächsischen Bevölkerung besser einschätzen, als es in Wirklichkeit ist. Zudem fehlen Unternehmen personelle und finanzielle Ressourcen für Analysen und die Umsetzung akzeptanzsteigernder Maßnahmen.

1. Einleitung: Die Bedeutung von Akzeptanz

Weltweit steigen Ressourcen-, Energie- und Rohstoffbedarf [1, 2, 3]. Allein durch Effizienzsteigerung, Einsparungen und Recycling kann der zukünftige Bedarf nicht gedeckt werden [4]. Entsprechend verfolgt Deutschland das Ziel, die Stromproduktion mittels erneuerbarer Energieumwandlungsanlagen (EEA) zu realisieren [5].

Erneuerbare Energie (EE) genießt in der Bevölkerung hohes Ansehen und gesellschaftliche Akzeptanz [6]. Dies gilt aber nur, bis eine Anlage vor Ort gebaut werden soll. In der Regel folgt auf die Bekanntgabe, dass ein Vorhaben geplant wird, bürgerschaftlicher Protest, welcher die Projektrealisierung verzögern oder sogar verhindern kann [7, 8, 9]. Folglich bestimmt die (In-)Akzeptanz der Bevölkerung die Zahl der EEA mit. Hinzu kommt, dass, um Strom aus EEA zu erzeugen, zu speichern und zu transportieren, u. a. Metalle wie Seltene Erden, Indium, Germanium und Platingruppenelemente in steigender Menge benötigt werden [10]. Sie müssen bergbaulich gewonnen werden, die erforderliche Akzeptanz für die heimische bergbauliche Rohstoffgewinnung schwindet in der deutschen Öffentlichkeit jedoch trotz hoher Umweltschutzstandards [11].

Die Realisierung des ökologischen Wandels hängt somit von der Auflösung des Spannungsfelds zwischen den Einzelinteressen der AnwohnerInnen, welche in ihrem Umfeld keine Anlagen möchten, und dem Gemeinwohl ab.

UnternehmerInnen müssen Akzeptanz in der Bevölkerung für Energie- und Rohstoffprojekte aktiv anstreben. Dazu muss die Ist-Situation bekannt sein:

- Der Grad der (In-)Akzeptanz der Bevölkerung und
- die Möglichkeiten der UnternehmerInnen durch Kommunikation mit den StakeholderInnen in Kontakt zu kommen, diese zu informieren, zu integrieren, zu beteiligen und so die Akzeptanz für Projekte zu erhöhen.

2. Theoretische Grundlage: Wie Akzeptanz zu definieren ist

Nachfolgend wird von Akzeptanz gesprochen, wenn eine positive Einstellung in Form einer passiven oder aktiven Zustimmung und Handlungsbereitschaft oder Handeln gegeben ist [12]. Inakzeptanz meint folglich eine negative Einstellung. Betrachtet wird in dieser Veröffentlichung die projektbezogene (In-)Akzeptanz; also die der Bevölkerung vor Ort, wenn ein Vorhaben umgesetzt werden soll. Zur Beurteilung der (In-)Akzeptanz wird sich der Akzeptanz-Inakzeptanz-Skala [13] in abgewandelter Form bedient (Tabelle 1, blau).

Bei der (In-)Akzeptanz handelt es sich um ein instabiles Konstrukt, welches sich im Zeitverlauf verändern kann [14]. Dies macht die Beurteilung ihrer Entwicklung herausfordernd, bildet aber gleichsam die Voraussetzung, kommunikative Maßnahmen strategisch zur Akzeptanzsteigerung einzusetzen. Kommunikationsmanagement dient dabei insbesondere dazu, den StakeholderInnen und ihren Bedürfnissen adäquate Aufmerksamkeit zukommen zu lassen [15]. Auf eine StakeholderInnenanalyse aufbauend werden Strategien und Maßnahmen entwickelt, umgesetzt und evaluiert [16].

3. Methodik: Befragung von Bevölkerung und Unternehmen bezüglich (In-)Akzeptanz

Für die Beurteilung der (In-)Akzeptanz wird sich der Umfrage bedient. Um repräsentative Ergebnisse zu ermöglichen, wird sie auf die Bevölkerung und Unternehmen im Freistaat Sachsen beschränkt. Die AnwohnerInnen sollen sich bezogen auf verschiedene Anlagearten einer der acht Stufen auf der Akzeptanz-Inakzeptanz-Skala zuordnen. Den Stufen ist jeweils ein Akzeptanzstufenwert von eins bis acht zugeordnet (Tabelle 1). Um eine bessere Übersicht über die (In-)Akzeptanz zu gewährleisten, sind die Stufen den drei Akzeptanzausprägungen *Akzeptanz*, *Gleichgültigkeit* und *Inakzeptanz* zugeordnet; entsprechend reicht der Akzeptanzausprägungswert von eins bis drei. Die Mittelwerte (\bar{x}) der durchschnittlichen

Akzeptanzstufe und durchschnittlichen Akzeptanzausprägung sind somit Ausdruck des Akzeptanzniveaus der Befragten.

Tabelle 1: Antwortmöglichkeiten, entsprechende Akzeptanzstufen/-ausprägungen samt Werten

| | | | | | | | | |
|---------------------|------------------------|--------------|-----------------|----------------------|-----------|------------------------|----------------|----------------------|
| Antwortmöglichkeit | setze mich dagegen ein | lehne ich ab | bin zwiespalten | ist mir gleichgültig | dulde ich | akzeptiere ich | befürworte ich | setze mich dafür ein |
| Akzeptanzstufe | Aktive Gegnerschaft | Ablehnung | Zwiespalt | Gleichgültigkeit | Duldung | Konditionale Akzeptanz | Zustimmung | Engagement |
| Akzeptanzstufenwert | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |

| | | | |
|--------------------------|-------------|------------------|-----------|
| Akzeptanzausprägung | Inakzeptanz | Gleichgültigkeit | Akzeptanz |
| Akzeptanzausprägungswert | 1 | 2 | 3 |

Mittels Telefon- und Onlinebefragung war es im Jahr 2019 möglich, die Akzeptanzstufe von 1.006 volljährigen Personen zu ermitteln, die nach einem mehrstufigen, geschichteten Zufallsprinzip ausgewählt wurden. Die Umfrageergebnisse sind für den Freistaat Sachsen repräsentativ.

Anfang des Jahres 2020 wurden in Sachsen ansässige oder tätige Unternehmen aus dem Energie- und Rohstoffsektor kontaktiert. Der Fragebogen wurde sowohl postalisch als auch als Onlineversion zur Verfügung gestellt. Die UnternehmensvertreterInnen wurden u. a. um ihre Einschätzung zur Akzeptanzstufe der Bürgerschaft gebeten sowie zu ihrem Umgang mit den StakeholderInnengruppen und der Gestaltung ihres Kommunikationsmanagements befragt.

4. Ergebnisse: Diskrepanz zwischen unternehmerisch empfundener und tatsächlicher (In-)Akzeptanz

Werden die Angaben der Bevölkerung als Akzeptanzausprägungen zusammengefasst, akzeptiert die sächsische Bevölkerung am ehesten eine Wasserkraft- und Solaranlage in ihrer Nähe, gefolgt von einer Geothermie-, Windkraft- und Biomasseanlage (Tabelle 2). Bei der Rohstoffgewinnung wird der Untertagebau etwas mehr als der Tagebau akzeptiert.

Tabelle 2: Projektbezogene Akzeptanzausprägung gegenüber Anlagearten (gewichtet); \bar{x} von eins für *Inakzeptanz*, zwei für *Gleichgültigkeit* bis drei für *Akzeptanz*

| Anlageart | Projektbezogene Akzeptanzausprägung | | | |
|-------------------|-------------------------------------|------------------|-----------|-----------|
| | Inakzeptanz | Gleichgültigkeit | Akzeptanz | \bar{x} |
| Wasserkraftanlage | 3,8% | 30,6% | 65,6% | 2,6 |
| Solaranlage | 7,1% | 30,2% | 62,7% | 2,6 |
| Geothermieanlage | 10,8% | 45,7% | 43,5% | 2,3 |
| WKA | 19,8% | 41,4% | 38,8% | 2,2 |
| Biomasseanlage | 18,0% | 46,4% | 35,6% | 2,2 |
| Untertagebau | 48,4% | 37,0% | 14,7% | 1,7 |
| Tagebau | 53,0% | 34,6% | 12,5% | 1,6 |

Werden die Angaben der Bevölkerung als Mittelwert der erfragten Akzeptanzstufen \bar{x} dargestellt und mit den Angaben der UnternehmensvertreterInnen zur vermuteten

ten Akzeptanzstufe der AnwohnerInnen verglichen, zeigen sich Abweichungen (Tabelle 3): Alle UnternehmerInnen schätzen die Akzeptanz der Bevölkerung besser ein, als sie ist. Dabei liegen die 31 VertreterInnen der EE-Branche mit ihrer Einschätzung näher an der Angabe der Bevölkerung als die 40 VertreterInnen aus dem Rohstoffsektor.

Tabelle 3: Projektbezogenes Akzeptanzniveau, von UnternehmensvertreterInnen vermutet und von der Bevölkerung angegeben (gewichtet); \bar{x} von eins für *Aktive Gegnerschaft*, zwei für *Ablehnung* usw. bis acht für *Engagement*

| Kategorie | Projektbezogenes Akzeptanzniveau | | | | |
|----------------|---|-----------|------------------------|-----------|----------|
| | von UnternehmensvertreterInnen vermutet | | Angabe der Bevölkerung | | Δ |
| | N | \bar{x} | N | \bar{x} | |
| EE | 31 | 5,4 | 1.006 | 5,0 | -0,4 |
| Rohstoffsektor | 40 | 4,5 | 1.006 | 3,1 | -1,4 |

Besteht Kontakt zu den AnwohnerInnen, ist er bei 56,7 % der Unternehmen aus dem EE-Sektor nach eigenen Angaben *selten*, bei 26,7 % *regelmäßig* und bei 16,7 % *häufig* (N = 30). Die meistgenannte Antwort bei den Rohstoffunternehmen lautet mit 52,6 % *regelmäßig*, 34,2 % geben *selten* und 13,2 % *häufig* an (N = 38). Neben der Häufigkeit wurde auch die Art des Kontakts, ob *zwangsweise*, *zufällig* oder *proaktiv* abgefragt. Die Angabe *zufällig* überwiegt bei den Unternehmen beider Kategorien (Tabelle 4).

Tabelle 4: Art des Kontakts zu AnwohnerInnen nach Kategorien; \bar{x} von eins für *zwangsweise*, zwei für *zufällig* bis drei für *proaktiv*

| Kategorie | Art des Kontakts zu AnwohnerInnen | | | | |
|----------------|-----------------------------------|-------------|----------|----------|-----------|
| | N | zwangsweise | zufällig | proaktiv | \bar{x} |
| EE | 29 | 6,9% | 48,3% | 44,8% | 2,4 |
| Rohstoffsektor | 37 | 18,9% | 43,2% | 37,8% | 2,2 |

Die AnwohnerInnen bzw. weitergefasst die Bevölkerung sind nur eine der relevanten StakeholderInnengruppen. Den UnternehmensvertreterInnen waren bei der schriftlichen Befragung *AnwohnerInnen* sowie *LokalpolitikerInnen vor Ort* als Anspruchsgruppen vorgegeben. Sie wurden gebeten, weitere zu nennen. Im Durchschnitt kam es zu einer zusätzlichen Nennung. In der Onlineumfrage wurden den Unternehmen neun Gruppen vorgegeben. Werden beide Umfragen zusammen ausgewertet, zeigt sich, dass drei Unternehmen zu keiner StakeholderInnengruppe Kontakt haben; alle drei entstammen dem Rohstoffsektor. Die Mehrzahl der Unternehmen interagiert aber mit *AnwohnerInnen*, *PolitikerInnen vor Ort* und *Behörden*.

Grundsätzlich sind bei Bau- und Infrastrukturprojekten die AkteurInnen häufig ähnlich [17], dennoch kann an dieser Stelle keine vollständige Liste der StakeholderInnen(-gruppen) erarbeitet werden. Die Art und Anzahl der Beteiligten unterscheidet sich je nach räumlichem und sachlichem Ausmaß des Projekts, nach Ausführungsfort-

schritt sowie Komplexität der Verfahren [18]. Eine allumfassende Liste wäre deshalb bei veränderten Rahmenbedingungen und speziell an einem anderen Ort obsolet. Als Anhaltspunkt wurde dennoch beispielhaft eine Übersicht von StakeholderInnengruppen zusammengestellt (Abbildung 1). Diese sind: *Bevölkerung, bürgerschaftliche Interessensgruppen, Exekutive, Legislative, Judikative, Verbände, unternehmensinterne StakeholderInnen, andere Unternehmen, Wissenschaft/Forschung sowie Medien*. Dabei lassen sich weitere Unterscheidungen bspw. nach der Nähe zum Unternehmen(-sprojekt) vornehmen, so gibt es in der Regel eine lokale, regionale, landes- und bundesweite Ebene. Auf der ersten, der lokalen Ebene sind unter *Bevölkerung* etwa die Anrainer- und OrtsanwohnerInnen zu nennen. Im regionalen Umfeld sind die EinwohnerInnen der Gemeinde, im landesweiten Umfeld die BürgerInnen des Bundeslands und auf Bundesebene die BürgerInnen Deutschlands aufzuführen.

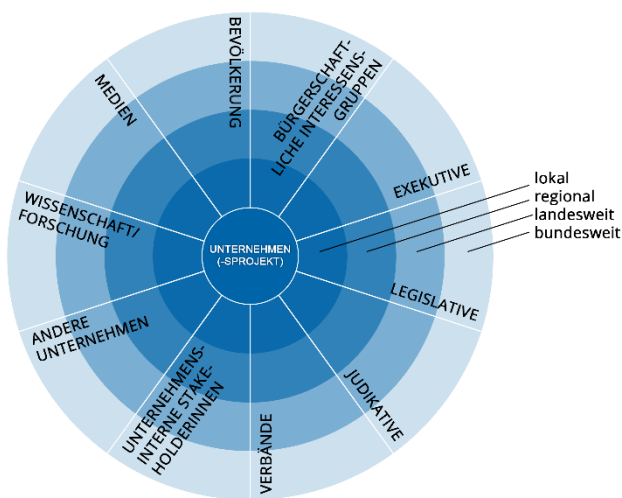


Abbildung 1: Übersicht über StakeholderInnengruppen

Von den EE-Unternehmen haben 7,4 % eine Presse-, Marketing- oder Öffentlichkeitsabteilung (N = 27); bei den Rohstoffunternehmen sind es 25,5 % (N = 47). 14,8 % der UnternehmerInnen aus dem EE- (N = 27) und 6,5 % aus dem Rohstoffsektor (N = 46) geben an, keine Öffentlichkeitsarbeit zu betreiben. Die anderen Unternehmen sind in der überwiegenden Anzahl mit einer Webpräsenz vertreten, Social-Media-Kanäle werden am seltensten genutzt (Tabelle 5)

Tabelle 5: Genutzte Formen der Öffentlichkeitsarbeit nach Kategorien (EE: N = 27, Rohstoffsektor: N = 46)

| Formen der Öffentlichkeitsarbeit | Kategorie | |
|----------------------------------|-----------|----------------|
| | EE | Rohstoffsektor |
| Website | 74,1% | 87,0% |
| Pressearbeit | 48,1% | 52,2% |
| Sponsoring | 33,3% | 67,4% |
| Veranstaltungen | 37,0% | 58,7% |
| Broschüre/-flyer | 48,1% | 39,1% |
| Gesprächsrunden | 33,3% | 34,8% |
| Werbeanzeigen | 33,3% | 30,4% |
| Social-Media-Kanäle | 18,5% | 21,7% |

Die höchste Bedeutung hat die Kommunikationsarbeit für die UnternehmerInnen des EE-Sektors in der Umsetzungsphase (N_{min} = 23, N_{max} = 26), für die RepräsentantInnen des Rohstoffsektors ist sie in der Betriebs- und Planungsphase am wichtigsten (N_{min} = 44, N_{max} = 46). Die Kommunikation in der Ideenphase wird von beiden Arten von UnternehmensvertreterInnen am geringsten bewertet. Als meistgenannte Antwort bzgl. der Probleme bei der Öffentlichkeitsarbeit geben 71,4 % der EE- (N = 21) und 39,5 % der RohstoffvertreterInnen (N = 43) zu wenig Personal bzw. zu geringes Budget an. Wiederrum 19 % der EE- und 34,9 % der RohstoffunternehmerInnen können von keinen Problemen berichten. Negative Resonanz in der Vergangenheit erlebten 4,8 % der EE- und 23,3 % der RohstoffvertreterInnen.

5. Zusammenfassung und Diskussion: Kommunikatives Potenzial liegt brach

Bezüglich der (In-)Akzeptanz zeigt sich, dass die BranchenvertreterInnen ein von der Realität abweichendes Gefühl bzgl. des Akzeptanzniveaus der sächsischen Bevölkerung haben. Der Vergleich macht deutlich, dass die Bevölkerung zukünftigen Projekten der Unternehmen kritischer gegenüberstehen wird, als die UnternehmerInnen dies bisher kennen oder vermuten. Die getroffenen Aussagen bleiben in ihrer Prognosewirkung allerdings begrenzt. Dies liegt u. a. daran, dass es schwierig ist, aus Befragungen auf die Wirklichkeit bei der Vorhabensumsetzung zu schließen. Die weitere Erforschung der „Zusammenhänge zwischen gegenwärtigen Einstellungen und späterem Verhalten“ [17] sowie der möglichen Diskrepanz zwischen beiden bleibt daher auch nach dieser Forschung eine Priorität.

Neben den AnwohnerInnen nannten die Befragten in der schriftlichen Umfrage in der Regel eine weitere StakeholderInnengruppe. Dies lässt vermuten, dass den UnternehmerInnen die Vielzahl der StakeholderInnen und damit auch deren Einfluss nicht bewusst sind. Ergänzt werden muss in diesem Zusammenhang, dass sich bereits die Interessen von AnwohnerInnen stark unterscheiden können: Während eine AnwohnerIn etwa von einer neuen Windkraftanlage persönlich profitiert, weil sie EigentümerIn des Grundstücks ist, sehen andere darin eine Gefahr, weil sie der Technik nicht vertrauen und Geräuschbelästigungen befürchten. Das Bewusstsein um solche Unterschiede, die auch in anderen Gruppen gegeben sind, erfordert neben einer detaillierten StakeholderInnenanalyse den regelmäßigen, proaktiven Austausch mit allen AkteurInnen. In der weiteren Forschung ist die Übersicht über die StakeholderInnengruppen zu spezifizieren, Key-StakeholderInnen herauszuarbeiten sowie die Gruppen z. B. anhand der Kriterien Macht und Interesse oder Bedrohungs- und Kooperationspotenzial zu bewerten.

