Wissenschaftliche Vertiefungskonferenz 2019

Tagungsband

Herausgeber:
Uwe Kloos, Hochschule Reutlingen
Natividad Martinez, Hochschule Reutlingen
Gabriela Tullius, Hochschule Reutlingen

Hochschule Reutlingen
Reutlingen University

INF
Informatik

WVK 2019

Tagungsband

Wissenschaftliche
Vertiefungskonferenz

27. November 2019, Reutlingen

Impressum

Anschrift:

Hochschule Reutlingen
Reutlingen University
Fakultät Informatik
Human-Centered Computing
Alteburgstraße 150
D-72762 Reutlingen

Telefon: +49 7121 / 271-4002
Telefax: +49 7121 / 271-4042

E-Mail: wvk@reutlingen-university.de
Website: wvk.reutlingen-university.de

Organisationskomitee:

Prof. Dr. rer. nat. Uwe Kloos, Hochschule Reutlingen
Prof. Dr.-Ing habil. Natividad Martinez, Hochschule Reutlingen
Prof. Dr. rer. nat. Gabriela Tullius, Hochschule Reutlingen

Hannah Büchner
Claushinrich Himstedt
Alexander Kahabka
Georgia König
Maximilian Malik
Marek Mauser

ISBN: **978-3-00-064236-4**

Copyright: © Hochschule Reutlingen, Reutlingen 2019
Herstellung und Verlag: Hochschule Reutlingen

Hochschule Reutlingen
Reutlingen University

Vorwort

Sehr geehrte Leserinnen und Leser,

wie aufregend! An unserer Hochschule Reutlingen organisieren die Studierenden des Masterstudiengangs Human-Centered Computing (HUC) eine eigene wissenschaftliche Konferenz! Diese durchaus sehr anspruchsvolle Tätigkeit hat zwei wichtige Effekte: Zum einen dient sie der Vorbereitung der Masterarbeit und zum anderen hilft sie zu begreifen, wie Wissenschaft funktioniert! Daher bin ich über die Ausrichtung der Wissenschaftlichen Vertiefungskonferenz 2019 sehr erfreut und sehr zufrieden mit der Veröffentlichung der Beiträge in diesem Tagungsband. Was wäre eine Konferenz ohne Inhalte? Unsere Studierenden haben zahlreiche Themen bearbeitet, Literatur recherchiert und komplexe Sachverhalte verständlich aufgearbeitet und dargestellt. Zu den diesjährigen Themen gehören Beiträge aus dem Bereich kognitiver Systeme, bei denen es darum geht, Marker von Gegenständen, Muskelaktivität oder Objekte in einer schwierigen dreidimensionalen Szenerie maschinell zu erkennen. Eine Arbeit befasst sich mit der Modulation von Radarwellen, um sie für Gestenerkennen einsetzen zu können. Einige Beiträge gehören in den Bereich der kollaborativen Systeme, bei denen Mensch und Maschine eng zusammenarbeiten, wie z.B. bei der Positionsbestimmung eines Roboters im PatientInnenkontext oder der Verbesserung von NutzerInnen-Oberflächen. Um große Datenmengen geht es bei der Anwendung von künstlicher Intelligenz im Mittelstand für Data Analytics und im Kontext von Advanced Analytics. Und nicht zuletzt gibt es zwei spannende Softwaretechnik-Themen: Eine Ausarbeitung beschäftigt sich mit der Migration eines Monolithen zu Microservices, die kleinste fachliche Bausteine eines Systems darstellen. Eine weitere Ausarbeitung ist die Konzeption eines DevOps-Reifegradmodells, das bewertet, wie gut ein DevOps-Entwicklungsteam arbeitet. Im Kontext eines Krankenhausinformationssystems wird in einer Ausarbeitung die Spezifikation von veralteten Komponenten erstellt, um diese störungsfrei ablösen zu können. Alles sind schwierige Fragestellungen.

Zunächst möchte ich allen AutorInnen danken, die mit Ihren Ausarbeitungen und Beiträgen die Konferenz und den Tagungsband mit wertvollem Inhalt gefüllt haben. Mein größter Dank gilt den MitgliederInnen des Organisationskomitees, insbesondere der Organisationsleitung Frau Hannah Büchner und Herrn Maximilian Malik für die professionelle Vorbereitung und Durchführung der Konferenz. So geht Wissenschaft! Auch ein herzliches Dankeschön an meine KollegInnen, den ProfessorInnen des HUC-Masterstudiengangs, die mit Ihrer Betreuung der Studierenden die Beiträge ermöglicht haben. Den LeserInnen wünsche ich viel Spaß bei der Betrachtung der Beiträge und unseren Studierenden weiterhin viel Erfolg und alles Gute bei Ihrem Studium und der anschließenden beruflichen Laufbahn.

Es grüßt Sie herzlich aus Reutlingen

Prof. Dr. rer. nat.
Christian Kücherer

Inhaltsverzeichnis

Object Recognition innerhalb eines kollaborativen Systems 1
Automatische Analyse von Gesichtsmuskelaktivität mit Hilfe von 2D- und 3D-Daten 12
Trennung der Dekodierung und Detektion von Markern zur Positionsüberwachung eines Patientenpositionierroboters in Echtzeit 22
Perspektivische Transformation zur Kalibrierung von Verkehrsüberwachungskameras 32
Vom Monolith zu Microservices - Best Practices und Herausforderungen am Beispiel des Webkonferenzsystems Accelerator 39
Analyse einer Dokumentenverarbeitung mit Microsoft Office im Krankenhausinformationssystem 49
Konzeption und Durchführung einer Umfragereihe zur Erfassung von DevOps-Aspekten in Softwareentwicklungsprojekten 60
Application of Artificial Intelligence to optimize forecasting capability in procurement 69
Advanced Analytics: Analyse und Ablauf der Pre-IMplementation 81
Frequenzmodulation-Parametrisierung für Radarwellen zur Erkennung von Gesten 92
Model-View Design Patterns 102

Autorenverzeichnis **110**

Object Recognition innerhalb eines kollaborativen Systems

Panagiotis Taxidis
Hochschule Reutlingen
Panagiotis.Taxidis@Student.Reutlingen-University.de

Abstract
In dieser Arbeit wird ein Ansatz vorgestellt, der die Verwendung von einer Object Recognition Software innerhalb eines kollaborativen Systems ermöglicht. Es werden verwandte Arbeiten betrachtet, die sich ebenfalls mit dem Thema Object Recognition beschäftigen und dabei auch eigene Lösungen vorstellen. Es wird das genaue Ziel der Object Recognition innerhalb des kollaborativen Systems beschrieben und daraus Konzept und Use Cases erstellt. Für die Object Recognition werden Anforderungen gesetzt und bereits existierende Projekte und Libraries analysiert. Die Projekte/Libraries, die diesen Anforderungen entsprechen, werden dann bei der Umsetzung verwendet.

Schlüsselwörter
Kollaboratives System, Brainstorming Session, Handwritten Text Recognition, Object Recognition, OpenCV, Tensorflow

CR-Kategorien

Human-centered-computing~Collaborative interaction

Computing-methodologies~Object recognition

1 Einleitung
Diese Arbeit beschäftigt sich mit der Verwendung von Object Recognition Software innerhalb eines kollaborativen Systems und basiert auf der Publikation von Merk et al. [1]. In [1] geht es um die Unterstützung der Ideenfindung in einer Co-Located Brainstorming-Session innerhalb einer Gruppe von drei bis fünf Leuten. Die Brainstorming-Session findet an einem Tisch mit einem Durchmesser von 80 cm statt. Dabei hat jeder Teilnehmer in der Session genügend Platz für sich, ist jedoch nicht zu weit von den anderen Teilnehmern entfernt, wodurch eine Gruppenatmosphäre besteht. Über dem Tisch befindet sich ein Projektor, der für jeden Teilnehmer die Position und die Größe der Arbeitsfläche projiziert (siehe Abbildung 1).

Betreuer/in Hochschule:

Prof. Dr. rer. nat. Gabriela Tullius
Hochschule Reutlingen
Gabriela.Tullius@Reutlingen-University.de

Matthias Merk, M. Sc.
Hochschule Reutlingen
Matthias.Merk@Reutlingen-University.de

Wissenschaftliche Vertiefungskonferenz
21. November 2019, Hochschule Reutlingen
Copyright 2019 Panagiotis Taxidis

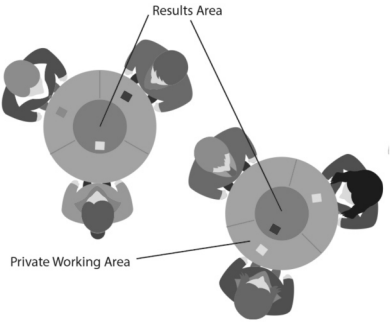


Abbildung 1: Projizierte Arbeitsflächen [1]

Der private Arbeitsbereich ist für jeden Einzelnen für seine privaten Notizen und Informationen reserviert, welche für eine Diskussion mit der Gruppe verwendet werden können, um verschieden Ideen zu kreieren. In der Mitte über den Arbeitsflächen der Teilnehmer, wird der Ergebnisbereich projiziert. Dort landen alle gemeinsamen Ideen der Gruppe. [1] Zukünftig wäre es auch möglich zu erfassen, welcher Teilnehmer einer Session am Sprechen ist, um LEDs am Arbeitsbereich der sprechenden Person anzusteuern. [1]

2 Ziel der Arbeit

Ziel dieser Arbeit ist es eine Co-Located Computer Supported Cooperative Work (CSCW) Umgebung zu erweitern. Es soll das kollaborative System aus Kapitel 1 (Einleitung) mit Object Recognition kombiniert werden, um die Ideenfindung innerhalb einer Brainstorming Session zu verbessern, indem die Ideen/Informationen in einer Datenbank gespeichert werden und wieder aufrufbar sind. Für die genaue Zielbeschreibung wurde ein Use Case erstellt (siehe Abbildung 2).

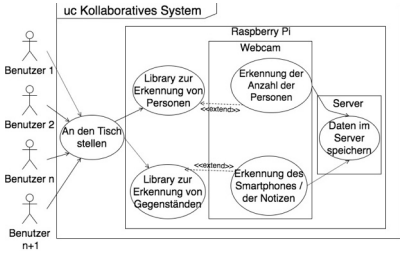


Abbildung 2: Use Case des Ziels der Object Recognition

Mit Hilfe der Object Recognition, soll es in dem kollaborativen System möglich sein Personen zu erkennen, die sich an den Tisch gestellt haben. Bei der Auswahl der Software zur Anwendung der Object Recognition, stellten die Kriterien die Genauigkeit der Erkennung und die Möglichkeit des Trainings mit Hilfe eines eigenen Datensatzes dar. Jede erkannte Person soll eine eigene ID bekommen, die in einer Datenbank gespeichert wird. Wenn es bei der Erkennung von Personen zu Verwechselungen mit der ID kommt, sollte auch die Erkennung der Smartphones über NFC möglich sein, sodass die ID einer Person eindeutig definiert werden kann. Über die Anzahl der erkannten Personen, sollen auch die projizierten Arbeitsflächen des Tisches angepasst werden. Sobald eine Person ein Objekt auf den Tisch gelegt hat, soll das Objekt automatisch erkannt und ebenfalls in der Datenbank mit einer ID gespeichert werden. Dabei können die erkannten Objekte ein Smartphone, Notizen oder andere Gegenstände sein. Bei Notizen soll es auch die Möglichkeit geben, die Handschrift zu erkennen. Es soll erkannt werden welches Objekt von welcher Person auf den Tisch gelegt wurde, damit über die ID der Person und die ID des Objekts eine Assoziation erstellt werden kann. Da jede Person ihren eigenen Bereich hat, sollen diese auch unabhängig voneinander bei Veränderungen beobachtet und erkannt werden.