Werden von den Unternehmen Maßnahmen der Öffentlichkeitsarbeit genutzt, sind diese überwiegend reaktiver Art und setzen auf das Pull-Prinzip: BürgerInnen müssen

etwa aktiv die Website der Unternehmen aufrufen, um Informationen zu erhalten. Damit werden nur diejenigen erreicht, die durch diesen Aufruf Interesse am Unternehmen(-sprojekt) ausdrücken. Die für die Akzeptanzgenerierung wichtige Gruppe der Unentschlossenen drückt ein solches Interesse jedoch nicht aus. Sie muss mit Maßnahmen aktiviert werden, die sie in ihrer Lebenswelt erreichen. Zudem fehlen Beteiligungsmaßnahmen; nur ca. ein Drittel der UnternehmerInnen aus beiden Kategorien nutzt Gesprächsrunden, wobei der Austausch und die Integration der Bevölkerung bei dieser Beteiligungsform lediglich auf der Stufe der Information stattfindet. Von kooperativen Verfahren wie BürgerInnenhaushalt und Runder Tisch oder Verfahren der Mitentscheidung wie bspw. BürgerInnengutachten ist keine Rede. Dabei gilt Beteiligung, neben Kommunikation, als zentraler Einfluss- und Erfolgsfaktor für die Akzeptanzgenerierung bzw. -steigerung. Gleichzeitig sind die Angaben zum Kommunikations- und Beteiligungsmanagement nachvollziehbar, wenn nur 2 von 27 UnternehmerInnen aus dem EE-Sektor und 12 von 47 UnternehmerInnen aus dem Rohstoffsektor angeben, dass sie über eine entsprechende Abteilung verfügen.

Dies könnte ebenso Auswirkungen auf die Beurteilung der Bedeutung von Kommunikationsarbeit haben: Ihr wird von UnternehmensvertreterInnen beider Gruppen in der Ideenphase nur geringe Wichtigkeit zugeschrieben. Dabei ermöglicht frühzeitige Kommunikation einerseits die öffentliche Meinung derart zu formen, dass unwahre Gerüchte entkräftet werden und keine übermäßige Beeinflussung anderer durch falsch informierte Personen stattfindet. Andererseits ermöglicht ein frühzeitiger Diskurs, inklusive entsprechender Beteiligungsformate, zeitnah Änderungen am Projekt vorzunehmen, die sich positiv auf die Akzeptanz positiv auswirken, wodurch die Gefahr juristischer Auseinandersetzungen, damit verbundener Verzögerungen der Realisierung und teurer Änderungsplanungen zum späteren Zeitpunkt verhindert werden. Dies beschleunigt den ökologischen Wandel.

Auch, wenn den UnternehmerInnen das Spannungsfeld und die Möglichkeiten der Akzeptanzsteigerung durch Kommunikation und Beteiligungsmanagement bewusst(er) gemacht werden, heißt das noch nicht, dass sie die entsprechenden personellen und finanziellen Mittel dafür aufbringen können; dies gilt insbesondere für kleine und mittlere Unternehmen. Dennoch bleibt festzuhalten, dass sich Unternehmen insgesamt „ihrer kommunikativen Wirkung bewusster werden“ [18] müssen.

Literaturverzeichnis

- [1] BP, Statistical Review of World Energy, Nr. 69 (2020), 3.
 [2] IEA, Electricity consumption – World 1990-2018, (o. J.).

- [3] G. Flach, H. Kautzleben, K. Steinitz, Sichere Versorgung der Menschheit mit Energie und Rohstoffen, Sitzungsberichte der Leibniz-Sozietät, Nr. 82 (2005), 167-168.
 [4] F.-W. Wellmer, J. Gutzmer, J. Kullik, B. Erlach, Die Energiewende braucht verlässliche Rahmenbedingungen für den Metallerzbergbau, GAIA Ökologische Perspektiven für Wissenschaft und Gesellschaft, Nr. 26/3 (2017), 233-234.
 [5] BMWi, Erneuerbare Energien, (o. J.).
 [6] AEE, Zustimmung für den Ausbau der Erneuerbaren Energien bleibt hoch, (2021).
 [7] D. Fuchs, S. Gözl, A. Graf, T. Gumbert, M. Klobasa, M. Ruddat, M. Sonnberger, Energiewende – Akzeptanz stärken – Forschungsprojekt KomMA-P, (2016), 6.
 [8] M. Althaus, Schnelle Energiewende – bedroht durch Wutbürger und Umweltverbände? Protest, Beteiligung und politisches Risikopotenzial für Großprojekte im Kraftwerk- und Netzausbau, Wissenschaftliche Beiträge TH Wildau, (2012), 103-104.
 [9] M. Uken, Der Totalausfall, ZEIT online, (2019).
 [10] acatech, Energiesysteme der Zukunft, (o. J.).
 [11] VDI, Mit der Zahl der Kumpel sinkt die Akzeptanz für die Steinkohle, (2001).
 [12] A. Sauer, F. Luz, M. Suda, U. Weiland, Steigerung der Akzeptanz von FFH-Gebieten, BfN-Skripte, Nr. 144 (2005), I-3 in Anlehnung an D. Lucke, Akzeptanz – Legitimität in der "Abstimmungsgesellschaft", (1995), 104.
 [13] A. Sauer, F. Luz, M. Suda, U. Weiland, Steigerung der Akzeptanz von FFH-Gebieten, BfN-Skripte, Nr. 144 (2005), I-3 in Anlehnung an G. Hofinger, Denken über Umwelt und Natur, Nr. 5/1 (2001), 21.
 [14] M. Schäfer, D. Keppler, Modelle der technikorientierten Akzeptanzforschung, Nr. 34 (2013), 25.
 [15] B. F. Schmid, B. Lyczek, Die Rolle der Kommunikation in der Wertschöpfung der Unternehmung, 2. Aufl. (2008), 101.
 [16] K. Spang, W. Clausen, Projektmanagement von Verkehrsinfrastrukturprojekten, (2016), 225.
 [17] M. Vonier, Nachhaltiges Kommunikationsmanagement zur Erlangung sozialer Akzeptanz von Bau- und Infrastrukturprojekten dargestellt anhand einer Fallstudie zweier Kraftwerksprojekte, (2013), 318.
 [18] E. Spiegel, Konsensfindung und Konfliktbewältigung durch Verhandlungslosungen in der Umweltpolitik, (2012), 8.
 [19] BSR Sustainability GmbH, Forschungszentrum Jülich GmbH, Fraunhofer ISI, Wuppertal Institut für Klima, Umwelt, Energie GmbH, Sozioökonomische Begleitforschung zur gesellschaftlichen Akzeptanz von Carbon Capture and Storage (CCS) auf nationaler und internationaler Ebene, (2008), 46.
 [20] D. Schermer, F. Niethammer, Vom Stakeholdermanagement zu CSR: Glaubwürdig kommunizieren, (2008), 97.

Biologische Systeme verstehen

FRET-assisted structural model of the GAAA RNA tetraloop receptor

Felix Erichson¹, Fabio D. Steffen², Richard Börner¹

¹Laserinstitut Hochschule Mittweida, University of Applied Sciences Mittweida,
09648 Mittweida, Germany

²Department of Chemistry, University of Zurich, 8057 Zurich, Switzerland

*Long-range tertiary interactions between RNA tetraloops and their receptors stabilize the folding of ribosomal RNA and support the maturation of the ribosome. Here, we use FRET-assisted structure prediction to develop a structural model of the GAAA tetraloop receptor (TLR) interaction and its dynamics. We build the docked TLR *de novo*, label the RNA *in silico* and compute FRET histograms based on MD simulations. The predicted mean FRET efficiency is remarkably consistent with single-molecule experiments of the docked tetraloop. This hybrid approach of experiment and simulation will promote the elucidation of dynamic RNA tertiary contacts and accelerate the discovery of novel RNA and RNA-protein interactions as potential future drug targets.*

1. Introduction

Increasingly, biomolecular structures are obtained from hybrid structural modeling [1]. In these integrative approaches, experimental restraints are collected to refine predicted models [2]. One such method is fluorescence resonance energy transfer (FRET), first described by Theodor Förster in the late 1940s [3]. As a molecular ruler, FRET is used as a distance constraint to filter a computationally generated structural ensemble. Performing the experiments on a single-molecule level circumvents ensemble averaging and dissects heterogeneities in structure and dynamics [4–6].

According to Förster, the energy transfer from an excited donor fluorophore to an acceptor is strongly distance dependent [3].

$$E(r) = \frac{R_0^6}{R_0^6 + r^6} \quad (1)$$

The Förster radius R_0 describes the distance where 50% of the energy is transferred from the donor to the acceptor molecule. Thus, given a FRET pair with known R_0 and a distribution of inter-dye distances, generated for example by a molecular simulation, a mean transfer efficiency $\langle E \rangle$ can be calculated and compared to experiments [7].

Sulfonated carbocyanines sCy3 and sCy5 are often used for FRET experiments of RNA [8]. The fluorophores consist of two heterocyclic moieties connected by a polymethine chain. The dyes are attached to the biomolecule by flexible linkers to allow free dye movement. Coupling to the nucleobases is commonly achieved by reacting an amine group with a succinimidyl ester derivative of the fluorophore [9].

In order to reproduce the rotational diffusion of the dyes *in silico*, molecular dynamics (MD) simulations can be performed. Here, the dyes are attached to a structural model and the dye movement and potential interactions with the biomolecule are monitored [10–13]. However, MD simulations are computationally expensive. An alternative is to compute the sterically accessible volume (AV)

of the dye [14, 15]. This geometrical approach approximates the free diffusion and rotation of the fluorophore when attached to a biomolecule. The dye is described by a spheroid with a linker of a given length and width. Subsequently, the possible steric locations of the dyes are calculated [15]. This model is extended by considering interactions of the dyes with the biomolecule. A contact volume (CV) is determined and included in the AV calculations [13]. The AV and accessible contact volume (ACV) models determine the dye distribution in a few seconds on regular desktop computers [14].

These FRET-assisted approaches can be used to guide structure predictions of proteins and RNA [7]. In the presence of monovalent ions, like Na^+ or K^+ , RNA folds by forming Watson-Crick (WC) base pairs (A-U, C-G) [16, 17]. Tertiary structure motifs like the GAAA tetraloop-receptor are promoted by divalent metal ions [5, 18–21]. These contacts are stabilized by non-WC interactions among those the A-minor motif, which is highly abundant in the large ribosomal subunit [22]. Tertiary elements connect RNA helices and are thus responsible for compacting the RNA fold. Several of these motifs were characterized by means of NMR and smFRET [20, 23].

Herein, we present a semi-automated approach to predict FRET values on the basis of *de novo* modeled RNA structures. For this purpose, we reconstructed the GAAA tetraloop receptor motif *de novo* based on the sequence from Downey et al. [20]. We predict the transfer efficiency of the FRET pair sCy3/5 by means of MD simulation and ACV calculations. We evaluate our approach using the published experimental smFRET histograms [24–26].

2. Methods

The secondary structure was predicted with RNAfold [27] (Fig. 1a) and an initial 3D *de novo* model was created by RNAComposer [28, 29] (Fig. 1b). Fluorophore ACV and mean transfer efficiencies $\langle E \rangle$ (Fig. 1d) were determined using FRETraj [13] in PyMOL [30]. The carbocyanine fluorophores for the ACV calculation were parametrized

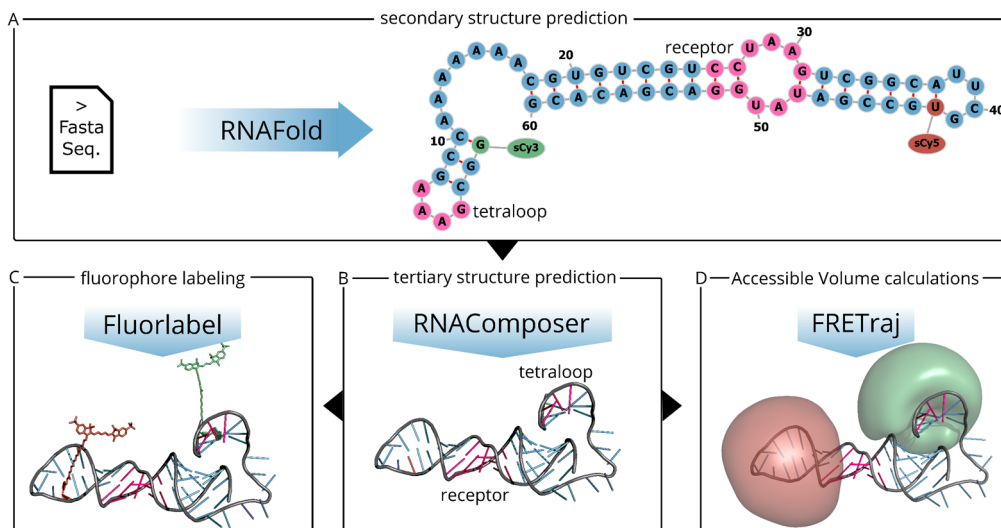


Figure 1 Pipeline for *de novo* model prediction and ACV calculation. (A) From the RNA sequence, a secondary structure is predicted which is then used to compose a three-dimensional *de novo* model (depicted in B). (C) Dyes are added with the tool Fluorlabel for a subsequent MD simulation. (D) Alternatively, accessible volumes are computed by FRETraj. Source code is available at github.com/RNA-FRETools.

according to [13, 15] and measurements in PyMOL (Tab. 1). For ACV calculation, the CV width was set to 3.0 Å and a CV occupancy of 0.25 was assumed. This corresponds to a moderate dye hinderance with a residual fluorescence anisotropy of 0.1 [13, 14].

All-atom MD simulation of the labeled RNA construct were performed using GROMACS 2020.4 [31]. For this purpose, the *de novo* models were labeled with fluorophores sCy3 and sCy5 using PyMOL and Fluorlabel [13] (Fig. 1c). The latest RNA-adjusted AMBER force field [32] with bsc0 [33] and χ OL3 [34, 35] correction was used. Molecules were solvated in a dodecahedral box filled with explicit TIP3P water. K^+ and Cl^- ions were randomly added to neutralize the net charge of the labeled RNA molecule. For the system representing the docked state, Mg^{2+} ions were added to a concentration of approximately 10 mM. In the equilibration phase (canonical and isothermal-isobaric ensembles, 3 ns each), the temperature was set to 298K with the velocity rescale thermostat and the pressure was set to 1 bar with the Parrinello-Rahman barostat. For production runs, bonds were constrained using the LINCS algorithm [36] and an integration time of 2 fs was used. Unbonded interactions were

Table 1: Overview of dye parameters for AV and ACV calculations derived from [13, 15, 30].

| Dye | R_1 (Å) | R_2 (Å) | R_3 (Å) | w_{link} (Å) | L_{link} (Å) |
|------|-----------|-----------|-----------|----------------|----------------|
| sCy3 | 8.0 | 3.0 | 1.5 | 4.5 | 20.3 |
| sCy5 | 10.0 | 3.0 | 1.5 | 4.5 | 23.0 |

processed with the Verlet scheme and a cut-off length of 1 nm. Long range electrostatics were processed with the particle mesh Ewald algorithm [37]. FRET burst distributions and transfer efficiencies were determined from MD trajectories with an *in silico* FRET algorithm called Fluorburst [13].

3. Results

The starting point for the modeling is the tetraloop receptor sequence taken from [20]. We additionally introduced a loop with the sequence AUUCGU to make the construct single-stranded (Fig. 1a). Further, we labeled the RNA with sCy3 and sCy5 (Fig. 1c). The latter was added at U42, which corresponds to the 5'-end label position of the original sequence. sCy3 was attached to the 5'-terminal G1. The secondary structure prediction shows that the GAAA tetraloop and the tetraloop receptor are formed (Fig. 1a). The subsequent RNAComposer 3D structure prediction (Fig. 1b) indicates a perfect overlap of the receptor domain (C18-G60) with a previously solved NMR structure (PDB-ID: 2i7z, overall RMSD = 2.12 Å) [23]. However, the tertiary contact between the GAAA tetraloop and the receptor is not formed and thus this

Table 2: Overview of calculated mean transfer efficiencies ($\langle E \rangle$) for AV, ACV and MD trajectories with $R_0 = 53$ Å and smFRET results from [24–26]. * initial structure from RNAComposer.

| Construct | $\langle E \rangle$ (MD) | $\langle E \rangle$ (AV) | $\langle E \rangle$ (ACV) | $\langle E \rangle$ (smFRET) |
|-----------|--------------------------|--------------------------|---------------------------|------------------------------|
| docked | 0.70 ± 0.08 | 0.67 | 0.72 | 0.68-0.70 |
| undocked | 0.03 ± 0.03 | 0.57 * | 0.61 * | 0.24-0.29 |

initial model represents one of many conformations of the undocked state. To model the bound state, we mapped the GAAA loop onto the NMR structure and remodeled the poly-A-linker (Fig. 2a). AV and ACV calculations were performed for the docked and the undocked state (Fig. 2b, bottom, Tab. 1). Predicted averaged transfer efficiencies ($\langle E \rangle$) are listed in Table 2.

An MD simulation of the docked state shows stable binding during the full runtime (1 μ s) with an overall RMSD of about 3.4 Å. Thus, no conformational dynamics are present on the timescale of the simulation. Conversely, in the undocked state rapid diffusion of the tetraloop away from the receptor domain is observed resulting in an

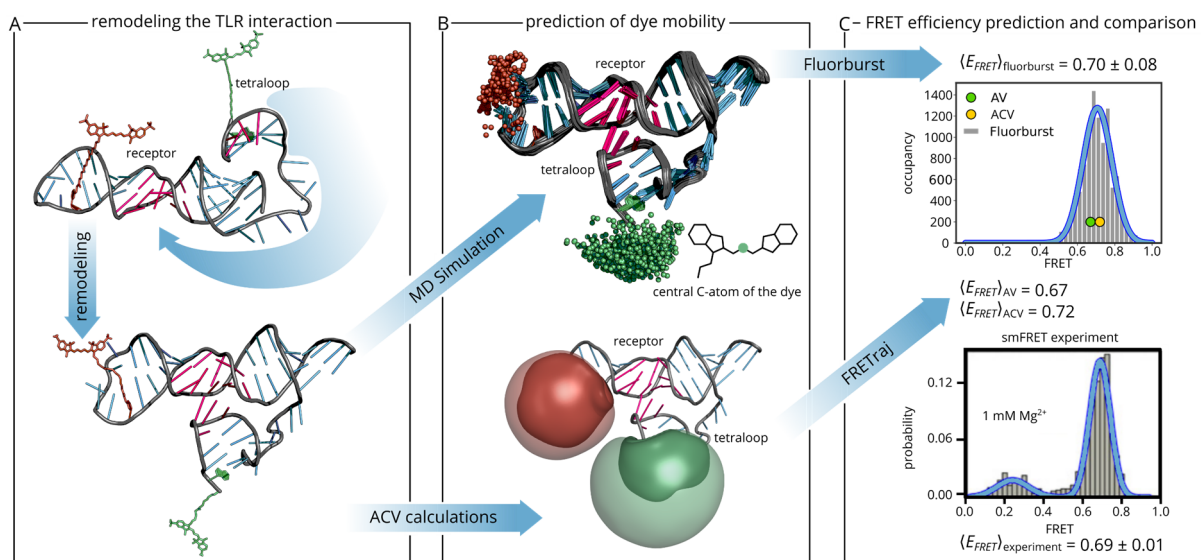


Figure 2: (A) The TLR structure was remodeled to the docked state by alignment with a published NMR structure (PDB: 2i7z). (B) MD simulation and ACV calculation were used to explore dye distributions. (C) The predicted mean FRET efficiencies (E) are compared with previous single-molecule experiments. FRET histogram adapted from [26].

RMSD of 9.1 Å for the entire structure and 11.3 Å for the tetraloop alone. Interestingly, the dye position of sCy3 distributes over a larger area than for sCy5 (Fig. 2b, top). The movement of the fluorophore sCy5 appears to be restricted as it interacts with U39 and A37 of the inserted loop. The mean transfer efficiency based on the MD trajectory of dye position and dipole orientations is $\langle E \rangle = 0.70 \pm 0.08$ for the docked state (Fig. 2c, Tab. 2) and 0.03 ± 0.03 for the undocked state (Tab. 2).

4. Discussion

The FRET values predicted for the docked state corresponds well to literature values of 0.68-0.70 [24–26]. The ACV slightly shifts the mean transfer efficiency to a higher value, thus bringing the dyes a bit closer together. Introduction of the additional loop exposes some interaction sites for the internal Cy5 dye. The remarkable agreement of simulation and experiment though suggests that the stacking tendency is similar to the blunt end present in the experiment. The mean transfer efficiency of the undocked state $\langle E \rangle = 0.03 \pm 0.02$ on the other hand does not align with literature values of 0.24-0.29 [24–26]. Neither MD nor AV and ACV calculations could yet reproduce the experimental results. The MD simulation of the undocked state sample conformations where the tetraloop diffuses freely without ever interacting with the receptor domain. The 0.2 state in the single molecule experiments likely arises from residual, unspecific interactions, which are not yet captured by the MD simulation.

5. Outlook

Here, we assumed a relative contact volume occupancy of 0.25. Taking the experimentally determined dynamic fluorescence anisotropy into account will enable us to accurately weight the ACV to refine the FRET predictions. To better describe the undocked state, multiple MD simulations of different start structures are necessary. Thus,

the movement of the RNA needs to be sampled more thoroughly to cover the continuum of undocked state conformations.

Acknowledgements

The authors thank the organizers of the Interdisziplinäre Wissenschaftliche Konferenz an der Hochschule Mittweida (IWKM) for giving them the opportunity to present their research. We thank Florian Heinke for insightful discussions. We also thank the University of Applied Sciences Mittweida and the Laserinstitut Hochschule Mittweida for financial support.

References

- [1] J. W. Parks et al., Single-molecule FRET-Rosetta reveals RNA structural rearrangements during human telomerase catalysis, *RNA*, 23 (2017), pp. 175–188.
- [2] H. M. Berman et al., Federating Structural Models and Data: Outcomes from A Workshop on Archiving Integrative Structures, *Structure*, 27 (2019), pp. 1745–1759.
- [3] T. Förster, Zwischenmolekulare Energiewanderung und Fluoreszenz, *Ann. Phys.*, 437 (1948), pp. 55–75.
- [4] B. Schuler, Single-molecule FRET of protein structure and dynamics - a primer, *J. Nanobiotechnology*, 11 Suppl 1 (2013), S2.
- [5] R. Börner et al., Metal ion induced heterogeneity in RNA folding studied by smFRET, *Coord. Chem. Rev.*, 327-328 (2016), pp. 123–142.
- [6] F. D. Steffen et al., Metal ions and sugar puckering balance single-molecule kinetic heterogeneity in RNA and DNA tertiary contacts, *Nat. Commun.*, 11 (2020), p. 104.
- [7] M. Dimura et al., Automated and optimally FRET-assisted structural modeling, *Nat. Commun.*, 11 (2020), p. 5394.

- [8] Q. Zheng et al., Ultra-stable organic fluorophores for single-molecule research, *Chem Soc Rev*, 43 (2014), pp. 1044–1056.
- [9] M. Levitus and S. Ranjit, Cyanine dyes in biophysical research: the photophysics of polymethine fluorescent dyes in biomolecular environments, *Q. Rev. Biophys.*, 44 (2011), pp. 123–151.
- [10] M. Hoefling et al., Structural heterogeneity and quantitative FRET efficiency distributions of polypyrrolines through a hybrid atomistic simulation and Monte Carlo approach, *PLoS ONE*, 6 (2011), e19791.
- [11] T. Graen, M. Hoefling, and H. Grubmüller, AMBER-DYES: Characterization of Charge Fluctuations and Force Field Parameterization of Fluorescent Dyes for Molecular Dynamics Simulations, *J. Chem. Theory Comput.*, 10 (2014), pp. 5505–5512.
- [12] R. B. Best et al., Quantitative interpretation of FRET experiments via molecular simulation: force field and validation, *Biophys. J.*, 108 (2015), pp. 2721–2731.
- [13] F. D. Steffen, R. K. O. Sigel, and R. Börner, An atomistic view on carbocyanine photophysics in the realm of RNA, *Physical chemistry chemical physics : PCCP*, 18 (2016), pp. 29045–29055.
- [14] M. Dimura et al., Quantitative FRET studies and integrative modeling unravel the structure and dynamics of biomolecular systems, *Curr. Opin. Struct. Biol.*, 40 (2016), pp. 163–185.
- [15] S. Kalinin et al., A toolkit and benchmark study for FRET-restrained high-precision structural modeling, *Nat. Methods.*, 9 (2012), pp. 1218–1225.
- [16] D. E. Draper, A guide to ions and RNA structure, *RNA*, 10 (2004), pp. 335–343.
- [17] J. Lipfert et al., Understanding nucleic acid-ion interactions, *Annu. Rev. Biochem.*, 83 (2014), pp. 813–841.
- [18] I. Tinoco and C. Bustamante, How RNA folds, *J. Mol. Biol.*, 293 (1999), pp. 271–281.
- [19] Zhichao Miao et al., RNA Structure: Advances and Assessment of 3D Structure Prediction, *Annual Review of Biophysics*, 46 (2017), pp. 483–503.
- [20] C. D. Downey et al., Metal ion dependence, thermodynamics, and kinetics for intramolecular docking of a GAAA tetraloop and receptor connected by a flexible linker, *Biochemistry*, 45 (2006), pp. 3664–3673.
- [21] S. A. Woodson, Compact intermediates in RNA folding, *Annual Review of Biophysics*, 39 (2010), pp. 61–77.
- [22] P. Nissen et al., RNA tertiary interactions in the large ribosomal subunit: the A-minor motif, *Proceedings of the National Academy of Sciences of the United States of America*, 98 (2001), pp. 4899–4903.
- [23] J. H. Davis et al., Role of metal ions in the tetraloop-receptor complex as analyzed by NMR, *RNA*, 13 (2007), pp. 76–86.
- [24] J. L. Fiore et al., Monovalent and divalent promoted GAAA tetraloop-receptor tertiary interactions from freely diffusing single-molecule studies, *Biophys. J.*, 95 (2008), pp. 3892–3905.
- [25] J. L. Fiore et al., Enthalpy-driven RNA folding: single-molecule thermodynamics of tetraloop-receptor tertiary interaction, *Biochemistry*, 48 (2009), pp. 2550–2558.
- [26] J. L. Fiore et al., The role of counterion valence and size in GAAA tetraloop-receptor docking/undocking kinetics, *J. Mol. Biol.*, 423 (2012), pp. 198–216.
- [27] R. Lorenz et al., ViennaRNA Package 2.0, *Algorithms for molecular biology : AMB*, 6 (2011), p. 26.
- [28] M. Popeno et al., Automated 3D structure composition for large RNAs, *Nucleic Acids Res.*, 40 (2012), e112.
- [29] M. Antczak et al., New functionality of RNAComposer: an application to shape the axis of miR160 precursor structure, *Acta Biochim. Pol.*, 63 (2016), pp. 737–744.
- [30] The PyMOL Molecular Graphics System: Schrödinger, LLC - Version 2.4.
- [31] E. Lindahl, GROMACS 2020.4 Source code: Zenodo.
- [32] W. D. Cornell et al., A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules, *J. Am. Chem. Soc.*, 117 (1995), pp. 5179–5197.
- [33] A. Pérez et al., Refinement of the AMBER force field for nucleic acids: improving the description of alpha/gamma conformers, *Biophys. J.*, 92 (2007), pp. 3817–3829.
- [34] P. Banáš et al., Performance of Molecular Mechanics Force Fields for RNA Simulations: Stability of UUCG and GNRA Hairpins, *J. Chem. Theory Comput.*, 6 (2010), pp. 3836–3849.
- [35] M. Zgarbová et al., Refinement of the Cornell et al. Nucleic Acids Force Field Based on Reference Quantum Chemical Calculations of Glycosidic Torsion Profiles, *J. Chem. Theory Comput.*, 7 (2011), pp. 2886–2902.
- [36] B. Hess et al., LINCS: A linear constraint solver for molecular simulations, *J. Comput. Chem.*, 18 (1997), pp. 1463–1472.
- [37] U. Essmann et al., A smooth particle mesh Ewald method, *J. Chem. Phys.*, 103 (1995), pp. 8577–8593.

Optimization of an optical setup for combined TIRF/confocal fluorescence microscopy

Hannes Michaelsen, Richard Börner

Laserinstitut Hochschule Mittweida, University of Applied Sciences Mittweida,
09648 Mittweida, Germany

This work deals with the construction of a microscope for combined total internal reflection fluorescence (TIRF) and confocal microscopy. It is especially designed for single-molecule fluorescence spectroscopy. The design of the microscope body is based on the miCube (Hohlbein lab, Wageningen University, NL). The excitation and detection pathways were adapted to allow both TIRF and confocal illumination as well as camera and point-detection for two color-channels to allow single-molecule Förster resonance transfer measurements.

1. Introduction

Fluorescence microscopes from common manufactures equipped with single-molecule detection capacities are very expensive (> 100 k€). Therefore, such microscopes have been designed as open source concept such as the miCube [1], the liteTIRF [2] and the smfBox [3] which achieve comparable good results for single-molecule FRET experiments [4]. Herein, we adapt the concept of the miCube for combined total internal reflection fluorescence (TIRF) and confocal microscopy with two color alternating laser excitation (ALEX) and two-color detection, i.e., detecting the donor and acceptor fluorescence, to enable single-molecule Förster resonance energy transfer (smFRET) measurements.

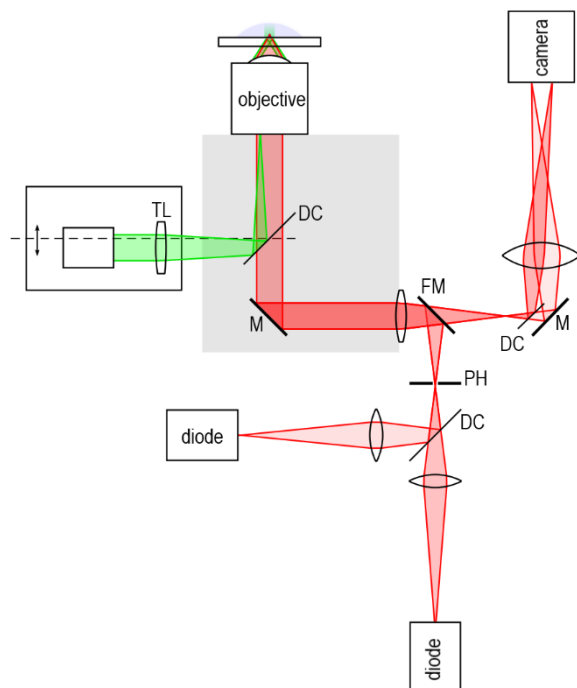


Figure 1. Excitation (green) and detection (red) beam path of the microscope. The excitation beam path includes the beam collimator (BC) with TIRF lens (TL) and y-travelling post (\leftrightarrow). To switch from TIRF to confocal mode remove TL and adjust the y-stage. The detection beam path includes a flipping mirror (FM) to switch between widefield camera-detection and confocal point-detection including a pinhole (PH) to reduce out-of-focus light. Both color-channels are separated via dichroic mirrors (DC).

Fluorescence microscopy is a widely used technique to study fluorescently labelled biomolecules. Due to the Stokes-shift only fluorescent light is detected while the excitation light is blocked from a dichroic mirror which separates the detection and excitation pathway.

For objective-based TIRF microscopy the parallel excitation beam is focused into the back focal plane of the high-numerical aperture ($NA > 1.4$) objective and is totally internal reflected at the interface between the sample (usually dissolved in water $n_{\text{water}} \approx 1,33$) and the cover slip (e.g. BK7 glass $n_{\text{glass}} = 1,51$). The emerging evanescent field excites the fluorophores only close to the surface due to its low penetration depth (< 100 nm). Therefore, this method is perfectly suited to suppress background fluorescence and to reach a high signal-to-background ratio. The fluorescent light is finally detected with an EMCCD or sCMOS camera.

In confocal microscopy the parallel excitation beam is focused with a high NA objective into the sample. The subsequently emitted fluorescent light is collected with the same objective. Thus, excitation and detection pathway overlay each other according to the confocal principle and build a "cuvette" with a volume in the fL range. Usually, out-of-focus light is blocked with a pinhole in the detection pathway to reduce the background signal, again, to reach a high signal-to-background ratio.

To switch the excitation pathway from TIRF to confocal and *vice versa*, the TIRF lens needs to be removed and the collimator stage is adjusted to reach a perfect symmetric filling of the back focal plan of the objective. The associated detection pathway can be chosen with a flipping or magnetic mirror post (compare Fig. 1).

2. Microscope design

The microscope setup is composed out of commercially available opto(mechanics) complemented with a home-build microscope body, sample holder and collimator stage. A complete list of all commercial parts can be found elsewhere [1]. The self-designed elements are partially based on miCube templates (cube and collimator). All newly designed elements are described herein.

The setup is split into four groups: the excitation beam path, the microscope body, the sample holder, and the detection beam path. The excitation path includes two 150 mW diode lasers (Oxxius) with 515 nm and 638 nm wavelength, respectively, both coupled into one optical single-mode fiber (Thorlabs). The reflective fiber collimator (Thorlabs) is set on a travelling stage and equipped with an iris aperture to allow under- and overfilling of the objective and a removeable (TIRF) lens to switch between TIRF and confocal excitation. The microscope body/cube design ensures that objective, fiber collimator and imaging tube lens are in correct distance to each other (Fig. 2). The central part of the cube is the dichroic mirror (515/638 DC, AHF analytics) which separates the excitation from the detection beam path.

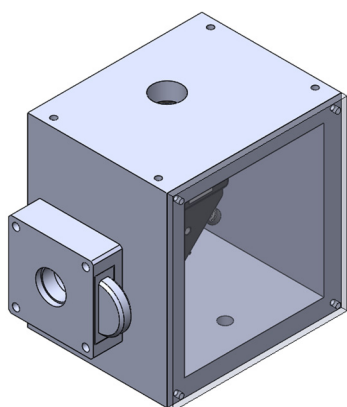


Figure 2: CAD model of the microscope body with the tube lens holder.

To keep costs low the sample holder is built out of commercially available elements with manual x-y-axis travelling stage and a motorized z-axis traveling stage (Fig. 3). The latter is used to easily adjust the focus position. The round insert of the sample holder is made of magnetic steel. This not only allows fast switching of samples but to easily keep the cover slip in place with magnets. A component overview is listed in Tab. 1.

Table 1. Component overview for the sample holder.

| title | quantity | Price (excl. VAT) |
|-------------------------|----------|-------------------|
| x-y-linear axis | 1 | € 793 |
| rail connector | 1 | € 35 |
| 66 mm rail | 1 | € 44 |
| rail slide | 1 | € 67 |
| motorized linear z-axis | 1 | € 810 |
| controller | 1 | € 616 |
| sample holder | 1 | home-build |
| total | | € 2365 |

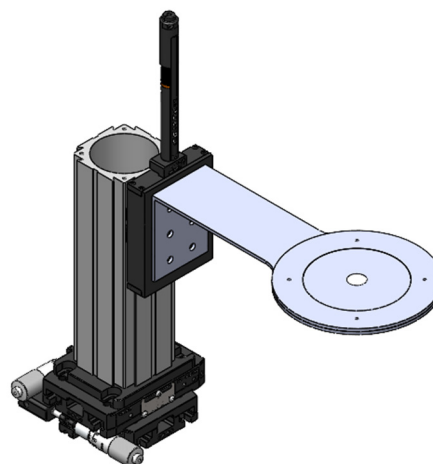


Figure 3: CAD-model of the sample holder with motorized z-stage, and manual x-y-stage.

The detection path consists out of two distinct beam paths for each detection mode. A flipping mirror or a magnetic mirror post allows easy and fast switching.

For TIRF microscopy the image is recorded via the prime BSI express camera from photometrics. The camera is equipped with a scientific CMOS sensor with a very high sensitivity and very low noise characteristics to allow high signal-to-noise ratio (SNR) mandatory for single molecule fluorescence detection. The field of view (FOV) is split into two-color channels via a DC mirror and imaged on the camera.

For confocal microscopy, a telecentric lens system allows easy adjustment of the pinhole. A DC mirror splits the beam into two-color channels. The fluorescence signal is measured with two avalanche photo diodes (APD, Picoquant).

Most of the parts except of the cube and the laser baseplate are printed out of ABS with a custom 3D printer. The cube and the laser baseplate were built out of aluminum due to the numerous tapped holes.

3. Conclusion

We have realized a combined TIRF- and confocal microscope for less than 50 k€. The most expensive parts are the CFI P-Apo 100x 1.45 oil immersion objective (Nikon), the laser (Oxxius), the APDs (Picoquant) and the sCMOS camera (Photometrics). All optical and optomechanical components cover only 20% of the budget. A good way to save money is the additive manufacturing of the cube etc. as the material costs are low and the CAD models are freely available.

The camera and the high NA objective are especially necessary for single-molecule microscopy. Additional three-axis piezo scanner for sample positioning and sample scanning will allow for confocal imaging capacities.

4. Acknowledgement

We thank the Laserinstitut Hochschule Mittweida, the University of Applied Sciences Mittweida and the Werner-Legat-Fund of the University of Zurich for financial support. We thank Alexander Horn supporting us with optomechanical components, Martin Eler for the additive manufacturing of optomechanical components and Klaus Grundmann for the conventional milling and construction of the microscope body.