Die Brainstorming Sessions finden nur für eine bestimmte Zeit statt, daher soll diese Arbeit es ermöglichen, alle Daten/Information älterer Sessions wieder in neuen Sessions aufrufen zu können. So können Personen, die über mehrere Tage an verschiedenen Brainstorming Sessions teilnehmen, immer wieder ihre mit der eigenen ID assoziierten Daten aufrufen, um direkt vom letzten Stand weiter zu arbeiten und neue Ideen zu entwickeln.

Mit der Verwendung von Object Recognition und der automatischen Dokumentation des Brainstormings innerhalb des kollaborativen Systems aus der Einleitung, kann ebenfalls eine Studie durchgeführt werden, um zu untersuchen welche Erkenntnisse gewonnen werden können.

3 Verwandte Arbeiten

Dieses Kapitel befasst sich mit wissenschaftlichen Arbeiten, die sich wie in Kapitel 2 mit Object- und Handwritting Recognition beschäftigen.

3.1 Real Time Object Recognition

In der Publikation von Chandan et al. [2] wird ein Single Shot Detector (SSD) basierter Algorithmus zur Erkennung von Objekten in einer Python-Umgebung mit OpenCV implementiert. [2]

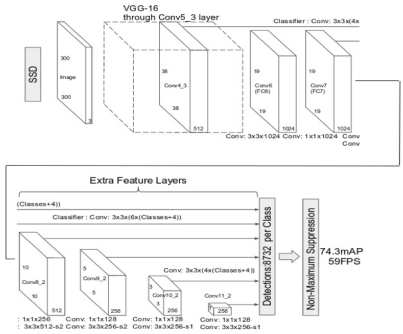


Abbildung 3: SSD Architektur[1]

Der SSD Algorithmus basiert auf der VGG-16-Architektur (siehe Abbildung 3) [1]. Hier wird eine Reihe von Standardboxen so eingestellt, dass sie faltungstechnisch (convolutional) über mehrere Feature-Maps geführt wird. Wenn ein Objekt innerhalb einer Box in einem Objektklassifizierer erkannt wurde, wird ein Wert generiert. Die Objektform wird so angepasst, dass sie der entsprechenden Box entspricht. Für jede Box wird der Formversatz und die Zuversichtlichkeit (confidence level) des erkannten Objekts vorhergesagt. Die Zuversichtlichkeit sagt aus, wie sicher sich ein System mit seiner Vorhersage des erkannten Objektes ist. [2]

Bei der Implementierung werden entsprechend Chandan et al. [2], die folgenden Methoden eingesetzt:

Frameunterscheidung (Frame differencing) - Die Frames werden in regelmäßigen Abständen von einer Kamera aufgenommen. Die Differenz wird aus den aufeinanderfolgenden Frames geschätzt.

Optischer Ablauf (Optical Flow) – Hier wird der Optische Ablauf von Bewegungsinformationen geschätzt und berechnet. Als nächstes wird für die Rauschunterdrückung ein selbstadaptiver

[1] https://towardsdatascience.com/learning-note-single-shot-multibox-detector-with-pytorch-part-1-38185e84bd79

Algorithmus verwendet. Der Algorithmus beinhaltet eine weitgehende Anpassung an die Anzahl und Größe der Objekte und hilft, zeitaufwändige und komplizierte Vorverarbeitungsmethoden (pre-processing) zu vermeiden.

Hintergrund-Subtraktion (Background substraction) – Alle Objekte innerhalb eines Frames werden lokalisiert. Hier wird der Vordergrund, der diese Objekte enthält, vom Hintergrund getrennt.

Sobald der Algorithmus in OpenCV und Python implementiert wurde, gab es mehrere Testversuche (siehe Abbildung 4).

Abbildung 4: SSD und OpenCV Object Detection [2]

Die Ergebnisse von Abbildung 4 haben gezeigt, dass der SSD Algorithmus in OpenCV hohe Zuversichlichkeitswerte erbracht hat. Dies bedeutet, dass die Objekte richtig erkannt wurden.

3.2 Handwritten Character Recognition

In der Publikation von Vaidya et al. [3] wird ein Verfahren zur Offline-Handschriftzeichenerkennung, mit Hilfe von Deep Neural Networks vorgestellt. Es wird ein auf der Bildsegmentierung basierendes System zur Handschrifterkennung entwickelt. Für die Bildverarbeitung, wird OpenCV verwendet und für das Training eines neuronalen Netzwerks, wird Tensorflow verwendet. Das System wurde mit der Programmiersprache Python entwickelt. [3]

Es gibt zwei Arten von Handschrifterkennungssysteme, Online und Offline. In einem Online-Handschrifterkennungssystem, wird die Handschrift des Benutzers beim Schreiben erkannt. Die Informationen wie die Reihenfolge, in der der Benutzer die Striche gemacht hat, sind ebenfalls verfügbar. Im Offline-Handschrifterkennungssystem, ist jedoch die Handschrift des Benutzers nur als Bild verfügbar. Was die Handschrifterkennung erschwert, ist die individuelle Handschrift eines jeden Menschen. Ebenfalls sorgen auch Großbuchstaben, Kleinbuchstaben, Ziffern und Sonderzeichen für eine erschwerte Handschrifterkennung. Daher ist ein großer Datensatz erforderlich, um ein nahezu genaues neuronales Netzwerkmodell zu trainieren. [3]

Eine Android Applikation, wird als Frontend verwendet, dass ein Bild mit einer Handschrift an einem Server sendet und den erkannten Text vom Server wieder empfängt. [3]

Der Server ist das Backend des Systems und ist für zwei wichtige Aufgaben zuständig. Der Server hosted das vortrainierte neuronale Netzwerkmodel, um Vorhersagen zu liefern und führt Bildverarbeitungsoperationen an dem empfangenen Bild aus, um den handschriftlichen Text zu erkennen. Auf dem Server existiert das mit Tensorflow trainierte neuronale Netzwerkmodell und das Python-Skript, das mit einer OpenCV-Bibliothek ausgestattet ist. [3]

An der Verarbeitung der empfangenen Bilder sind entsprechend Vaidya et al. [3], folgende Schritte beteiligt:

Schritt 1 - Vorverarbeitung (Pre-processing) - Hier wird das Rauschen aus dem Bild durch eine Medianfilterung entfernt.