5. Literature

- [1] Martens, K. J. A. et al. Visualisation of dCas9 target search in vivo using an open-microscopy framework. *Nature communications* 10, 3552; 10.1038/s41467-019-11514-0 (2019).
- [2] Auer, A. et al. Nanometer-scale Multiplexed Super-Resolution Imaging with an Economic 3D-DNA-PAINT Microscope. *Chemphyschem : a European journal of chemical physics and physical chemistry* 19, 3024–3034; 10.1002/cphc.201800630. (2018).
- [3] T. Craggs, B. Ambrose, J. Cully, J. Baxter, J. Shewring, B. Bateman, M. Aaldering und E. Steele, „smfBox: Open Source Single-Molecule FRET,“ Available online (Last access March 08, 2021): <https://craggs-lab.github.io/smfBox/index.html>.
- [4] Hellenkamp, B. et al. Precision and accuracy of single-molecule FRET measurements-a multi-laboratory benchmark study. *Nature methods* 15, 669–676; 10.1038/s41592-018-0085-0 (2018).

AnthroWorks3D: Digitalisierung von Skelettfundstücken und die virtuell-osteologische Untersuchung

Hanjo Tim Fritzscht¹, Fabian Schmiedel¹, Marie Heuschkel¹, Marleen Mohaupt¹, Dirk Labudde¹, Toralf Kirsten¹, Birgit Grosskopf², Anna Lena Flux², Stefanie Meier³, Patricia van der Burgt³

¹Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

²Georg-August-Universität Göttingen, Wilhelmsplatz 1, 37073 Göttingen

³Landesamt für Archäologie Sachsen, Zur Wetterwarte 7, 01109 Dresden

In der Anthropologie werden körperliche Überreste von Menschen historischer und rezenter Zeiträume mit dem Ziel der Aufdeckung ihrer Lebensumstände untersucht. Das Projekt AnthroWorks3D der Hochschule Mittweida verbindet Methoden der Videospielentwicklung und Osteologie, um das Knochenmaterial schonend und hochauflösend durch eine 3D-Scanpipeline zu digitalisieren und virtuell, ortsunabhängig und parallel zu vermessen mit dem Ziel, den Verschleiß am Knochenmaterial zu minimieren und die Fundstücke möglichst vielen Forschenden zugänglich zu machen. Die virtuelle Vermessung wurde in einem ersten Test in Zusammenarbeit der Abteilung für historische Anthropologie und Humanökologie des Johann-Friedrich-Blumenbach-Instituts für Zoologie und Anthropologie der Universität Göttingen evaluiert. Dabei schätzen über die Hälfte der befragten Anthropologen den Prototypen als Alternative zur physischen Untersuchung ein, besonders in Anwendungsfeldern, in denen das Knochenmaterial nur schwer zugänglich ist.

1. Einleitung

1.1. Die osteologische Untersuchung

Osteologie beschreibt die Untersuchung und Vermessung von Skelettfunden. In der Forensik ist die Osteologie vor allem in der Fallarbeit von Bedeutung, um beispielsweise bei Skelettfunden zwischen menschlichen und tierischen Knochen zu unterscheiden. Außerdem können Verletzungen und Veränderungen an den Knochen Hinweise auf die Todesursache oder Krankheiten des Individuums bieten. [1]

In der historischen und prähistorischen Anthropologie ist das Hauptziel der osteologischen Untersuchung das Ermitteln der individuellen Lebensdaten und -umstände eines Individuums. Dazu zählen zum Beispiel das Sterbealter, das Geschlecht und der Gesundheitszustand sowie genetische und umweltbedingte Besonderheiten des Individuums. So ist es möglich, durch die Untersuchung einer Vielzahl von Einzelschicksalen der Individuen, Aussagen über die Lebensbedingungen und Lebensweise der jeweiligen historischen Population zu treffen. [2]

1.2. Die osteologische Untersuchung im virtuellen 3D-Raum

Bei der physischen osteologischen Untersuchung entstehen für die Untersuchenden aber mehrere, mit herkömmlichen Mitteln nur schwer vermeidbare logistische Probleme. Beispielsweise können die Skelettfunde nur ortsgebunden untersucht werden und müssen teilweise über weite Strecken transportiert werden. Während dieses Transportes, der Untersuchung selbst und der damit

einhergehenden Vorbereitungen werden die Skelettfunde unweigerlich verschliffen. Dies ist besonders bei alten und porösen Funden problematisch und kann zu Abweichungen der Befunde bei der Untersuchung führen. [3]

Im Projekt AnthroWorks3D nehmen wir uns dieser Probleme an. Dabei wird ein digitaler Zwilling des Skelettfunds erstellt, um diesen dann in der virtuellen Untersuchungsumgebung in 3D darzustellen. Durch diesen digitalen Zwilling ist es nun möglich, die Funde ortsungebunden und parallel zu untersuchen. So bleibt der Fund auch nach der Beisetzung des Originals für weitere Untersuchung zugänglich. Der Lösungsansatz des Projektes nutzt die Methoden des Game Development im Bereich der 3D-Grafik, um die durch eine photogrammetrische 3D-Scanpipeline digitalisierten Fundstücke mittels 3D-Echtzeitrendering bereitzustellen und so die Infrastruktur für die virtuelle osteologische Untersuchung zu schaffen.

1.3. Übertragbarkeit auf andere Wirtschafts- und Forschungsbereiche

Der Fokus bei der Entwicklung von AnthroWorks3D liegt auf der Digitalisierung und Darstellung von menschlichen Skelettmaterial. Der Ansatz ist allerdings erweiterbar auf vergleichbare Anwendungsfälle. Da die entwickelte 3D-Rekonstruktionspipeline keinen 3D-Scanner, sondern lediglich handelsübliche Spiegelreflexkameras und Open-Source-Software für den 3D-Rekonstruktionsvorgang verwendet, ist sie gut übertragbar. Limitationen der 3D-Rekonstruktionsmethode stellen dabei transpa-

rente, homogene und stark reflektierende Oberflächen dar.

2. Die Digitalisierung des Knochenmaterials

2.1. Aufnahme des Bildmaterials

Der digitale Zwilling wird aus einer Reihe von Fotos vom gescannten Objekt erzeugt. Dazu wird der Knochen von drei Kameras aus drei verschiedenen Höhen aufgenommen. Um auch die Unterseite rekonstruieren zu können, wird der Knochen auf eine Halterung aus Acrylglas gelegt. Die gesamte Halterung steht dabei auf einem Drehteller, um den Knochen in einer 360°-Aufnahme fotografieren zu können. Dabei ist auch das Aufnehmen mehrerer und fragmentierter Knochen in einem Scandurchgang möglich.



Abbildung 1: Scansetup

Insgesamt werden bei diesem Vorgang 159 Bilder im JPEG-Format aufgenommen. Dafür werden im Schnitt sechs Minuten pro Aufnahmedurchgang benötigt. Da die Knochen lediglich auf einen Glastisch gelegt werden müssen, ist der Scansvorgang für das Material sehr schonend, wodurch das Original unbeschadet an die Sammlung zurückgegeben werden kann.

2.2. Rekonstruktion in Meshroom

Zur 3D-Rekonstruktion des Modells aus den Bildern wird die Open-Source Software „Meshroom“ [4] verwendet. Für die Rekonstruktionseinstellung wurde ein Template entwickelt, welches sich für die Digitalisierung von Knochenmaterial eignet. So ist es dem Nutzer der Pipeline nun möglich die Bilder aus dem vorhergegangenen Schritt per Drag-and-Drop in die Software einzufügen und die Rekonstruktion zu starten. Im Rekonstruktionsvorgang werden Features auf dem Knochenmaterial erkannt, um die einzelnen Bilder zu matchen und daraus das 3D-Modell zu generieren.

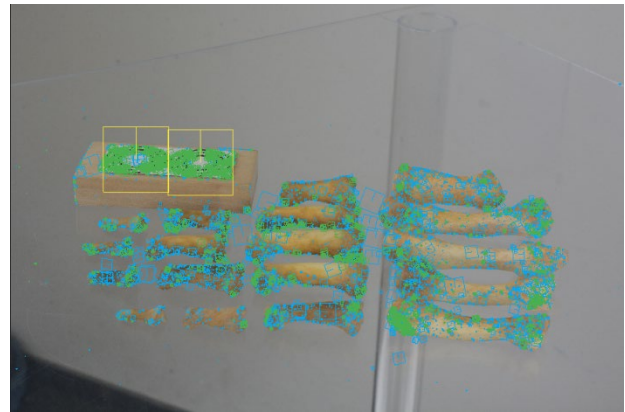


Abbildung 2: Hand und Referenzobjekt mit erkannten Sift (blau), Akaze (grün) und CCTag3 (gelb) Features

Die Berechnung des 3D-Modells benötigt im jetzigen Setup mit einem Intel Core i7-8750H CPU und einem NVIDIA GeForce GTX 1070 Grafik-chip im Schnitt sieben Stunden.

2.3. Retopology in Meshlab

Da die im letzten Schritt entstandenen Modelle aus sehr vielen Polygonen bestehen und zur Untersuchung in Echtzeit dargestellt werden sollen, muss die Polygonanzahl der Modelle verringert werden. Nur so kann eine flüssige Darstellung gewährleistet werden. Dazu wird die Open Source Software „Meshlab“ [5] verwendet. Meshlab bietet eine einfache Möglichkeit die Polygonanzahl auf einen gewählten Maximalwert zu reduzieren. Im selben Arbeitsschritt werden eventuelle Scanfragmente entfernt. Das Ergebnis der Rekonstruktionspipeline ist ein detailliertes 3D-Modell des Knochens bestehend aus einem Mesh im OBJ-Format und einer Textur im PNG-Format. Die Durchführung der einzelnen Arbeitsschritte der Pipeline ist bereits zum Großteil automatisiert.

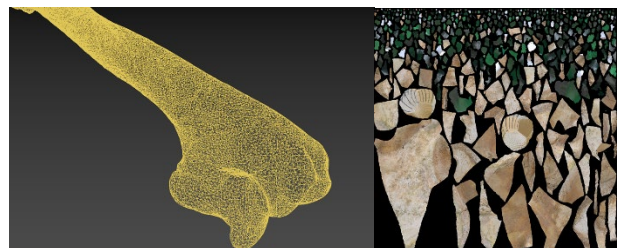


Abbildung 3: Erstelltes Mesh und Textur eines Humerus



Abbildung 4: Digitaler Zwilling des Humerus

3. Virtuelle osteologische Untersuchung des digitalen Zwillings

3.1. Virtuelle Untersuchungsumgebung

Um eine virtuelle Untersuchung der aus der Scanpipeline entstandenen Modelle zu ermöglichen, wurde ein Prototyp mit der Game Engine „Unity“ [6] entwickelt, da durch diese bereits eine Echtzeitrenderpipeline zur Darstellung der 3D-Modelle bereitgestellt wird. Der Prototyp ermöglicht die Darstellung der Modelle in Echtzeit und die Interaktion mit diesen.

3.2. Platzierung der digitalen Knochen im Skelett

Der aktuelle Prototyp zur Untersuchung erlaubt es diese digitalen Zwillinge als 3D-Objekt zu importieren. Beim Import ist es möglich Informationen zum jeweiligen importierten Knochen bzw. Knochenfragment festzuhalten. Die Bewegungswerkzeuge erlauben es dann den Knochen in die richtige Position im Skelett einzupassen.

Dabei erhält der Nutzer eine Orientierung durch einen Platzhalter, in welchen er das jeweilige Knochenfragment einpassen kann. Der Auslegevorgang ist weitaus effizienter als im Physischen, da die Knochenmodelle nicht der Schwerkraft unterliegen und die aktuelle Position in Relation zu den anderen Knochen abgespeichert wird.

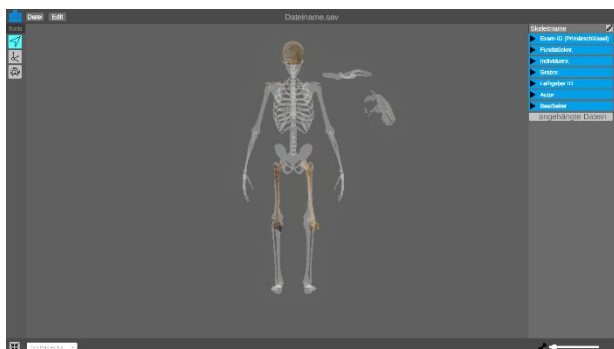


Abbildung 5: Angeordnete Knochen im Platzhalter-Skelett

3.3. Messungen am digitalen Zwilling

Die Messwerkzeuge des Prototyps erlauben es Distanzmessungen, Umfangsmessungen und Winkelmessungen an den Knochenmodellen durchzuführen. Dabei ergaben erste Experimente, in welchen Messungen an den physischen Knochen mit Messungen am digitalen Zwilling verglichen wurden, dass die bisher implementierten Messwerkzeuge eine vergleichbare Genauigkeit wie die physischen Messwerkzeuge liefern. Dem Nutzer wird es ermöglicht Messungen mit Bezeichnungen zu versehen, um diese schnell wiederzufinden. Die Messungen werden kategorisiert in Listen je nach Art der Messung zusammengefasst und können direkt auf der Oberfläche des Knochens dargestellt werden. Weiterhin ist es möglich die Messungen auf dem Knochen ein- und auszublenden, sowie die Kameraansicht auf diese zu fokussieren. So kann direkt nachvollzogen werden, welche

Messungen von einem vorhergegangenen Nutzer getätigt wurden.

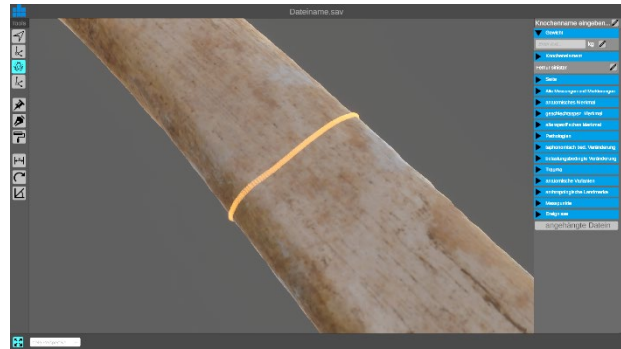


Abbildung 6: Beispielhafte Umfangsmessung

3.4. Markierungen

Um Besonderheiten auf dem Knochen zu dokumentieren, stehen dem Nutzer eine Reihe von Markierungswerkzeugen zur Verfügung. Diese können wie Messungen kategorisiert sowie ein und ausgeblendet werden. Zudem ist eine automatische Fokussierung der Ansicht auf die jeweilige Markierung möglich. Weiterhin können zusätzliche Informationen und Metadaten zu den einzelnen Markierungen und Messungen abgespeichert werden.

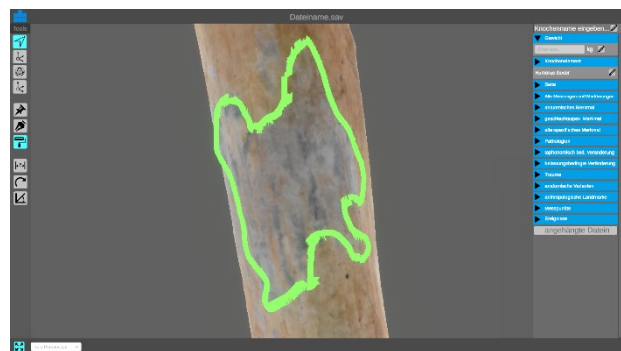


Abbildung 7: Beispiel einer Markierung auf dem Knochen

3.5. Anhängen von Dateien

Zur Einbeziehung der Ergebnisse und Dokumentation aus der physischen Untersuchung ist es möglich zu jedem importierten Knochen, sowie zum gesamten Skelett Dateien anzuhängen. Die angehängten Dateien haben keine Beschränkung nach Dateityp oder Datengröße und können mit einem Knopfdruck in den Eigenschaften des jeweiligen Knochens bzw. Knochenfragments aufgerufen werden. So können beispielsweise PDF-Dateien, Fotos von den realen Knochen oder andere Formate einfach auffindbar archiviert werden.

4. Ergebnisse der Evaluation des ersten Prototyps

Die Evaluation des ersten Prototyps wurde im Zeitraum vom 25.08.2020 bis zum 05.10.2020 in vier Experimentdurchgängen mit insgesamt sieben Anthropologen und Anthropologie-Studenten in Zusammenarbeit mit der Abteilung für historische Anthropologie und Humanöko-

logie des Johann-Friedrich-Blumenbach-Instituts für Zoologie und Anthropologie der Universität Göttingen durchgeführt. Dabei wurde der erste Prototyp auf User Experience, Richtigkeit der Messungen und Programmfehler getestet und im Nachhinein wurden zusätzlich Umfragen und Interviews mit den Testern der Anwendung durchgeführt. Es stellte sich heraus, dass die Anthropologen sich eine bedingte Nutzung im Arbeitsalltag vorstellen könnten. Aufgrund von einigen Nachteilen gegenüber der physischen Untersuchung wie beispielsweise der fehlenden Haptik der Knochen, kann die Anwendung zum aktuellen Zeitpunkt die physische Untersuchung nicht ersetzen und vorerst nur ergänzen. Dennoch gaben etwas mehr als die Hälfte der Befragten an, dass der Prototyp im Vergleich zur physischen Untersuchung eine gute Alternative mit Verbesserungspotential ist. Im Vergleich der physischen Messungen mit den Messungen im Prototyp wurde überprüft, wie zuverlässig die einzelnen Messwerkzeuge im Prototyp im Vergleich zu ihrem physischen Äquivalenten abschneiden. Daher wurde die Standardabweichung der Messungen der gleichen Messstrecken im Physischen und am Scan miteinander verglichen. Dabei hat sich herausgestellt, dass Messungen mit dem Distanztool im Prototyp zuverlässigere Werte wie der physische Tastzirkel liefern. Außerdem schneidet das Distanztool zuverlässiger bei Durchmesser-, Breiten- und Höhenmessungen als der Gleitzirkel ab. Das Winkeltool im Prototyp misst zuverlässiger als das Mandibulometer im Physischen und Umfangsmessungen im Prototyp haben eine ähnliche Zuverlässigkeit wie Messungen mit einem Maßband im Physischen. Zur Quantifizierung der User Experience wurden die Anthropologen in einer „AttrakDiff“-Analyse [7] mit einer Kontrollgruppe verglichen, die bereits über ein Jahr Erfahrung mit etablierten 3D-Anwendungen hatten. Dabei stellte sich heraus, dass sich ein positiver Trend User Experience einstellte. Im Vergleich zu etablierter 3D-Software gibt es allerdings noch Verbesserungspotential in verschiedenen Bereichen. Besonders die Navigation durch den 3D-Raum wird von den Anthropologen als schwieriger wahrgenommen als von der 3D-Testgruppe.