Schritt 2 - Konvertierung in Graustufen (Conversion to Gray-Scale) – Als nächstes wird das Bild in Graustufen umgewandelt. Die Umwandlung in Graustufen ist notwendig, da verschiedene Personen Stifte

mit unterschiedlichen Farben und unterschiedlichen Intensitäten verwenden. Die Arbeit an Graustufenbildern reduziert auch die Gesamtkomplexität des Systems.

Schritt 3 - Schwellenwert (Thresholding) - Sobald ein Bild in Graustufen umgewandelt wird, ist der handschriftliche Text dunkler als sein Hintergrund. Mit Hilfe der Schwellenwerte kann dadurch der dunklere Bereich (Handschrift) des Bildes, von dem helleren Bereich (Hintergrund) getrennt werden.

Schritt 4 - Bildsegmentierung (Image Segmentation) – Der Text von einer Person wird in Zeilenform geschrieben, dadurch wird nach dem Threshold der Text in einzelne Zeilen segmentiert. Jede einzelne Zeile, wird dann in einzelne Wörter unterteilt. Schließlich wird jedes Wort in einzelne Zeichen unterteilt, wobei zunächst das geschätzte Bild invertiert wird, so dass der Hintergrund in den Vordergrund tritt und umgekehrt. Beim Scannen, wird die Summe der Pixel in jeder Bildzeile berechnet. Die Summe der Pixel ist Null, wenn alle Pixel einer bestimmten Zeile schwarz sind. Die Summe ist ungleich Null, wenn weiße Pixel in einer Reihe vorhanden sind. Dadurch werden die invertierten Zeichen erkannt, indem die weißen Pixel gezählt werden. Die einzelnen Zeichen werden an das vortrainierte neuronale Netzwerkmodell übergeben, um Vorhersagen zu machen. Auf diese Weise wird der endgültige vorhergesagte Text, als Ergebnis an den Benutzer zurückgesendet.

Anhand dieser Schritte, ist das System dieser Publikation lediglich in der Lage englischsprachigen, handschriftlichen Text zu erkennen. Momentan werden auch nur Buchstaben und Ziffern erkannt. Um kursiven Text und weitere Sprachen zu erkennen, müssten in Zukunft neue Modelle trainiert werden. [3]

4 Konzept und Anforderungen

Wie bereits in Kapitel 2 beschrieben, wird in diesem Kapitel für das kollaborative System, ein genaueres Konzept erstellt.

Neben dem Projektor, der über den Arbeitstisch angebracht ist und für die Projizierung der Arbeitsfläche zuständig ist, soll eine Webcam angebracht werden. Diese dient dazu, alle Personen und Objekte/Notizen am Arbeitsbereich der Brainstorming Session zu erkennen. Die Webcam, sollte dabei mindestens eine Auflösung von 3840x2160 (4K) Pixeln haben. Je höher die Auflösung ist, desto besser werden später die Objekte erkannt. Die Webcam ist an einem Raspberry Pi angeschlossen, auf dem die Software für die Objekterkennung läuft, dazu mehr in den Unterkapiteln.

4.1 Erkennung der Teilnehmer

Für die Implementierung der Object Recognition innerhalb des kollaborativen Systems, wurde der Use Case aus Kapitel 2 verwendet, welcher die Funktionsweise beschreibt (siehe Abbildung 2).

Im Fall des Use Cases von Abbildung 2, stellen sich drei bis fünf Personen an den Tisch. Hier läuft bereits die Object Recognition Software vom Raspberry Pi im Hintergrund. Auf dem Raspberry Pi selber, läuft ein Linux basiertes Betriebssystem, das alle erforderlichen Skripte für die Object Recognition ausführt. Die Webcam ist an dem Raspberry Pi angeschlossen und das Raspberry Pi selber, hat eine Verbindung zu einem Server/Datenbank.

Sobald die Anzahl der Personen, die sich an den Tisch gestellt haben, erkannt wurde, wird diese in der Datenbank gespeichert und dazu verwendet, die projizierten Arbeitsflächen mittels des Beamers am Tisch zu bestimmen.

4.2 Erkennung der Objekte

Sobald eine Brainstorming-Session angefangen hat, können die Teilnehmer verschieden Objekte/Notizen auf den Tisch legen, diese werden dann über die Webcam erkannt (siehe Abbildung 2). Die Erkennung kann dabei live stattfinden oder innerhalb von einem bestimmten Zeitintervall ausgeführt werden. Welche Variante sich dabei mehr lohnt, soll in Zukunft über eine Studie getestet werden.

Die erkannten Objekte werden alle in einer Bilddatei markiert und gespeichert. Die einzelnen markierten Objekte, werden als nächstes in eigene Bilddateien ausgeschnitten und in der Datenbank gespeichert.

4.3 Erkennung von Notizen

Für die Erkennung von Notizen wurde der Use Case von Kapitel 2 verfeinert (siehe Abbildung 5).

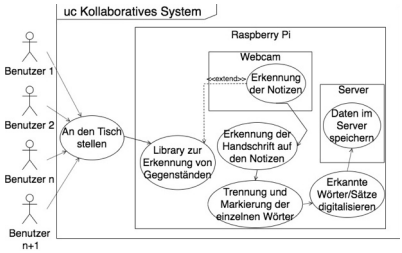


Abbildung 5: Use Case Erkennung von Notizen

Sobald es sich unter den erkannten Objekten um Notizen handelt, soll als nächstes die Handschrift auf den Notizen erkannt werden. Dabei werden die einzelnen Wörter markiert. Die erkannten Wörter/Sätze, sollen dabei als String in einer Datenbank abgelegt werden. Um die Notizen mit dem dazugehörigen Text zu finden, sollen beide in der Datenbank mit einander assoziiert werden.

4.4 Anforderungen

Anhand der vorher beschriebenen Use Cases, wurden Anforderungen (siehe Tabelle 1) erstellt. Die Anforderungen wurden vom Autor so definiert, dass innerhalb der erweiterten CSCW Umgebung, alle Schritte der zuvor beschriebenen Use Cases, ohne Einschränkungen erfüllt werden können.

Tabelle 1. Anforderungen

ID	Anforderungen
1	Open Source – Verwendete Projekte/Libraries sollen Open Source sein, damit diese frei von Lizenzeinschränkungen sind und der Code auf eigene Verhältnisse angepasst werden.
2	Erkennung von Personen, Gegenstände und Handschriften – Personen und Gegenstände sollten erkannt werden, damit eine Assoziation der beiden möglich ist. Bei Gegenständen wie Notizen, die Texte enthalten, sollte eine Handschrifterkennung stattfinden können.
3	Linux kompatibel – Da die Skripte für die Object Recognition auf dem Raspberry Pi ausgeführt werden sollen, sollten diese mit dem Linux Betriebssystem des Raspberry Pi kompatibel sein.
4	Eigene Daten trainieren – Die verwendeten Projekte/Libraries, sollten die Möglichkeit bieten eigene Daten zu trainieren, um die Genauigkeit der Erkennung von Objekten wie z.B. Notizen zu erhöhen.
5	Erkennung innerhalb eines Zeitintervalls oder Live – Das Skript für die Object Recognition, sollte die Möglichkeit bieten Objekte Live über eine Webcam zu erkennen oder nach einer bestimmten Zeit, über einen gespeicherten Frame.
6	Erkennung zwischen 0,5 - 5 Sekunden – Die Erkennung sollte

	nicht zu lange dauern, da sich in einer Brainstorming Session, der aktuelle Zustand schnell verändern kann.
7	Kommunikation an Server – Die verwendeten Projekte/Libraries, sollten die Kommunikation an einem Server nicht einschränken, sodass die erkannten Objekte digitalisiert und auf dem Server abgespeichert werden können.
8	Stabilität – Das Skript für die Object Recognition, sollte während einer Brainstorming Session im Hintergrund laufen, ohne dass es zum Absturz kommt, damit die erfolgreiche Erkennung der Objekte und Speicherung der Daten garantiert werden kann.
9	Einfach zu verwenden – Das Skript sollte leicht zu verwenden sein, indem es einmal gestartet wird und alle weiteren Aufgaben, automatisiert im Hintergrund abgearbeitet werden.

5 Stand der Technik

Um die Realisierung der Use Cases von Kapitel 4 zu ermöglichen, werden in diesem Kapitel die existierenden Techniken untersucht, welche sich ebenfalls mit Object Recognition befassen.

5.1 ImageAI

ImageAI ist eine Python-Bibliothek, die es Entwicklern, Forschern und Studenten ermöglicht, Anwendungen und Systeme mit eigenständigen Deep Learning und Computer Vision-Funktionen zu erstellen.[2]

ImageAI unterstützt eine Liste von verschiedenen Machine Learning Algorithmen für die Bildvorhersage, die benutzerdefinierte Bildvorhersage, die Objekterkennung, die Videoerkennung, die Videoobjektverfolgung und für das Training für Bildvorhersagen. [4]

ImageAI unterstützt auch die Objekterkennung, die Videoerkennung und die Objektverfolgung mit RetinaNet, YOLOv3 und TinyYOLOv3, die mit dem COCO-Datensatz trainiert wurden. [4]

Für die Verwendung von ImageAI ist die folgende Software notwendig. [4]

- Python 3.5.1 (oder neuere Versionen)
- Tensorflow 1.4.0 (oder neuere Versionen)
- OpenCV
- Keras 2.x

ImageAI bietet sehr leistungsstarke Methoden, um die Objekterkennung an Bildern durchzuführen und jedes Objekt aus dem Bild zu extrahieren. Die Objekterkennungsklasse bietet auch Optionen zur Anpassung der Leistung und der Echtzeit-Verarbeitung. [4]

Alle ImageAI-Implementierungen, können auf jedem Rechner mit schwächerer CPU Leistung (Dual Core, 2,3 GHz) ausgeführt werden. Die Verarbeitungsgeschwindigkeit für Operationen wie Bildvorhersagen, Objekterkennungen, usw. sind jedoch auf der CPU eher langsamer und nicht für Echtzeitanwendungen geeignet. Um Echtzeit Computer-Vision Operationen mit hoher Leistung durchzuführen, sollten Rechner mit einer Leistungsstarken GPU (z.B. GeForce RTX 2080 Ti) verwendet werden. Für die Computer Vision Operationen, verwendet ImageAI als Basis das Grundgerüst von Tensorflow. [4]

5.2 YOLO

You Only Look Once (YOLO), ist ein relativ neuer Ansatz zur Objekterkennung. Ältere Projekte für die Object Recognition verwenden Klassifizierer, um die Objekteerkennungen durchzuführen. In

[2] https://imageai.readthedocs.io/en/latest/

YOLO prognostiziert ein einzelnes neuronales Netzwerk Bounding Boxen und Klassenwahrscheinlichkeiten, direkt aus Vollbildern in einer Auswertung. Da die gesamte Erkennungspipeline ein einziges Netzwerk ist, kann sie End-to-End direkt auf die Erkennungsleistung optimiert werden. [5]