5. Praktische Relevanz des Prototyps

Die virtuelle Untersuchung kann die physische Untersuchung an realen Knochen zum aktuellen Zeitpunkt nicht ersetzen, allerdings hat sie hohes Potential die etablierten Untersuchungsmethoden zu unterstützen. Besonders bei der Untersuchung von im Physischen schwer zugänglichen Skeletten bietet die virtuelle Untersuchung einen großen Mehrwert. So können einfach Messungen durchgeführt werden, um einen ersten Eindruck von einem Individuum zu erhalten, ohne die physischen Knochen an den Untersuchungsort transportieren zu müssen. Durch die virtuelle Herangehensweise können Vergleichspopulationen und besondere Individuen auf der ganzen Welt parallel und ortsungebunden untersucht werden. Der digitale Zwilling gewährleistet außerdem

eine dauerhafte Verfügbarkeit der Knochen, ohne die physischen Originale einbehalten zu müssen. Besondere Funde können so archiviert werden und möglichst vielen Forschenden zugänglich gemacht werden. Dies ist besonders für die Lehre interessant. Studenten, die keinen Zugang zu einer Sammlung haben, wird es so ermöglicht Untersuchungen an Skeletten zu trainieren, was besonders in Zeiten der Pandemie an Relevanz gewonnen hat. Es wird ihnen so auch ermöglicht an seltenen prähistorischen Funden Erfahrungen zu sammeln - ohne die Gefahr die Funde zu beschädigen - und ihre Ergebnisse mit den Lehrenden digital auszutauschen.

6. Zukunft der Entwicklung

Der Ansatz unseres Projektes sieht vor, die physische Untersuchung zu digitalisieren und in den virtuellen Raum zu bringen, um die virtuelle Untersuchung dann schrittweise zu automatisieren und mit den durch die Digitalisierung eröffneten Möglichkeiten zu erweitern. Daher sehen die nächsten Entwicklungsschritte das Anbinden einer Datenbank vor, wodurch es ermöglicht werden soll, mehrere Skelette auch über Sammlungen hinweg miteinander zu vergleichen. Dabei soll eine Benutzer- und Versionsverwaltung implementiert werden, um die Untersuchungshistorie sichtbar zu machen und Zugriffs- und Bearbeitungsrechte zu verteilen. Weiterhin soll es möglich sein sich einen Eindruck der Knochen in einer Web Preview zu machen. Zudem bietet das beim Scanvorgang entstandene Bildmaterial einen Ansatzpunkt für eine automatische Annotation des Skelettmaterials durch maschinelles Lernen. Weiterführend soll auch das Innere der Knochen digitalisiert werden, indem CT-Scans und Photogrammetrische Scans zusammenfinden.

Literaturverzeichnis

- [1] Verhoff, M. A., Kreutz, K., Jopp, E., & Kettner, M. (2013). Forensische Anthropologie im 21. Jahrhundert. *Rechtsmedizin*, 23(2), 79-84.
- [2] Grupe, G., Harbeck, M., & McGlynn, G. (2015), *Prähistorische Anthropologie*, Springer Spektrum, 249
- [3] Grupe, G., Harbeck, M., & McGlynn, G. (2015), *Prähistorische Anthropologie*, Springer Spektrum, 217-221
- [4] AliceVision (2020). Meshroom (2020.1.0 64-bit). [Software]. AliceVision Association. URL: <https://alicevision.org> (Zugriff am 02. März 2021).
- [5] ISTI - CNR (2020). MeshLab (2020.06). [Software]. Institute for Information Science and Technologies "Alessandro Faedo". URL: <https://www.meshlab.net> (Zugriff am 02. März 2021).
- [6] Unity Technologies (2020). Unity (2019.3.6f1). [Software]. Unity Technologies. URL: <https://unity.com/> (Zugriff am 03. März 2021).
- [7] User Interface Design GmbH (2020). AttrakDiff. [Software]. User Interface Design GmbH. URL: <http://attrakdiff.de/> (Zugriff am 02. März 2021)

Engineering cyanobacteria for direct solar chemical and fuel production from CO₂

Peter Lindblad

Uppsala University, Department of Chemistry-Ångström, SE-751 20 Uppsala, Sweden

*Cyanobacteria, prokaryotic microorganisms with basically the same oxygenic photosynthesis as higher plants, are becoming excellent green cell factories for sustainable generation of renewable chemicals and fuels from solar energy and carbon dioxide. In the presentation I will visualize the concept green cell factories by introducing and discussing two examples: (i) engineering cyanobacteria to produce the important bulk chemical and potential blend-in biofuel butanol from sunlight and carbon dioxide, so called photosynthetic butanol, and (ii) generation of a functional semisynthetic [FeFe]-hydrogenase linking to the native metabolism in living cells of the unicellular cyanobacterium *Synechocystis* PCC 6803.*

1. Introduction

Cyanobacteria, prokaryotic microorganisms with oxygenic photosynthesis, evolved on earth billions of years ago. They convert solar energy, carbon dioxide and water into chemical energy while releasing O₂ into the atmosphere. In addition, some strains are able to convert N₂ into ammonia. Compared to other oxygenic photosynthetic organisms, cyanobacteria possess the highest solar energy capturing efficiency with corresponding adequate carbon dioxide concentrating mechanisms and carbon dioxide fixation. With their modest nutrient requirements and ability to grow on non-arable land and thereby not compete with e.g. food or fodder production, cyanobacteria have emerged as potential green cell factories for sustainable generation of truly carbon neutral, or even carbon negative, renewable chemicals and fuels. The rapid progress in synthetic biology and genetic engineering have made it possible to custom design microorganisms including cyanobacteria. This may include deletion of native and addition of new capacities as well as modified and optimized metabolic flow towards desired product(s). Until today, cyanobacteria have been engineered, as proof-of-concept, to synthesize numerous non-native products.

2. Engineering cyanobacteria to produce photosynthetic butanol

Cyanobacteria do not produce butanol naturally. However, a combination of synthetic biology and system biology strategies have been employed to enable photosynthetic butanol production, i.e. butanol produced from solar energy and carbon dioxide, in a limited number of model cyanobacterial strains.

Butanol is a four-carbon alcohol (C₄H₉OH) occurring in four structural isoforms: 1-butanol, 2-butanol, isobutanol and *tert*-butanol. It is mainly used as solvent, intermediate in chemical synthesis, and as fuel. The global market for this important bulk chemical and excellent blend-in fuel is very large, projected to reach USD 5.6 billion by 2022. Most butanol is presently as produced from

fossil resources, fossil butanol. Additionally, there are biological routes for renewable butanol production, mainly to produce 1-butanol and isobutanol. Existing bio-based 1-butanol is produced from starch, sugar, or cellulose such as wheat, beet, corn and wood. Products of these fermentation processes additionally include acetone and ethanol. It is the understanding that the existing, and further developed, bio-butanol market will grow significantly in the near future. Photosynthetic butanol directly from solar energy and carbon dioxide will be the most sustainable and carbon dioxide neutral production of this important bulk chemical and fuel.

Adding *kivd* and *adh* on a self-replicating plasmid into cells of *Synechocystis* PCC 6803 (*Synechocystis*) introduced the ability to form isobutanol in the cyanobacterial cells in a direct process from solar energy and carbon dioxide [1]. *Kivd* was then identified as the bottleneck for isobutanol production. Single modifications of selected amino acids in the active center of the enzyme, based on knowledge-based modelling, resulted in significantly improved versions of *Kivd* with higher catalytic activities in the cells [2]. Using the best isobutanol producing strain in long-term experiments (as detailed below) resulted in a maximal cumulative titer of 911 mg photosynthetic isobutanol per L with a maximum rate of 43.6 mg photosynthetic isobutanol per L and day [3].

By extensive work, we designed and constructed cyanobacteria that form significant levels of photosynthetic 1-butanol in a direct process [4]. Based on wild-type cells of *Synechocystis*, we introduced the ability to produce 1-butanol from acetyl-CoA and pyruvate, a pathway consisting of six individual genes introduced into three different locations in the genome. In addition, we increased the flow of carbon (C) from the Calvin-Benson-Bassham cycle to the product by introducing the enzyme phosphoketolase and by removing the ability to form the storage product acetate - all with the aim to increase the amount of 1-butanol being formed. Our systematic work generated many different genetically modified cyanobacterial strains with progressively higher 1-butanol production, from wild-type cells via strain BOH1 to the final

strain BOH78 with the highest photosynthetic 1-butanol production, Figure 1 [4].

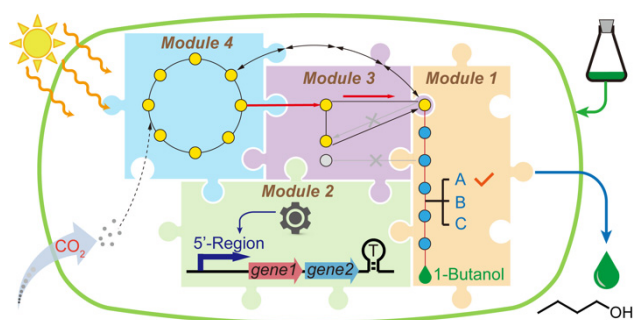


Fig. 1: Systematic modular engineering of the unicellular cyanobacterium *Synechocystis* PCC 6803 resulted in cells with significant levels and rates of photosynthetic 1-butanol production from solar energy and carbon dioxide [4]

When strain BOH78 was grown on smaller scale (25 mL) in long-term trials (at least 1 month) with daily manual adjustments of pH and cell culture removal (removing 2.5 mL (10%) of cell culture replaced by 2.5 mL of new medium) every second, respectively each day, the maximum amount of 1-butanol in the cell culture was 1.86 and 2.13 grams of photosynthetic 1-butanol per L (*in-bottle* titer), respectively. If we add the produced amount of 1-butanol that was removed every second, or each day, the total production became 3.0 and 4.8 grams photosynthetic 1-butanol per L (*cumulative* titer), respectively, with a maximum rate of 302 mg photosynthetic 1-butanol per L and day [4]. This was, within the European Photofuel consortium, increased to a rate of 600 mg photosynthetic 1-butanol per L and day with a carbon partitioning of 60%, i.e. 60% of the fixed carbon ends up in the secreted product butanol (instead of as biomass) [5]. This is a very high production of a solar fuel/chemical in a direct process from solar energy, carbon dioxide and water – it holds great promise for the future.

3. Generation of a functional semisynthetic [FeFe]-hydrogenase in cyanobacteria

Almost all cyanobacteria contain an inherent capacity to produce hydrogen gas catalyzed by a nitrogenase and/or a bi-directional hydrogenase. Wild type (WT) cells of *Synechocystis* harbor a single, bidirectional, oxygen sensitive [NiFe]-hydrogenase encoded by the *hox*-genes. Hydrogen production occurs both under dark, fermentative conditions and, in a short burst, during the transition from darkness to light. In darkness, the hydrogen production is dependent on external or internal fermentation substrate(s). In light, the native hydrogenase can utilize electrons from PSII-catalyzed water splitting and/or fermentative sources, although it is rapidly inactivated by oxygen generated by PSII. However, neither nitrogenase nor [NiFe] hydrogenase can be considered Nature's premier H₂ production catalysts. Indeed, the [FeFe]-hydrogenases (HydA), found in some eukaryotes and in prokaryotes outside the cyanobacteria phylum,

are known to have catalytic activities that vastly outperform any hydrogen producing enzyme naturally occurring in cyanobacteria. Moreover, the [FeFe]-hydrogenases are also much less prone to catalyze the oxidation of H₂ to protons and electrons. Thus, the introduction of such an [FeFe] hydrogenase is expected to result in significant enhancement in H₂ production from the cyanobacterial host. An heterologous expression of HydA is complicated by the requirement for efficient co-expression of its dedicated maturation machinery.

We expressed the [FeFe]-hydrogenase HydA1 from the green algae *Chlamydomonas reinhardtii* (CrHydA1) in WT and a hydrogenase deficient mutant ($\Delta hoxYH$) of *Synechocystis* PCC 6803 (*Synechocystis* WT and *Synechocystis* Δhox , respectively) using a broad host range shuttle vector [6]. To reliably express the *hydA1*-gene in the cyanobacterial host strains, we used the Ptr_{core} promoter together with a bicistronic design adapter. The strain *Synechocystis* Δhox offers a unique opportunity to examine the activity of a heterologous hydrogenase without interference from any native hydrogen metabolism. A functional semisynthetic [FeFe]-hydrogenase was formed by adding a synthetic [2Fe] subcluster mimic [Fe₂(adt)(CN)₂(CO)₄]²⁻ (complex 1, adt: -SCH₂NHCH₂S-), the active site of the [FeFe]-hydrogenase, to the cell cultures, Figure 2 [6].

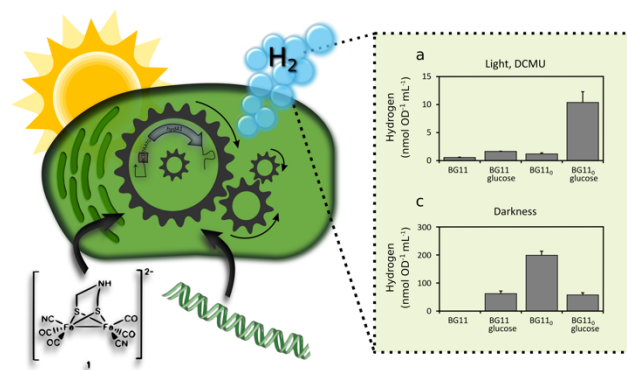


Fig 2: A functional, semi-synthetic and heterologously expressed [FeFe]-hydrogenase is generated in *Synechocystis*, resulting in cells with sustained hydrogen production [6]

The activated hydrogenase evolved hydrogen both in light and in darkness with an activity directly linked to the metabolic status of the cell. This was the first example of synthetic activation of a metalloenzyme within any photosynthetic microorganism.

In following work, we extended the work of artificial activation to two newly discovered [FeFe]-hydrogenases, *Sm* HydA from *Solobacterium moorei* and *Tam* HydS from *Thermoanaerobacter mathranii*, in cells of *Synechocystis* [7]. We demonstrated *in vivo* H₂-evolution activity from both enzymes after synthetic activation by the diiron subcluster mimic. *Synechocystis* cells carrying the activated *Sm* HydA exhibits clear H₂ production under different environmental conditions, and the semisynthetic enzyme remains stable *in vivo* for several days [7].

Acknowledgements

The author acknowledges financial support by the European Union Horizon 2020 Framework Program under the grant agreement number 640720 (Photofuel), the Kamprad Family Foundation for Entrepreneurship, Research & Charity (Photosynthetic butanol), the NordForsk NCoE program "NordAqua" (project number 82845), and the Swedish Energy Agency (CyanoFuels, project number P46607-1). Special thanks to Dr Gustav Berggren and his research group as well all to all past and present members of the Microbial chemistry research group at the Department of Chemistry-Ångström, Uppsala University.

References

- [1] R. Miao, X. Liu, E. Englund, P. Lindberg, P. Lindblad, *Metabolic Engineering Communications*, 5 (2017), 45-53.
- [2] R. Miao, H. Xie, F.M. Ho, P. Lindblad, *Metabolic Engineering*, 47 (2018), 42-48.
- [3] R. Miao, H. Xie, P. Lindblad, *Biotechnology for Biofuels*, 11 (2018), 267.
- [4] X. Liu, R. Miao, P. Lindberg, P. Lindblad, *Energy & Environmental Science*, 12 (2019), 2765-2777.
- [5] J. Wichmann, K.J. Lauersen, N. Biondi, M. Christensen, T. Guerra, K. Hellgardt, M. Kuronen, P. Lindberg, C. Rösch, I.S. Yunus, P. Jones, P. Lindblad, O. Kruse, *Trends in Biotechnology*, 39 (2021), 323-327.
- [6] A. Wegelius, N. Khanna, C. Esmieu, G.D. Barone, F. Pinto, P. Tamagnini, G. Berggren, P. Lindblad, *Energy & Environmental Science*, 11 (2018), 31633167.
- [7] A. Wegelius, H. Land, G. Berggren, P. Lindblad, *Cell Reports Physical Science* (2021), doi: 10.1016/j.xcrp.2021.100376.

Development of a genetic biomonitoring test for the investigation of pollinator-plant-interactions

Lisa Prudnikow¹, Robert Leidenfrost¹, Catrin Westphal², Röbbe Wünschiers¹

¹Hochschule Mittweida, Technikumplatz 17, 09648 Mittweida

²Georg-August-Universität Göttingen, Grisebachstr. 6, 37077 Göttingen

There is a world-wide decline in biodiversity recorded. Especially insects and accompanying pollinators are threatened. When the foraging behaviour of pollinators is understood in detail, future crop and floral pollination services can be sustained and it is possible to establish projects for the conservation of pollinators and plant biodiversity. With the use of nanopore sequencing methods it is possible to detect pollen species that were collected by pollinators by their genetic information. In this study, a protocol for portable nanopore sequencing of DNA from pollen that was collected by honey bees, bumble bees and wild bees is being designed. DNA-metabarcoding is used to identify species within the mixed DNA sample. The ITS2-region will be used as a barcode. We will investigate pollen preferences of three pollinator species by placing their hives or nests at the same. Based on the results, landscape management schemes are developed that target pollen preferences and nutritional requirements of managed and wild social bee species as well as solitary wild bees.

Ein weltweiter Rückgang der Artenvielfalt ist zu verzeichnen. Besonders Insekten und damit einhergehend Bestäuber, sind bedroht. Wenn das Nahrungssuchverhalten von Bestäubern im Detail verstanden wird, können zukünftig die Bestäubungsleistungen von Feldfrüchten und Blumen aufrechterhalten werden. Außerdem können Projekte zum Erhalt von Bestäubern und der Pflanzenbiodiversität etabliert werden. Mit dem Einsatz der Nanoporen-Sequenzierung ist es möglich, Pollenarten, die von Bestäubern gesammelt wurden, anhand ihrer genetischen Information zu erkennen. In dieser Studie wird ein Protokoll zur portablen Nanoporen-Sequenzierung von DNA aus Pollen, der von Honigbienen, Hummeln und Wildbienen gesammelt wurde, entwickelt. Wir verwenden DNA-Metabarcoding, um Arten innerhalb der gemischten DNA-Probe zu identifizieren. Dazu wird die ITS2-Region als Barcode verwendet. Mit einer Biodiversitätsanalyse wird untersucht, wie die drei Bestäuberarten die umliegenden Pollenressourcen nutzen, wenn sich ihre Stöcke bzw. Nester am selben Standort befinden. Basierend auf den Ergebnissen werden Landschaftsmanagementpläne entwickelt, die auf die Pollenpräferenzen und den Nährstoffbedarf von bewirtschafteten und wilden sozialen Bienenarten sowie solitären Wildbienen abzielen.

1. Introduction

In the past 20 years, high-throughput DNA sequencing (HTS) has proven to be a powerful tool for tracking species biodiversity. Taxonomic identities can be identified based on short DNA sequences [1]. Traditional monitoring processes such as observing or counting individuals are less precise and require high time effort (due to geographical or seasonal factors) and exceptional taxonomic expertise. Due to less standardised sampling techniques of traditional monitoring methods, for large-scale analyses, genomic methods are more suitable. And as species declines are becoming more and more serious, a reliable method is needed [2]. With this in mind, environmental DNA (eDNA) analyses are appropriate in the sense that species under study are not injured. A mixture of DNA from different organisms that can be found in an environmental sample like faeces, water, soil or other sediments is called eDNA. It has proven to be a useful material for biodiversity monitoring [3] as interactions between organisms and the environment can be characterized [4].