Das Basis Modell von YOLO, ist in der Lage Bilder in Echtzeit mit 45 Frames Per Second (FPS) zu verarbeiten. Eine kleinere Version des Netzwerks, Fast YOLO, kann 155 Bilder pro Sekunde verarbeiten. Im Vergleich zu anderen Detektionssystemen macht YOLO mehr Lokalisierungsfehler, ist aber weniger anfällig für Falscherkennungen im Hintergrund. [5]

YOLO trainiert auf Vollbildern und optimiert direkt die Erkennungsleistung. Dieses vereinheitlichte Modell, hat die folgenden Vorteile gegenüber herkömmlichen Methoden der Objekterkennung. [5]

YOLO ist extrem schnell, da die Rahmenerkennung als Regressionsproblem betrachtet wird. Dadurch wird auch keine komplexe Pipeline benötigt. Das neuronale Netzwerk, läuft zu Testzeiten auf einem neuen Bild, um die Erkennung vorherzusagen. YOLO betrachtet das ganze Bild, wenn es um Vorhersagen geht. YOLO sieht das gesamte Bild während des Trainings und der Test Zeit, so dass es implizit kontextuelle Informationen über die Klassen, sowie deren Aussehen kodiert. [5]

YOLO lernt durch allgemeine Darstellungen von Objekten. Da YOLO sehr allgemein ist, ist es weniger wahrscheinlich, dass es bei der Anwendung auf neue Domänen oder unerwartete Eingaben zusammenbricht. [5]

YOLO hat jedoch die Einschränkung, dass es nahlegende Objekte schwer erkennen kann, da eine Bounding Box immer nur für ein Objekt verwendet werden kann, ebenfalls sind Objekte, die eine ungewöhnliche Perspektive haben, schwer zu erkennen. [5]

5.3 Tensorflow

Tensorflow ist ein Framework für Deep Learning, um neuronale Netze zu modellieren. Mit großen Datensätzen und Rechnern, die in der Lage sind, diese zu verarbeiten und zu optimieren, haben neuronale Netze in den letzten Jahren an Popularität gewonnen. Sie sind in der Lage, aus Datensätzen zu lernen und trainierte Modelle aus den Daten zu erstellen. Die Tensorflow API wird verwendet, um mehrere Objekte in Echtzeit-Videostreams oder Bildern zu erkennen. [6]

Tensorflow besitzt als Standard bereits einen vortrainierten Datensatz. Dieser ist jedoch nur allgemein auf alltägliche Objekte trainiert. Der Datensatz erkennt Gegenstände wie Flaschen, Blumen, Notebooks und andere Objekte die im Alltag vorkommen (siehe Abbildung 6).

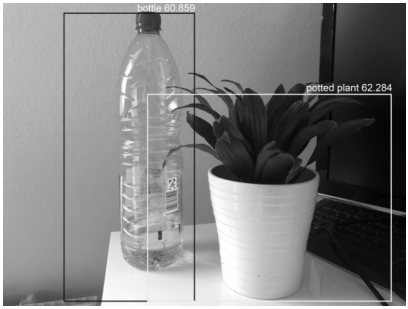


Abbildung 6: Tensorflow Objekt Erkennung

Falls jedoch eine spezielle Erkennung von bestimmten Objekten notwendig ist, muss ein neuer Datensatz speziell auf die gewünschten Objekte erstellt bzw. trainiert werden. Als Beispiel können die Notizen von Kapitel 3.3 genommen werden. Damit der neue Datensatz in der Lage ist Notizen bzw. Post-it Notes zu erkennen, müssen ungefähr 150 Bilder [6] von Notizen erstellt werden und für das Training verwendet werden. Der Einsatz im Projekt hat gezeigt, dass die 150 Bilder verschiedene Varianten von den Notizen enthalten sollten. Hier sollten die Bilder aus verschiedenen Perspektiven mit verschiedenen Lichtverhältnissen aufgenommen werden. Ebenfalls sollten leere und vollständige Notizen

aufgenommen werden, so wie Notizen bzw. Post-it Notes aus verschiedenen Farben. Je mehr Bilder verwendet werden die unterschiedlich sind, desto höher und akkurater wird die Erkennung von den Objekten sein.

5.4 OpenCV

Open Computer Vision (OpenCV), ist eine Open-Source-Bibliothek für die Bild- und Videoanalyse, die ursprünglich vor mehr als zehn Jahren von Intel eingeführt wurde. OpenCV verfügt über mehr als 2500 optimierte Algorithmen. Die OpenCV-Bibliothek (seit Version 2.2) ist in mehrere Module unterteilt, wobei jedes Modul im Allgemeinen zu einer Gruppe zugeordnet werden kann, die sich mit einem Bildverarbeitungsproblem befasst. [7]

Da eine Matrix eine grundlegende Datenstruktur in OpenCV ist, erscheint es sinnvoll, dass jedes Element der Matrix ein Pixel repräsentiert. Bei einem Grauwertbild, sind die Pixelwerte selbst jedoch typischerweise unsignierte 8-Bit-Werte, während bei einem Farbbild, drei solche 8-Bit-Werte pro Pixel verwendet werden, um drei Grundfarben, d.h. Kanäle (Rot, Grün und Blau) zu beschreiben. Um auf jedes einzelne Matrixelement zugreifen zu können, müssen nur Zeile und Spalte angeben werden. [7]

Unter OpenCV gibt es unteranderem folgende Module[3], die auf Matrixelemente zugreifen können.

- Image Processing (imgproc) - ein Bildverarbeitungsmodul, das lineare und nichtlineare Bildfilterung, geometrische Bildtransformationen, Farbraumkonvertierung, Histogramme usw. beinhaltet.
- Video Analysis (video) - ein Videoanalysemodul, das Bewegungsschätzung, Hintergrundsubtraktion und Objektverfolgungsalgorithmen beinhaltet
- Object Detection (objdetect) - Erkennung von Objekten und Instanzen der vordefinierten Klassen (z.B. Gesichter, Augen, Tassen, Personen, Autos usw.)

5.5 Handwritten Text Recognition (HTR) & Word Segmentation

Die Handwritten Text Recognition (HTR) wurde mit Tensorflow implementiert und wurde mit dem IAM Offline-HTR-Datensatz trainiert. Dieses Neuronale Netzwerk erkennt den Text von Bildern, welche segmentierte Wörter enthalten (siehe Abbildung 7).[4]

Abbildung 7: Handwritten Text Recognition (HTR)[4]

Testversuche im Projekt haben gezeigt, dass HTR jedoch nur in der Lage ist einzelne Wörter zu erkennen. Versucht man innerhalb eines Bildes einzelne Wörter eines Satzes zu erkennen, werden diese entweder nicht erkannt oder es wird nur ein Bruchteil der eigentlichen Anzahl erkannt. Damit HTR richtig funktioniert, ist daher die Word Segmentation[5] notwendig.

Die Word Segmentation ist in der Lage, jedes Wort innerhalb eines Satzes in einem Bild zu erkennen und mit einer Bounding Box zu markieren (siehe Abbildung 8). [5]

[3] docs.opencv.org/4.1.1/d1/dfb/intro.html

[4] github.com/githubharald/SimpleHTR

[5] github.com/githubharald/WordSegmentation

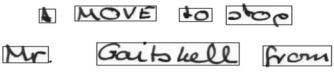


Abbildung 8: Word Segmentation[5]

Bei der Verwendung werden die einzelnen erkannten Wörter, jeweils als eigene Bilder gespeichert und können somit an das HTR übergeben werden. Allerdings gibt es momentan das Problem, dass die Erkennung der Wörter vom HTR, zurzeit nur mit der englischen Sprache richtig funktioniert. Bei der deutschen Sprache werden die Wörter häufig falsch erkannt. Hier müsste ein neues Datenmodell mit deutschen Wörter trainiert werden (siehe auch Kapitel 3.2).

6 Umsetzung

Für die Umsetzung wurden zwei Prototypen entwickelt, der erste Prototyp wurde mit OpenCV entwickelt und der zweite mit Tensorflow. Der Grund dafür war, dass beide alle Anforderungen erfüllt hatten (siehe Tabelle 2).

Tabelle 2. Erfüllte Anforderungen

Anforderungen	Open CV	Tensor Flow
1	Ja	Ja
2	Ja	Ja
3	Ja (durch Python)	Ja (durch Python)
4	Ja	Ja
5	Ja	Ja
6	Ja	Ja
7	Ja (unabhängig)	Ja(unabhängig)
8	Ja	Ja
9	Ja	Ja

6.1 OpenCV und Tensorflow

Beim Start der beiden Prototypen, hat OpenCV auf eine Bilddatei zugegriffen und Tensorflow auf den Live-Feed von der Webcam.

Aufgrund der Verwendung einer Bilddatei mit OpenCV, war die CPU-Last viel geringer und die Erkennung auch schneller. OpenCV hat dabei alle erkannten Objekte mit einer Bounding Box markiert und den Prozentwert über die Bounding Box geschrieben. Nach der Erkennung wurde eine Kopie der originalen Bilddatei erstellt, die jedoch alle Bounding Boxes beinhaltet hat. Über die Koordinaten der Bounding Boxes, ist OpenCV in der Lage die einzelnen Objekte auszuschneiden und als eigene Bilddateien zu speichern.

Mit Tensorflow läuft der Ablauf ähnlich wie in OpenCV ab. Hier werden über die Webcam Frames vom Live-Feed für die Erkennung gespeichert. Aufgrund des Live-Feeds, ist die CPU Last extrem hoch und die Erkennung kann dadurch auch mehrere Sekunden länger dauern. Um das schnellste Ergebnis zu bekommen, wäre eine Leistungsstarke GPU notwendig. Als nächstes werden auch bei Tensorflow, die erkannten Objekte in eigene Bilddateien ausgeschnitten und gespeichert.

Ab diesem Schritt ist der Ablauf der beiden Prototypen gleich. Die Bilddateien der erkannten Objekte, werden in eine Base64 Codierung umgewandelt und in eine Datenbank gespeichert. Eine ID wird automatisch vergeben und der Name des Eintrags in der Datenbank, entspricht dem Namen des erkannten Objektes mit einem Zeitstempel. Bei erkannten Objekten, bei denen es sich um Notizen handelt, werden die Libraries aus Kapitel 4.5 verwendet. Aber wie bereits erwähnt, gibt es noch Probleme mit der Genauigkeit, was in der Zukunft behoben werden soll.

Ebenfalls soll es zukünftig möglich sein, erkannte Personen mit IDs zu assoziieren und wiederzuerkennen. Dadurch wird es möglich sein, für eine Person die Daten einer Brainstorming-Session wieder aufrufen zu können. Ein weiterer wichtigerer Schritt, wäre es auch eigene Datensätze zu trainieren, um die Bestmögliche Erkennung der Objekte zu garantieren. Über eine zukünftige Studie

und mehreren Testversuchen, sollen auch mögliche Probleme behoben und Verbesserungen implementiert werden.

7 Fazit und Ausblick

Anhand der Anforderungen, wurden die zwei Prototypen aus Kapitel 6 implementiert. Die Prototypen laufen alle stabil im Hintergrund ab und erkennen auch die Objekte, die im aktuellen Datenmodell vorhanden sind. Zukünftig müsste ein neues Datenmodell mit einem eigenen Datensatz trainiert werden, welches z.B. speziell nur Bilder mit Notizen enthält. Dadurch wird sich die Genauigkeit der erkannten Objekte (z.B. Notizen) erhöhen. Die beiden Prototypen, müssen in Zukunft in einer Brainstorming Session getestet werden, um diese zu bewerten um den besseren Prototypen auszuwählen. Dabei soll auch herausgefunden werden, welche Funktionen verbessert werden müssen, z.B. die Handschrifterkennung und die Erkennung der Personen. Hier muss untersucht werden, wie genau die Erkennung der Personen ist und ob es zu Verwechselungen kommt. Wenn ja, dann ist es nötig die IDs der Personen, über den NFC-Chip des Smartphones zu identifizieren. Sobald alle Funktionen problemlos ablaufen, soll eine Studie durchgeführt werden, welche zeigen soll, wie sich das Verhalten und die Ergebnisse der Teilnehmer einer Brainstorming Session verändern, sobald das kollaborative System um Object Recognition erweitert wird.