Great need exists in the analysis of plant-pollinator networks, not only in agricultural landscapes, but also in urban ecosystems [5; 6]. Pollinators are indispensable, due to their pollination services (Figure 1). This is expressed not only by the growth of important crops but also by the genetic mixing of wild and cultivated plants [7]. Honey bees for example, store pollen as an energy resource to supply their colony. In the form of pollen loads, they transport the captured pollen grains from the flower to their hive [1]. In contrast, mason bees collect pollen on their abdomen and store it in a clump at their nest site for their offspring [8]. eDNA analyses based on collected pollen determines the plant species most used by pollinators. With this diet analysis based on pollen-DNA, the diversity of the surrounding flora can be measured and evaluated [4].

To identify species from a mixed-DNA sample so called DNA-metabarcoding is used. Therefore, not a whole DNA strand is used, but only one DNA-segment, so called barcode. An ideal barcode is as short as possible (<100-150 bp), displays a highly variable sequence and it is en-



Figure 1: Bumble bee (*Bombus terrestris*) during foraging. Pollen grains start to accumulate on her legs to form loads

closed by two conserved regions serving as primer motifs [3]. Realizing this approach with mixed community samples, it allows species detection and diversity measures. That means, a barcode represents a species and its quantity within a mixed sample due to its variable sequence. DNA-Metabarcoding is performed with polymerase chain reaction (PCR) and subsequent HTS-methods such as next generation- or nanopore sequencing [9]. For floral taxonomic issues, barcodes such as the internal transcribed spacer region (ITS2) or the chloroplast genes *rbcl* or *matK* and others are used. ITS2 reveals a high level of discrimination at species level and is therefore often used [10]. With metabarcoding of pollen DNA, plants visited by pollinators can be identified [11]. And this kind of information can be used to promote adequate floral resources at the appropriate place and time for an efficient habitat management [6]. For pollen metabarcoding, it requires a good cell breakage method to reach the pollen contents such as the DNA [12]. The pollen wall of seed plants, called sporoderm, is composed by two layers: the inner intine and the outer exine. The exine, mainly consists of the polymer sporopollenin and is very robust as it is acetolysis- and decay-resistant [13]. These morphological traits enable the preservation of the pollen nutrients [12]. As pollen contains carbohydrates, lipids, vitamins, minerals and all the basic amino acids it is of great importance for pollinators health. Thus, the quality of honey bee products like honey or propolis for example are influenced by pollen composition and its diversity. So, to identify the visited plants is also of importance to guarantee consumers and products safety [11] next to the evaluation of the plant environment [4]. As already mentioned, genetic biodiversity monitoring can be realized with HTS methods such as nanopore sequencing. Devices from Oxford Nanopore Technologies (ONT) like the MinION can be used to perform a portable sequencing process in field or areas without laboratory infrastructure [9]. This nanopore-based sequencing technology allows rapid analyses of DNA samples on site and avoids transport to distant laboratories. For eDNA studies this is an advantage [14]. The fact that DNA can be decoded is achieved with the

membrane potential. Using a motor protein, single-stranded DNA passes through nanopores embedded in a bilipid double layer membrane. When an external unilateral voltage is applied, the ion current can be detected in a measured manner, which is characteristically blocked by the DNA (Figure 2) [15]. With the help of this technology a genetic biodiversity test is to be developed, to evaluate the floral biodiversity and the foraging behaviour of pollinators. The protocol that will be designed in this process should be useful for ecological questions. It is the aim to create a protocol, that allows DNA extraction, sequencing and data analysis right at the study site. At the same time during this study, the generated data will be used for further research questions: 1. What are the main pollen species used by honey bees, bumble bees and wild bees in the respective habitats (urban, urban edge, village, forest edge, near field) when these pollinators are located at the same site? 2. To what extent does pollen use overlap among the four pollinators and do they share important pollen resources? 3. How do the main pollen sources of the three pollinator species change over the time of April-August? 4. Which properties has the recorded pollen diet of the pollinators concerning bee health and the amino acid composition of certain pollen species?

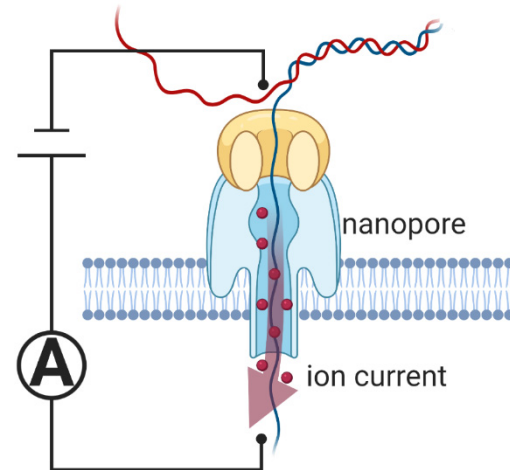


Figure 2: Principle of nanopore sequencing. The ion current in a nanopore protein embedded in a membrane is characteristically blocked when DNA transfers it (created with *biorender.com*)

2. Methods

2.1 Study Area and Pollen Sampling

Sampling of collected pollen from wild bees, honey bees and bumble bees is taking place in the vicinity of the cities of Chemnitz and Frankenberg. At these sites honey bee hives of private beekeepers are installed. For this study a bumble bee colony and a nesting aid for wild bees is installed at each of these sites as well to record the resource use of the three pollinator species. From the middle of April until the end of August, every two weeks pollen loads resp. wild bees stored pollen will be yielded. From honey bee colonies, pollen loads will be collected with pollen traps. Wild bees stored pollen bags

will be collected with a spatula. For the time of sampling, pollen will be yielded from the cells of the plates from the nesting aid. From the bumble bee colonies, pollen loads will be collected with insect bottles.

2.2 Pollen disruption and DNA extraction

As pollen is a very robust natural material, an efficient disruption method needs to be applied. Therefore, a series of tests with a ball mill and a vortex device was carried out with commercial pollen loads and tulip pollen. The size and type of ceramic and steel beads as well as the speed and time of the ball mill treatment were varied in this process. The number resp. weight of pollen loads was varied, depending on the manufacturer's instructions from the used DNA-extraction kit or literature references. The success of the mechanic pollen disruption was checked with a microscope. First, only samples with completely destroyed pollen grains were used for a subsequent DNA extraction. Later samples with half cracked pollen were used as well for a better assessment of the experimental conditions. For DNA extraction three different extraction kits were used (GenElute™ Plant Genomic DNA Miniprep Kit from Sigma-Aldrich, MasterPure™ Complete DNA and RNA Purification Kit from epicentre and the DNeasy® Plant Mini Kit from Qiagen). The success of the DNA extraction from the disrupted pollen loads and tulip pollen was checked with measuring DNA purity and concentration. Therefore, UV absorbance was measured with a UV-Vis Spectrophotometer (NanoVue™ plus GE Healthcare GmbH). Additionally, the concentration of the extracted pollen DNA was measured with a fluorometer (Qubit® 2.0, Invitrogen Corporation).

2.3 Nanopore Sequencing and Read-Processing

For the plant identification of the mixed pollen samples, the ITS2 region, as a standard plant barcode [10] is applied for the amplification of the pollen DNA. For sequencing, the amplified ITS2 region of the pollen DNA is fragmented. Thus, multiple samples (12-24) can be sequenced at the same time, the DNA library will be prepared with EXP-NBD104 resp. EXP-NBD114 and SQK-LSK109 kits (ONT). Nanopore sequencing is carried out with the MinION Sequencer and the Flongle Adapter using the R9.4.1 Flongle flow cell (Figure 3).



Figure 3: Flongle adapter attached to MinION sequencer from ONT for sequencing prepared DNA samples

Resulting data files are base-called using the ONT-software GUPPY. The generated sequence data are further analysed for subsequent BLASTs. For the sequence data analysis, the datasets are examined with bash command line-based programming. For this, several steps must be performed in series [9]. In the first step, basecalled reads are demultiplexed, to assemble the sequence data to sampling modules (one file of sequence data corresponds to the pollen collected by one pollinator species at a certain time). Therefore, the third-party application DeepBinner resp. Porchop is used [16]. For a first quality check of the generated reads (number of reads, read length, per base sequence quality, per base N content, GC content per base etc.) NanoPack [17] and FastQC [18] are applied. VSEARCH is used as well for checking phred scores for an additional quality assessment of the raw nanopore generated reads [19]. With USEARCH [20] reads are filtered with a mean phred score of a minimum threshold (default resp. value depending on first quality check) and minimum read length (70 bp).

2.4 ITS2-BLAST and Biodiversity analysis

For plant species monitoring based on sequencing methods, a sequence database of the used barcode is necessary for subsequent BLAST analyses. Therefore, a global ITS2 database [21] is used in this project. As the sampling of the pollen loads will be carried out in Saxony, the database was filtered with a list of Saxon plants. To create a complete list of all plants that have been described in Saxony, three publications from the *Sächsisches Landesamt für Umwelt, Landwirtschaft und Geologie* were merged [22, 23]. These are ferns and seed plants, woody plants and herbaceous plants. Furthermore, the ITS2-Database was filtered with a list of German Plants [24]. Boxes for trap nesting bees and bumble bees as well as the beehives for honey bees are placed in resp. near urban ecosystems. As in that regions non-native plants could be possibly exist due to private gardening [5], the unfiltered global ITS2 plant database is used for BLAST's as an additional control. For Biodiversity analysis the software RStudio (version 1.2.5001) is used. After BLASTing the generated reads against the ITS2-databases, it is calculated how often plant species occur in the samples. With this numerical values pollen richness, Shannon- and Simpson diversity index is calculated to describe samples' pollen type diversity (vegan R-package). To compare them for the three pollinator species Kruskal-Wallis is carried out (psych R-package). Pollen-use networks for each pollinator per sample are created as well (bipartite R-package).

3. Results and Discussion so far

The commercial pollen loads are very hard in their appearance. As these loads are sold as human food, it is possible that they are specially treated to be preservable like freeze-dry or heating. Consequently, it can be assumed that the cell components are somehow denatured or unusable. This assumption was confirmed during the following DNA extraction experiments within this

test series. Microscopic analysis showed that in some samples (both tulip and commercial Pollen loads) every pollen grain was damaged in a way after the treatment. In other samples half of the pollen grains were still intact or only cracked. This could have been observed when the samples were disrupted with both the ball mill and the vortex device (Figure 4).

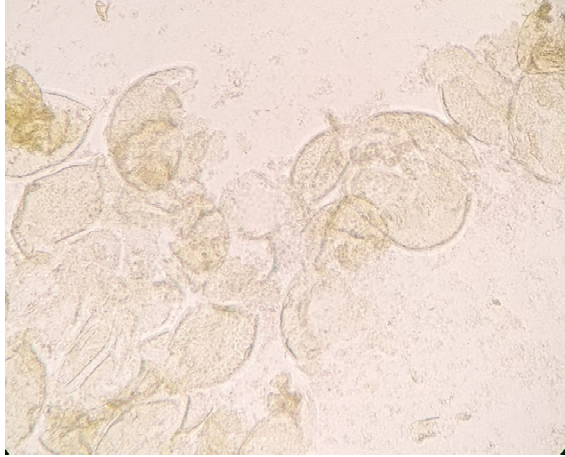


Figure 4: Tulip pollen loads that were mechanically disrupted.

After pollen disruption and DNA extraction of 62 samples it can be summarized: when using DNeasy® Plant Mini Kit from Qiagen the best results can be achieved for tulip pollen disrupted with ceramic beads in 200 µl lysis buffer with a vortex device. But it has to be kept in mind, that fresh pollen loads, that are yielded from pollinators are treated with pollinators salvia to get the ball shape of a pollen load [1]. Or that the pollen that is collected from wild bees stored in shape of a pollen bag, is treated with salvia and nectar as well [8]. In this part of the experiment, this fact was not considered yet, due to the autumn and winter season.

References

- [1] Kamo, T.; Kusumoto, Y. & Tokuoaka, Y. et al. (2018) *Applied Entomology and Zoology* 53:353-361.
- [2] Hardulak, L.; Morinière, J. & Hausmann, A. et al. (2020) *Molecular Ecology Resources*. 20(6):1542-1557
- [3] Taberlet, P.; Bonin, A. & Zinger, L. et al. (2018) Oxford University Press.
- [4] Ruppert, K., Kline, R. & Rahman, S. (2019) *Global ecology and conservation*. 17:e00547
- [5] Udy, K.; Reininghaus, H. & Scherber, C. et al. (2020) *Ecosphere*. 11(2).
- [6] Bertrand, C.; Eckert, P. & Ammann, L. et al. (2019) *Journal of Applied Ecology*. 56:2431-2442
- [7] Bender, S.; Chemnitz, C. & Klein, A.-M. et al. (2020) Heinrich-Böll-Stiftung.
- [8] Kornmilch, J. (2010) Deutsche Bundesstiftung Umwelt.
- [9] Krehenwinkel, H.; Pomerantz, A. & Prost, S. (2019) *Genes*. 10(11):858.
- [10] Yao, H.; Song, J. & Liu, C. et al. (2010) *PLoS one*. 5(10):e13102.
- [11] Galimberti, A.; De Mattia, F. & Bruni, I. et al. (2014) *PLoS One*. 9(10):e109363.
- [12] Yang, Y.; Zhang, J. & Zhou, Q. et al (2019) *J Sci Food Agric*. 99(12):5350-5357
- [13] Halbritter, H.; Ulrich, S. & Grímsson, F. et al. (2018) Springer Verlag.
- [14] Maestri, S.; Cosentino, E. & Paterno, M. et al. (2019) *Genes*. 10(6):468.
- [15] van Dijk, E.; Jaszyszyn, Y. & Naquin, D. et al. (2018) *Trends in Genetics*. 34(9):666-681.
- [16] Wick, R. (2018) Porechop. Retrieved from <https://github.com/rrwick/Porechop>
- [17] De Coster, W.; D'Hert, S. & Schultz, D. et al. (2018) *Bioinformatics*. 34(15):2666-2669.
- [18] Andrews, S. (2020) FastQC. Retrieved from <https://github.com/s-andrews/FastQC>
- [19] Rognes, T.; Flouri, T. & Nichols B. (2020) VSEARCH. Retrieved from: <https://github.com/torognes/vsearch>
- [20] Edgar, R. C. (2018) USEARCH. Retrieved from <https://drive5.com/usearch/download.html>
- [21] Förster, F. (2015) eukaryota.all.fasta.bz2. Retrieved from: https://github.com/BioInf-Wuerzburg/ITS2_database_update_2015/blob/master/data/2015/eukaryota.all.fasta.bz2.
- [22] "Artenreferenzliste Sachsen". Unknown author. Retrieved from: <https://www.natur.sachsen.de/artenreferenzliste-sachsen-21920.html>
- [23] „Gebietseigene Pflanzen“. Artenlisten Sachsen. Unknown author. Retrieved from: <https://www.natur.sachsen.de/gebietseigene-pflanzen20826.html>
- [24] Ludwig, G. & Schnittler, M. (1996) *Schriftenreihe für Vegetationskunde* 28

Which pollen do bees like best? - Analysis of mixed pollen samples with metabarcoding

Svenja Bänsch^{1,2}, Teja Tschardt³, Röbbke Wünschiers⁴, Catrin Westphal¹

¹ Functional Agrobiodiversity, University of Göttingen, Germany

² Centre for Biodiversity Monitoring, Zoological Research Museum Alexander Koenig Bonn, Germany

³ Agroecology, University of Göttingen, Germany

⁴ Department of Biotechnology and Chemistry, Mittweida University of Applied Sciences, Germany

Bees are important pollinators for many crops and wild plants, however, they are threatened. Among other factors, pollen nutrition can affect reproduction and bee health. We studied pollen foraging preferences two major pollinator species, honey bees and bumble bees, in agricultural landscapes during the flowering of oilseed rape and strawberry. Mixed pollen samples were investigated using ITS2 metabarcoding and we compared the quantitative outcomes to microscopic pollen grain counts. We found more different plant genera in bumble bee compared to honey bee samples. Oilseed rape was the most frequent resource for both species, while the frequency of strawberry was higher in honey bee than bumble bee samples. The relationship between amplicon reads and microscopic pollen counts was positive but pollen-type specific. Conservation schemes should promote flower-rich landscapes in abundance and diversity to support different pollinator species. Metabarcoding could provide similar quantitative information to microscopy but those results are pollen type specific and susceptible to bias throughout the process.

1. Introduction

Plants need to be pollinated to reproduce. Among insects, bees are the major and most effective pollinators. It has been shown that the seed set, yield and fruit quality of crops and wild plants can be greatly improved when insect-pollinated [1-3].

However, pollinator abundance and species richness are in decline. One major reason is agricultural intensification which is accompanied with e.g. the fragmentation of foraging habitats like structural elements, and homogeneous cropping of crops with low quality pollen such as cereals and maize. Those changes reduce the forage and nesting possibilities but also the diversity and abundance of host plant species [4, 5].

Particularly western honey bees (*Apis mellifera*) but also bumble bees (mainly *Bombus terrestris*) are managed to boost pollinator numbers in the field [6,7]. Both bee species are generalists and visit a great variety of plant species [8].

To support pollinators in agricultural landscapes many factors such as pollen and nectar sources, nesting opportunities and bee health must be considered [9]. With this study we focus on the pollen foraging preferences. Pollen is especially important to feed the bee larvae and the quality and diversity of pollen can affect the development of colonies [9-12].

To identify the pollen resource of bees, pollen identity can be investigated using metabarcoding and microscopy [13, 14]. In comparison to microscopy, metabarcoding can provide a higher taxon accuracy and has a higher throughput with a predictable cost and time frame [14-16]. However, metabarcoding is often limited

in pollen quantification due to over- or underestimation caused by e.g. polymerase chain reaction amplification or barcode copy number [17]. While pollen identification with the microscope needs great expertise and is a time-consuming task [15], target crop pollen can be counted by non-experts when pollen grains have a characteristic surface structure.

We investigated the pollen resource use of managed honey bee and bumble bee colonies next to an economically important crop, strawberry, with varying amounts of oilseed rape (OSR) in the surrounding agricultural landscapes. OSR is a mass-flowering crop (i.e. many flowers that provide large amounts of pollen and nectar) in comparison to strawberry that have a rather inconspicuous flowering.

With this study we want to answer following questions:

- 1) Which resources do bees use in the agricultural landscapes and how do their pollen preferences differ? Do they use strawberry and OSR as frequent resources?
- 2) Are quantitative results of metabarcoding are comparable with microscopic pollen grain counts?

2. Methods

2.1. Pollen collection

We placed two honey bee and bumble bee colonies next to nine strawberry fields in agricultural landscapes. The study landscapes were in vicinity to the cities Göttingen and Kassel, Germany. Small and large honey bee and bumble bee colonies were observed respectively. We recorded flower cover data of main crop resources as

this can affect the foraging of bees. The flower availability of OSR is quantified as the product of land cover within 2000 m surrounding the colonies and the flower cover within the nearest field during each observation date. The flower cover in strawberry fields was quantified by counting open flowers along a 2 m row within the field.

During the flowering of strawberry and OSR, we collected pollen at four different times in May and June 2016 from each colony respectively. Both bee species transport collected pollen as pollen loads on their legs (Fig. 1). The pollen from honey bees were collected using a pollen trap that can be installed in front of the colony. When honey bees enter the colony, they will lose their pollen loads. During each observation date, we placed the trap in front of the colony for 30 minutes. Pollen from bumble bees were collected by catching, if possible, five homecoming bumble bees per observation date. Pollen was carefully removed from their legs, and bees were released.

Pollen from each date were pooled per colony and mixed with 70 % ethanol (ratio 1:4 pollen:ethanol) to get a homogenous sample. From each sample we prepared 1 ml aliquots in 1.5 ml reactions tubes and let them dry with open lids in a fume cupboard for 72 h.



Fig. 1: Honey bee with pollen load on her hind leg

2.2. Metabarcoding

We used metabarcoding of ITS2-region PCR amplicons to quantify the richness of pollen collected by small and large honey bee and bumble bee colonies.

DNA from ~ 0.015 g dried pollen was extracted using the Qiagen DNEasy Plant Mini Extraction Kit following modifications of the protocol of Bänisch *et al.* [13].

We created a local data base of plants that are known to occur in Lower Saxony [18]. Therefore, we extracted plant-specific ITS2 sequences from a global eukaryote database [19]. This subset was used to create a magicBLAST database (version 1.4; [20]) and all paired reads that both aligned to a database entry (plant ITS2 sequence) with at least 50bp and a similarity >98 % were kept.

For each matching read, we calculated an alignment score by multiplying the alignment length with the alignment identity; the scores for the forward and reverse read were summed to get the final score for each read-pair. Read-pairs that matched several entries were ordered by this score and only the top scoring match (plant species) per read was counted. As some plant species

have very similar ITS2 sequences and we therefore cannot unambiguously distinguish them on a species level, we decided to use amplicon read-based identification at the genus level only.

We assume that *Brassica* equals OSR and *Fragaria* equals strawberry. We are aware that other species belong to these genera but during the observation period we were unable to assign any further species to these genera during recordings in flowering habitats (data not included here).

2.3. Microscopy

From each aliquot we prepared pollen slides to count the pollen grain numbers of main crop species (strawberry and OSR) to compare the quantitative outcomes with metabarcoding results. One aliquot was diluted with water (1:4 ratio pollen:water) and one drop of the pollen-water mixture was added to an observation microscopic slide together with a drop Kayer's gelatine stained with fuchsine. For each slide, we counted 200 pollen grains at 400 x magnification and categorized them as strawberry, OSR or others.

2.4. Statistics

Pollen diversity

Our analysis of pollen diversity is based on the presence/absence of ITS2 sequences of certain plant genera only. Pollen richness represents the number of plant genera occurring in one pollen sample.

To analyse the effects of bee species on pollen diversity we used generalised linear mixed effect models using the glmmTMB package [21]. First, we created a global model with pollen richness as response variable and following explanatory variables: bee species (honey bee / bumble bee), colony size (small and large honey bee and bumble bee colonies respectively), flower availability of OSR and strawberry. Colony ID nested in study landscape was included as random term. The final model was fitted with a negative binomial error. We applied the multimodel inference approach using the function dredge (package MuMIn; [22, 23]) that created a list of all possible model combinations where we limited the number of parameters in each model to three to avoid overfitting. We considered all models that have a delta AICc < 2 (second Akaike Information Criterion) compared to the best fitting model (dAICc = 0). The relative importance of each explanatory variable was assessed using the sum of Akaike weights ($\sum w_i$) over all candidate models which included the respective variables (function importance, package MuMIn; [23]). In this manuscript, we only considered parameters that were within the best fitting models with $\sum w_i > 0.5$ (possible range of $\sum w_i$ 0-1). Differences between species were tested using a posthoc test (function emmeans) with an alpha of 0.05 [24].

In order to describe the pollen resource utilization of honey bee and bumble bee colonies we show the five

most common plant genera in pollen samples summarized over colony sizes and the whole study period (i.e., frequency). These presence / absence data are merely described and not statistically analysed.

Comparison of metabarcoding and microscopy

To test for associations between quantitative outcomes of ITS2 metabarcoding and microscopy, we fitted linear regression models for log ITS2 amplicon reads + 1 versus log microscopic pollen grain counts + 1 for strawberry and OSR respectively.

All statistical analyses were performed with the software R version 3.6.2 [25].

3. Results

Pollen diversity

We found a significantly higher number of plant genera in bumble bee than honey bee samples ($\Sigma wi = 1$; $p < 0.05$; Fig. 2). The importance (i.e. Σwi) of flower availability and colony size was below 0.5 and hence, not considered as meaningful results in this manuscript.

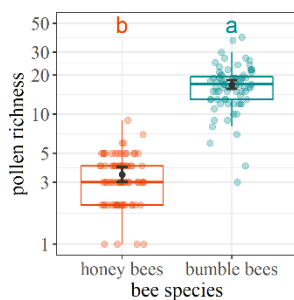


Fig. 2: Effect of bee species (a, $\Sigma wi = 1$) on pollen richness (i.e., number of plant genera). Different letters indicate significant differences obtained from post hoc Tukey tests (significance level of 0.05). Note that pollen richness is shown on a log scale.

In total, we found 32 different plant genera ($n=77$ pollen samples) in honey bee samples while we found 109 different plant genera in bumble bee samples ($n=75$ pollen samples). The five most frequent plant genera found in honey bee and bumble bee samples are shown in Tab. 1. *Brassica* (OSR) was the most frequent plant genera used by both species (honey bees: in 60 out of 77 pollen samples; bumble bees: in 59 out of 75 pollen samples). *Fragaria* (strawberry) was frequently found in honey bees samples as well (in 57 samples) while other plant genera were less frequently found (*Rosa*: 19, *Salix*: 12, *Crataegus*: 8). In bumble bees we found *Acer* and *Monotropa* (in 57 samples) as almost equally often as *Brassica*. *Fragaria* was found in less than half of the pollen samples (26) like *Sedum* that was found in 25 samples.

Tab. 1: Five most common pollen resources of honey bee and bumble bee colonies across colony sizes, all landscapes and observations dates. Frequency is the number of samples in which the ITS2 sequence was detected and is based on presence/absence data of plant genera in pollen samples.

| Bee species | Taxa | Frequency |
|----------------------|------------------|-----------|
| Honey bees n = 77 | <i>Brassica</i> | 60 |
| | <i>Fragaria</i> | 57 |
| | <i>Rosa</i> | 19 |
| | <i>Salix</i> | 12 |
| | <i>Crataegus</i> | 8 |
| Bumble bees n=75 | <i>Brassica</i> | 59 |
| | <i>Acer</i> | 57 |
| | <i>Monotropa</i> | 57 |
| | <i>Fragaria</i> | 26 |
| | <i>Sedum</i> | 25 |

Comparison of metabarcoding and microscopy

We found positive relationships between the number of ITS2 amplicon reads and microscopic pollen counts being stronger for strawberry pollen ($R^2 = .69$) than for OSR pollen ($R^2 = .15$). The positive intercepts for both plant genera (i.e., when microscopic pollen counts = 0) indicate that ITS2 metabarcoding was able to detect pollen when microscopy failed (Fig. 3).

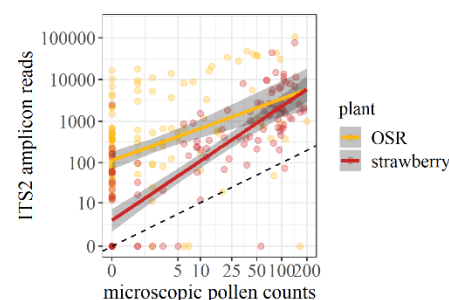


Fig. 3: Relationships between quantitative outcomes of ITS2 amplicon reads and microscopic pollen counts for strawberry and OSR pollen. The regression was calculated for strawberry and OSR pollen separately as follows: $\text{lm}(\log(\text{no. of reads} + 1) \sim \log(\text{microscopic counts} + 1))$. Solid lines show the regression lines, and dotted line the angle bisector with intercept = 0 and slope = 1

4. Discussion

Pollen diversity

Our target crop species OSR and strawberry, that were flowering during the whole study period, were continuously visited by honey bees, but also bumble bees used in particular OSR quite frequently. Particularly, bumble bees (*Bombus terrestris*) exploited a greater richness of plant genera in agricultural landscapes compared to honey bees. Hence, bumble bees may benefit from a greater diversification of flower resources in the landscape. At the same time, they do potentially contribute to the pollination of a greater number of plant species.

Both species are generalist, that means, that they can collect pollen and nectar from a great variety of plant species [26]. However, honey bees often focus on a few, preferably mass-abundant flowering species. They alternate flower resources only to a limited extent during foraging trips [27, 28] which might be due to their ability to communicate resources via the waggle dance [29].

Large amounts of pollen came from woody and herbaceous plant genera which can be found in, for example, hedgerows and field groves and have been highlighted by other studies as important sources for bee nutrition and colony growth [11, 28].

Our results have to be interpreted with caution since we cannot entirely conclude honey bees and bumble bees used the identified plant genera as major pollen resources. Some pollen may have been collected while foraging for nectar. The entire analyses of the data presented in this study including detailed analyses on quantitative results from the microscopic experiment can be found in Bänisch *et al.* [13].

Comparison of metabarcoding and microscopy

Our findings of a pollen type-specific relationship between ITS2 metabarcoding reads and microscopic pollen counts is in accordance with previous findings by Smart *et al.* [16]. To confirm a general positive relationship of both methods, further research is needed to take potentially confounding factors like pollen type identity, amount of pollen and defined compositions into account [16,30,31]. Further, other factors such as DNA extraction, amplification biases and contamination of the samples can affect the quantitative outcomes of metabarcoding as well [30, 31].

5. Conclusions

From the data presented here, we conclude that honey bees and bumble bees use OSR as a major foraging resource. Particularly honey bees preferred crop resource over others while bumble bees visited also other plant genera frequently. Hence, conservation management should not only focus on increasing flower availability in landscapes but should also consider the diversity of foraging preferences of different bee species.

While ITS2 metabarcoding is a suitable method to identify the richness of plant genera in mixed pollen samples, quantitative outcomes were restricted. Although we found a positive association of quantitative outcomes of ITS2 amplicon reads and microscopic pollen grain counts, the results were pollen type specific, weak and large proportions of variation were not explained. Both methods, metabarcoding and microscopy, can deliver powerful results but should be chosen carefully and in a targeted manner.

6. Acknowledgements

We would like to thank Bertram Brenig and his team for the support with DNA isolation and sequencing, and Doreen Gabriel for her statistical advice.

7. References

- [1] Klein, A.-M., Vaissière, B. E., Cane, J. H., Steffan-Dewenter, I., Cunningham, S. A., Kremen, C., & Tscharntke, T. (2007). *Proceedings of the Royal Society B: Biological Sciences*, 274, 303–313.
- [2] Bänisch, S., Tscharntke, T., Gabriel, D. & Westphal, C. (2020). *Journal of Applied Ecology*, 58 (2), 476–485
- [3] Wietzke, A., Westphal, C., Gras, P., Kraft, M., Pfohl, K., Karlovsky, P., Pawelzik, E., Tscharntke, T., & Smit, I. (2018) *Agriculture, Ecosystems and Environment*, 258, 197– 204.
- [4] Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W. E. (2010). *Trends in Ecology & Evolution*, 25, 345– 353.
- [5] Vaudo, A. D., Tooker, J. F., Grozinger, C. M., & Patch, H. M. (2015). *Current Opinion in Insect Science*, 10, 133– 141.
- [6] Gosterit, A., & Baskar, V. C. (2016). *Insectes Sociaux*, 63, 609– 614.
- [7] Velthuis, H. H. W., & van Doorn, A. (2006). *Apidologie*, 37, 421– 451.
- [8] Kleijn, D., Winfree, R., Bartomeus, I., Carvalheiro, L. G., Henry, M... Potts, S. G. (2015). *Nature Communications*, 6, 7414.
- [9] Vaudo, A. D., Farrell, L. M., Patch, H. M., Grozinger, C. M., & Tooker, J. F. (2018). *Ecology and Evolution*, 8, 5765– 5776.
- [10] Thorp, R. W. (2000). *Plant Systematics and Evolution*, 222, 211–223.
- [11] Kämper, W., Werner, P. K., Hilpert, A., Westphal, C., Blüthgen, N., Eltz, T., & Leonhardt, S. D. (2016). *Landscape Ecology*, 31, 2245– 2258.
- [12] Brodschneider, R., & Crailsheim, K. (2010). *Apidologie*, 41, 278– 294.
- [13] Bänisch, S., Tscharntke, T., Wünschiers, R., Netter, L., Brenig, B., Gabriel, D. & Westphal C. (2020). *Molecular Ecology*, 29(24), 5003–5018.
- [14] Bell, K. L., Burgess, K. S., Botsch, J. C., Dobbs, E. K., Read, T. D., & Brosi, B. J. (2019). *Molecular Ecology*, 28, 431– 455.
- [15] Keller, A., Danner, N., Grimmer, G., Ankenbrand, M., von der Ohe, K., ..., Steffan-Dewenter, I. (2015). *Plant Biology*, 17, 558– 566.
- [16] Smart, M. D., Cornman, R. S., Iwanowicz, D. D., McDermott-Kubeczko, M., ..., Otto, C. R. V. (2017). *Environmental Entomology*, 46(1), 38– 49.
- [17] Peel, N., Dicks, L. V., Clark, M. D., Heavens, D., Percival-Alwyn, L., Cooper, C., Davies, R. G., Leggett, R. M., & Yu, D. W. (2019). *Methods in Ecology and Evolution*, 10, 1690– 1701.

- [18] Garve, E. (2007). Verbreitungsatlas der Farn- und Blütenpflanzen in Niedersachsen und Bremen. Verbreitungskarten. Nds. Landesbetrieb für Wasserwirtschaft, Küsten- und Naturschutz, Hannover. Retrieved from
- [19] Förster, F. (2015). Eukaryota.all.fasta.bz2. Retrieved from https://github.com/BioInf-Wuerzburg/ITS2_database_update_2015/blob/master/data/2015/eukaryota.all.fasta.bz2
- [20] Boratyn, G. M., Thierry-Mieg, J., Thierry-Mieg, D., Busby, B., & Madden, T. L. (2019). BMC Bioinformatics, 20, 405.
- [21] Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Mächler, M., & Bolker, B. M. (2017). The R Journal, 9(2), 378–400.
- [22] Burnham, K., & Anderson, D. (2002). New York: Springer.
- [23] Barton, K. (2018). MuMIn: Multi-model inference. R package version 1.40.4.
- [24] Lenth, R. (2017). emmeans: Estimated marginal means, aka least-squares means. Retrieved from <https://cran.r-project.org/package=emmeans>
- [25] R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing.
- [26] Amiet, F., & Krebs, A. (2012). Bienen Mitteleuropas: Gattungen, Lebensweise, Beobachtung (1 st ed.). Bern: Haupt Berne.
- [27] Keller, I., Fluri, P., & Imdorf, A. (2005). Bee World, 86, 27–34.
- [28] de Vere, N., Jones, L. E., Gilmore, T., Moscrop, J., Lowe, A., Smith, D., Hegarty, M. J., Creer, S., & Ford, C. R. (2017). Scientific Reports, 7, 42838.
- [29] Couvillon, M. J. (2012). Insectes Sociaux, 59, 297–306.
- [30] Baksay, S., Pornon, A., Burrus, M., Mariette, J., Andalo, C., & Escaravage, N. (2020). Scientific Reports, 10, 4202.
- [31] Pornon, A., Escaravage, N., Burrus, M., Holota, H., Khimoun, A., Mariette, J., Pellizzari, C., Iribar, A., Etienne, R., Taberlet, P., Vidal, M., Winterton, P., Zinger, L., & Andalo, C. (2016). Scientific Reports, 6, 27282.

Long reads for microbial bioaerosol surveillance

Robert Maximilian Leidenfrost¹, Dierk-Christoph Pöther², Henriette Mietke³,
Udo Jäckel² & Röbbbe Wünschiers¹

¹Mittweida University of Applied Sciences, Technikumplatz 17, 09648 Mittweida

²Federal Institute of Occupational Safety and Health, Nöldnerstraße 40-42, 10317 Berlin

³State Operating Company for Environment and Agriculture, Waldheimer Straße 218, 01683 Nossen

Regulation concerning occupational or environmental safety and health in Germany offers limited options with respect to risk assessment for airborne micro-biological agents based on molecular analysis. The knowledge to identify and quantify potentially harmful biological agents in bioaerosols is however required for risk assessment. Subsequently, there is a need for development and evaluation of new analysis methods. Here we evaluate long DNA-sequence reads for profiling microbial bioaerosols and their suitability for near-real time field investigation.

After nucleic acid extraction from a bacterial mock community, we compared a) full length reads vs. 16S analysis and b) Oxford Nanopore Technologies' high-accuracy vs. fast base calling algorithm. We applied the workflow to a bioaerosol sample generated in an agricultural environment.

1. Introduction

Bioaerosols are ubiquitous and very complex airborne particles of biological origin. They contain parts of and complete microorganisms, such as bacteria and fungi, or viruses. These may be harmful to the health of human beings. Subsequently, they are relevant in the context of environmental and occupational health, but also within the scope of biosafety, biosecurity and environmental protection.

The accompanying diagnostics are to date limited with respect to quantitative and qualitative approaches and thus the development of new methodologies enabling the molecular surveillance of the microbial load of bioaerosols is necessary. This work contributes to the field by evaluating nanopore sequencing in aforementioned context.

2. Methods

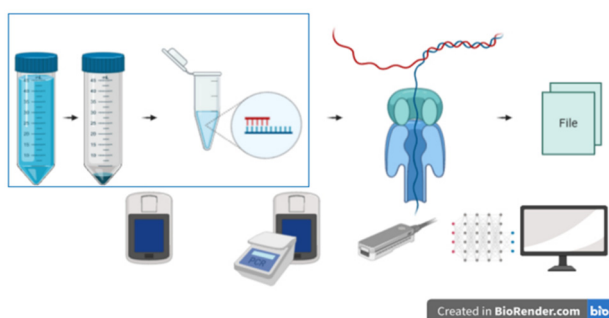


Figure 1: Bioaerosol sample analysis workflow.

Previously investigated samples [1] and a bioaerosol sample of agricultural origin (L-65) were subjected to comparative 16S-DNA sequencing for 72 hours using the 16S Barcoding Kit (SQK-RAB204, ONT, Oxford, United Kingdom) on a single flowcell (FLO-MIN106D, ONT, Oxford, United Kingdom). Following PCR failure and troubleshooting, protocols were modified to use Q5[®]

High Fidelity 2X Master Mix (NEB, Frankfurt a. M., Germany) and molecular grade water instead of proprietary RAP buffer. Custom 16S databases/indices were built for classification with Kraken2 and Centrifuge from 16S RefSeq Nucleotide sequence records (NCBI).

3. Results and Discussion

Sequencing pooled and barcoded samples generated 16.01 million reads equivalent to 22.1 giga bases.

| | FAST | HACC |
|--------------------|---------|---------|
| Length mean [bp] | 1,270.6 | 1,263.6 |
| Length median [bp] | 1,552.0 | 1,562.0 |
| Q mean | 9.0 | 10.7 |
| Q median | 9.5 | 11.2 |

Table 1: Read metrics comparison for fast (FAST) and high accuracy (HACC) base calling.

Targeted 16S sequencing using the commercial ONT kit requires only <10 ng genomic DNA input as opposed to our previously established full genome sequencing approach [1], which is a clear advantage for the highly limited DNA quantities extractable from bioaerosol sample material.

Analysis of the four benchmarking samples revealed inferior performance with only 4 out of 12 (hetero-genomic conc.) and 6/7 out of 12 (equi-genomic conc.) true positive bacterial species on genus level with a lower limit of detection of $5 \cdot 10^6$ to $5 \cdot 10^7$ genomes. In contrast, with the previously established whole genome pipeline [1] we were able to identify 12 out of 12 true positive bacterial species (equi-genomic conc.) and 5 (9) true positive bacterial species (hetero-genomic conc.) down to $2,5 \cdot 10^6$ ($2,5 \cdot 10^4$) genomes. Additionally, quantitative differences, possibly due to PCR bias, could be observed.

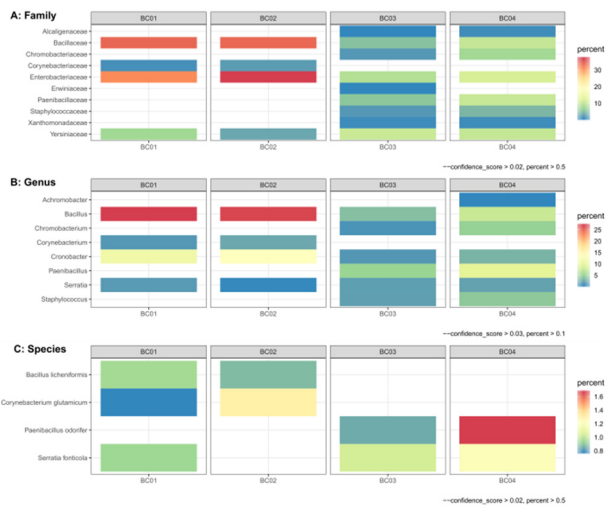


Figure 2: Kraken2 analysis of the four 16S benchmark samples on A: Family, B: Genus and C: Species level.

At last, sample L-65 was subjected to cursory analysis and representatives of *Faecalibacterium*, *Lactobacillus*, *Megamonas*, *Romboutsia* and *Staphylococcus* could be reliably identified, of which some have possible importance in environmental health. These findings match previous investigation (Mietke, personal communication).

4. Conclusion

This work compares our previously established whole genome analysis [1] workflow for investigating the genomic composition of bioaerosol samples to a commercially available 16S kit. In order to successfully employ the commercially available kit, several modifications had to be made. While the lesser input requirements for the 16S kit are highly beneficial for analysis of bioaerosol samples, this comes at the cost of inferior resolution.

Acknowledgements

RL acknowledges funding through the European Social Fund (ESF) PhD Scholarship, Grant #100316182. RW received funding from the Saxon State Ministry of Science and Art and the BMBF-funded Saxony5 Initiative. The funders had no role in the design of the study, the collection, analysis and interpretation of data and the writing of the manuscript.

References

- [1] Leidenfrost, R.M., Pöther, DC., Jäckel, U. et al. Benchmarking the MinION: Evaluating long reads for microbial profiling. *Sci Rep* 10, 5125 (2020). <https://doi.org/10.1038/s41598-020-61989-x>

Learning by doing: Integrating bioprocess engineering and undergraduate education

Robert Maximilian Leidenfrost¹, René Kretschmer¹

¹Mittweida University of Applied Sciences, Technikumplatz 17, 09648 Mittweida

Microorganisms are biotechnologically applicable in medicine, agriculture and industry. More precisely, microbial utility includes e.g. the ancient usage in brewery, the production of high value compounds or wastewater treatment, sometimes under very extreme conditions. While the original principles remain the same, a high degree of optimization and, e.g., modern computer-guided control and monitoring facilitate new or improved bioprocesses.

Since 2014, students from Mittweida University of Applied Sciences volunteer to participate in the annual 99€-Bioreactor Challenge hosted by the Netzwerk Bioverfahrenstechnik Dresden e.V.. Each year, there is a new and different challenge. Integrating and applying their biological and engineering knowledge, the students construct a bioreactor to face off against competing teams. Here we present a review of the past five years of hands-on engineering conducted by biotechnology undergraduates.

1. Background

Bioprocess engineering applies principles of engineering to biological systems. Highly interdisciplinary, it integrates disciplines such as mechanical and electrical engineering or computer sciences with the field of biology, including, but not limited to, chemical processes and, e.g., physiological mechanisms. Among its purpose is the design and development of equipment and processes used, e.g., for the manufacturing of food, feed, high value compounds or wastewater treatment.

The design, development and operation of bioreactors is a prominent subject of bioprocess engineering. Here we present successfully competing bioreactors developed by Mittweida University of Applied Sciences Biotechnology undergraduates for a nationwide bioreactor challenge held annually by the *Netzwerk Bioverfahrenstechnik Dresden e.V.*

Universal challenge key features include the cost limit of 99 Euros (provided by sponsors), and the requirement for greater than 24 hour of independent operation.

2. Challenge timeline

| | Aim | Win criteria |
|------|---|---|
| 2015 | Phototropic cultivation of <i>Rhodobacter sphaeroides</i> | Highest absorption >800nm and highest biomass |
| 2016 | Fed-batch cultivation of <i>Xanthophyllomyces dendrorhous</i> | Highest absorption at 480nm |
| 2017 | Wastewater treatment with <i>Trametes versicolor</i> | Highest waste removal and highest Laccase enzyme activity |
| 2018 | Carotenoid production with <i>Anoxybacillus flavithermus</i> | Highest pigment production and highest biomass |
| 2019 | High cell density fermentation of <i>Vibrio natriegens</i> | Highest biomass yield and highest biomass |

Table 1: Aims and win criteria for the 99€ bioreactor challenge timeline covering a wide range of microbes and bioprocess engineering applications.

3. Made in Mittweida

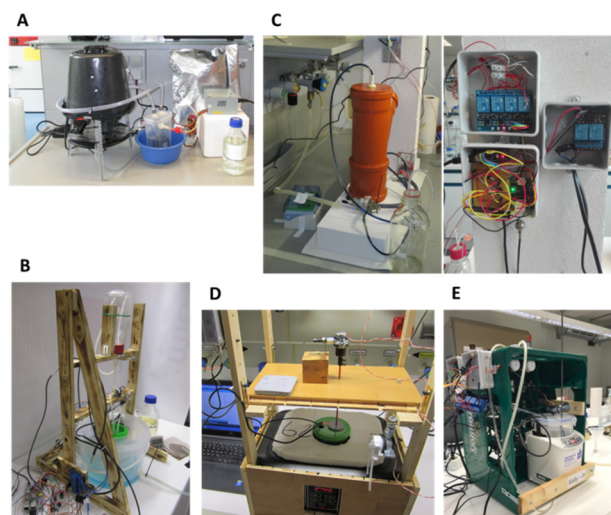


Figure 1: Bioreactors successfully competing. A: *Eimerreaktor* (2015), B: *Pink Candy Jar 4.0* (2016), C: *Funginator* (2017) with ingress protection for wastewater environment use, D: *Das Gelbe vom Ei* (2018) and E: *Salty-Jo* (2019)

4. Conclusion

Volunteering biotechnology undergraduates are challenged with failsafe engineering on scarce funding, with especially the former being of paramount importance.

Despite a lack of formal education in, e.g., embedded systems engineering, they gain a deeper insight into modern computer guided control and monitoring of bioprocesses on a technical level.

Acknowledgements

The authors kindly thank the Netzwerk Bioverfahrenstechnik Dresden e.V. for hosting the 99€-bioreactor challenge and all sponsors for financial support. We especially acknowledge and do thank our committed and successful students.

Into the genome of the German Sheep Poodle

Robert Maximilian Leidenfrost, Nils Schön, Rebecca Prause & Röbbbe Wünschiers
Mittweida University of Applied Sciences, Technikumplatz 17, 09648 Mittweida

The German Sheep Poodle is a breed of old German herding dogs. They are mid-sized, friendly and intelligent working dogs. Currently, breeders undertake efforts to preserve the German Sheep Poodle.

No exception to other dog breeds or even mammals, the breed is affected by cryptorchidism, which is the absence of one or both testes from the scrotum in males. The disease is a common defect of the male genital tract, and associated with several complications, such as, e.g., reduced fertility, or an increased risk for tumors. It affects up to 10% of all male puppies. A hereditary basis for cryptorchidism in dogs is suggested and heritability of cryptorchidism may be influenced through genetic selection. However, studies on canine cryptorchidism and its underlying causes remain scarce. We contribute to the understanding of canine cryptorchidism by taking a look into German Sheep Poodle genetics and the genome.

1. Background

Cryptorchidism describes a disease, in which one or both testes do not descend into the scrotum properly [1]. With a prevalence of up to 10% in male pups, depending on race and inbreeding [2], cryptorchidism is one of the most common birth defects of the male genital tract [3] and can also affect other mammals, including humans [4], making it a widespread problem. Amongst other disorders, it is associated with a 13.6 times higher risk for the development of malignant neoplasms and is correlated with an impaired spermatogenesis [5]. Despite its associated health risks and accompanying economic damage, resulting from surgery and losses in breeding [1], studies on canine cryptorchidism and its causes are relatively rare.

The aim of our study is to contribute to the understanding of canine and eventually human cryptorchidism by taking a look into German Sheep Poodle genetics and its genome.

2. Multifold approach for integrated analysis

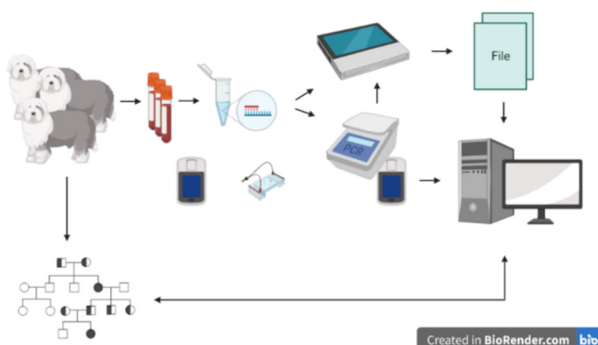


Figure 1: Methodology for investigating sheep poodle genome.

Working towards an integrated analysis of cryptorchidism in German Sheep Poodle is based on a multifold approach including 1) the creation and maintenance of an in-house relational database with known cryptorchidism-associated genetic variants including suggestions for standardization in reporting [6, 7, 8, 9], 2) the creation of a biorepository containing a wide-ranging

collection of German Sheep Poodle blood specimen (long-term storage) and corresponding hereditary data for pedigree analysis, 3) establishment of candidate gene amplification for sequencing [7, 8, 9] and 4) sequencing of candidate gene amplicons [8, 9] and whole genome sequencing of German Sheep Poodle employing second and third-generation sequencing technology.

3. Methods

A database containing all collected dog samples is implemented using standard office software solution and cross-referenced with known pedigree information (*Institut für ganzheitliche Schafpudelnzucht*, personal communication).

Genomic DNA is extracted from all stored blood samples using either MagSi-DNA (magtivio b.v., The Netherlands), MasterPure (EpiCentre, United States), Monarch gDNA (NEB, Germany), Monarch HMW (NEB, Germany) NucleoSpin (Macherey-Nagel Germany, Germany) extraction kits according to respective protocols. DNA quality is assessed photometrically (NanoVue, General Electric, United States), fluorometrically (Qubit, Invitrogen, Germany) and by agarose gel electrophoresis.

Cursory whole genome sequencing of a male German Sheep Poodle (library LSK-SQK108/EXP-NBD103, FLO-MIN106, ONT, United Kingdom, according to respective protocol) is performed as proof of concept.

4. Results and Discussion

Following the successful establishment of our in-house database [7-9] and candidate gene amplicon sequencing [7-9], our current focus is on expanding experimental data. Collected samples currently include 143 dogs, of which 78 are male, 64 are female and 1 is of unknown sex. Three quarters of the male dogs in our collection are non-cryptorchidic, while 14 animals are affected (Table 1). To our knowledge, we possess the largest collection of samples from German Sheep Poodle to date, including pedigree information (Figure 1).

| State | n |
|-------------------|----|
| No cryptorchidism | 58 |
| Both testes | 2 |
| Left testis | 5 |
| Right testis | 4 |
| Unknown | 9 |

Table 1: Summary statistics of cryptorchidism state in male dogs available from our repository.

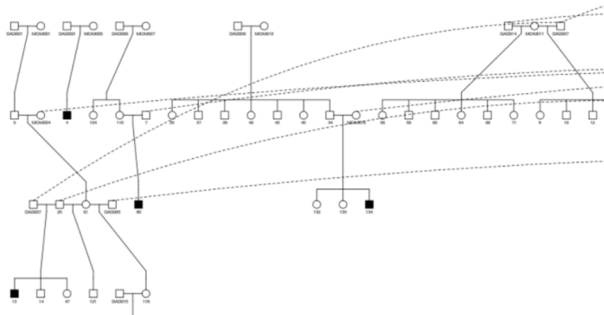


Figure 1: Pedigree information from the German Sheep Poodle sample collection (partial). Cryptorchidic dogs are displayed in black.

Reliable data generated by next- and third generation sequencing requires high quality and pure nucleic acid templates. Subsequently, we looked at DNA quantity, purity and quality parameters for 121 samples extracted with a variety of kits from samples collected since 2017 for assessment.

| Kit | Mean conc. NanoVue [ng/µl] | Mean conc. Qubit [ng/µl] | A260/230 | A260/280 | n |
|--------------|----------------------------|--------------------------|----------|----------|----|
| MagSi-DNA | 36.0 | 23.5 | 1.67 | 1.83 | 12 |
| MasterPure | 194.0 | 56.3 | 1.59 | 1.72 | 84 |
| Monarch gDNA | 57.2 | 39.4 | 2.02 | 1.78 | 10 |
| Monarch HMW | 125.0 | 65.9 | 2.07 | 1.80 | 5 |
| NucleoSpin | 56.7 | 34.5 | 2.18 | 1.77 | 10 |

Table 2: Overview of DNA yield and quality parameters for 121 samples extracted with a variety of commercial DNA extraction kits.

All kits were able to generate sufficient quantities of high-quality genomic DNA for our purposes (Table 2). The predominantly used MasterPure Complete DNA Extraction kit exhibited highest median yields, but was most laborious.

An initial insight into the genome of the German Sheep Poodle is given by 192,891 nanopore long sequence reads, equivalent to 1.8 Gb (Mean read length 9,490 bp, N50 17,162 bp; as assessed by NanoOK; longest aligned read ~91k bp, ~95.4% of all reads with alignments to CanFam3.0 reference). This implies a mean coverage of ~0.7 as compared towards the dog reference genome.

5. Conclusion & Outlook

We are making progress in elucidating German Sheep Poodle genetics and its genome. Increasing our knowledge on German Sheep Poodle genetics through expanding our databases is accompanied by gathering further experimental evidence regarding previously established candidate genes related to cryptorchidism. A high-quality draft genome of the German Sheep Poodle is in preparation.

Acknowledgements

RL acknowledges funding through the European Social Fund (ESF) PhD Scholarship, Grant #100316182. RW received funding from the Saxon State Ministry of Science and Art and the BMBF-funded Saxony5 Initiative. The funders had no role in the design of the study, the collection, analysis, interpretation of data and writing of the manuscript. The authors thank the *Institut für ganzheitliche Schafpudelmzucht*, GENEWIZ GmbH, Leipzig for NGS support, and especially all sheep poodles for financial support and dog blood samples.

References

- [1] Mattos, M. R. F., Simões-Mattos, L., & Domingues, S. F. S. (2000). Cryptorchidism in dog. *Ciência Animal*, 10(1), 61-70.
- [2] Birchard S.J. & Nappier, M. (2008). Cryptorchidism. *Compendium (Yardley, PA)*, 30(6), 325-36
- [3] Wood, H. M., & Elder, J. S. (2009). Cryptorchidism and testicular cancer: separating fact from fiction. *The Journal of urology*, 181(2), 452-461.
- [4] Mahmud, M. A., Onu, J. E., Shehu, S. A., Umar, M. A., Bello, A., & Danmaigoro, A. (2015). Cryptorchidism in mammals—a review. *Global Journal of Animal Scientific Research*, 3, 128-135.
- [5] Virtanen, H. E., Bjerknes, R., Cortes, D., Jørgensen, N., Rajpert-De Meyts, E., Thorsson, A. V., & Main, K. M. (2007). Cryptorchidism: classification, prevalence and long-term consequences. *Acta Paediatrica*, 96(5), 611-616.
- [6] Urh, K., & Kunej, T. (2016). Molecular mechanisms of cryptorchidism development: Update of the database, disease comorbidity, and initiative for standardization of reporting in scientific literature. *Andrology*, 4(5), 894-902.
- [7] Prause, R., Leidenfrost, R., & Wünschiers, R. (2019). Towards an integrated analysis of cryptorchidism in German Sheep Poodle. 20. Nachwuchswissenschaftlerkonferenz, Conference
- [8] Prause, R., Leidenfrost, R., & Wünschiers, R. (2019). Nanoporesequenzierung von Kryptorchismus-assoziierten Genen zur Identifizierung von genetischen Variationen in Schafpudeln. Tag der Biodiversität, Workshop
- [9] Prause, R. (2019). Nanopore sequencing of cryptorchidism associated genes for identification of genetic variants in German Sheep Poodle., Thesis

Eine Reihe von Publikationen der 26. IWKM ist im Rahmen von durch EU-, Bund,- oder dem Freistaat Sachsen geförderten Projekte entstanden. Wir danken insbesondere folgenden Förderern und Projektträgern hinsichtlich der Mitfinanzierung der Vorhaben:



EINE GEMEINSAME INITIATIVE VON



Bundesministerium
für Bildung
und Forschung



Gemeinsame
Wissenschaftskonferenz
GWK



Bundesministerium
für Wirtschaft
und Technologie



STAATSMINISTERIUM
FÜR WISSENSCHAFT
KULTUR UND TOURISMUS



Freistaat
SACHSEN

STAATSMINISTERIUM
FÜR WIRTSCHAFT
ARBEIT UND VERKEHR



Freistaat
SACHSEN



Die Hochschule wird mitfinanziert durch Steuermittel auf der Grundlage des vom Sächsischen Landtag beschlossenen Haushaltes.