8 Literaturverzeichnis

[1] M. Merk, G. Tullius, P. Hertkorn. Automating faciliation and group building tasks in co-located ideation sessions using multiple ideation tables equipped with projection and communication abilities. https://swuxlab.reutlingen-university.de/2019/04/29/swuxlab-auf-der-chi-2019-in-glasgow/;Glasgow, Scotland UK, CHI 2019

[2] G. Chandan, A. Jain, H. Jain, Mohana. Real Time Object Detection and Tracking Using Deep Learning and OpenCV. https://ieeexplore.ieee.org/document/8597266;Coimbatore, India, IEEE 2018, DOI 10.1109/ICIRCA.2018.8597266

[3] R. Vaidya, D. Trivedi, S. Satra, Prof. M. Pimpale. Handwritten Character Recognition Using Deep-Learning. https://ieeexplore.ieee.org/document/8473291;Coimbatore, India. IEEE 2018, DOI 10.1109/ICICCT.2018.8473291

[4] M. Olafenwa, J. Olafenwa. ImageAI, an open source python library built to empower developers to build applications and systems with self-contained Computer Vision capabilities, Website, 2018. Online verfügbar unter https://github.com/OlafenwaMoses/ImageAI; Besucht am 14.09.19

[5] J. Redmon, S. Divvala, R. Girshick, A. Farhadi. You Only Look Once: Unified, Real-Time Object Detection, Website 2016. Online verfügbar unter https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf; Besucht am 14.09.19

[6] R. Phadnis, J. Mishra, S. Bendale. Objects Talk - Object Detection and Pattern Tracking Using TensorFlow. https://ieeexplore.ieee.org/document/8473331;Coimbatore, India, IEEE 2018, ISBN 978-1-5386-1974-2

[7] I. Culjak, D. Abram, T. Pribanic, H. Dzapo, M. Cifrek. A brief introduction to OpenCV. https://ieeexplore.ieee.org/document/6240859;Opatija, Croatia, IEEE 2012, ISBN 978-953-233-068-7

Automatische Analyse von Gesichtsmuskelaktivität mit Hilfe von 2D- und 3D-Daten

Tim Wiedemer
Hochschule Reutlingen
Tim.Wiedemer@Student.Reutlingen-University.de

Abstract

In dieser Arbeit wird ein Programm für die automatische Analyse der Gesichtsmuskelaktivität mithilfe von 2D- bzw. 3D-Daten in Echtzeit entwickelt. Hierfür wird das Facial Action Coding System mit seinen Action Units verwendet. Es soll gezeigt werden, ob der gewählte Ansatz für weitere Untersuchungen vertieft werden kann, indem die Aktivitäten von drei verschiedenen Muskelaktivitäten im Gesicht analysiert werden. Diese wurden mit einem Motion Capture System sowie einer Webcam aufgenommen. Der Ansatz eignet sich für die Analyse von 2D- und 3D-Daten. Das Motion Capture System bietet allerdings mehr Vorteile und wird deshalb in weiteren Arbeiten gegenüber der 2D-Datenaufnahme bevorzugt.

CCS Concepts

• **Computing methodologies** → **Motion capture**; *Visual analytics*; • **Theory of computation** → *Constraint and logic programming*; • **Mathematics of computing** → Solvers;

Keywords

markerbasierte Gesichtsmuskelanalyse, nichtmarkerbasierte Gesichtsmuskelanalyse, Facial Action Coding System, eingeschränkte quadratische Optimierung

1 Einleitung

Seit Jahrzehnten nimmt die Forschung im Bereich der Gesichts- und Emotionsanalyse und der digitalen Gesichtsmodellierung stark zu [1]. Die Anwendungsbereiche dieser Forschungsergebnisse sind breit gefächert. So haben nicht nur die Film- und Spieleindustrie ein großes Interesse am Verstehen und Modellieren von Gesichtsausdrücken, sondern auch die Medizin oder Forensik [2], etwa für Diagnosenfindungen oder Gesichtsrekonstruktionen. Auch in der Mensch-Maschinen-Interaktion und in der Robotik gewinnt die Gesichtsanalyse zunehmend an Bedeutung [3]. Um die Gesichts- und Emotionsanalysen objektiv, robust und kostengünstig zu halten, wurden über die Jahre unterschiedliche 2D- und auch 3D-basierte visuelle sowie auf Akustik basierende Lösungsansätze entwickelt [1].

1.1 Ziel der Arbeit

Das Ziel dieser Arbeit ist es, mit dem von Choe und Ko [4] vorgestellten Algorithmus ein Programm zu entwickeln, welches die Gesichtsmuskelaktivität von zuvor festgelegten Action Units in Echtzeit messen bzw. schätzen kann. Dabei werden 2D- und 3D-Daten generiert und im direkten Vergleich auf ihre

Betreuer Hochschule: Prof. Dr.-Ing. Cristóbal Curio
Hochschule Reutlingen
Cristobal.Curio@Reutlingen-University.de
Salma Thalji M.Sc.
Hochschule Reutlingen
Salma.Thalji@Reutlingen-University.de

Wissenschaftliche Vertiefungskonferenz
27. November 2019, Hochschule Reutlingen
Copyright 2019 Tim Wiedemer

Eignung für den bestehenden Ansatz getestet [4]. Die gewonnenen Erfahrungen und Ergebnisse sollen in einer darauf aufbauenden Arbeit verwendet werden.

1.2 Aufbau der Arbeit

In dieser Ausarbeitung wird in Kapitel 2 zunächst das Facial Action Coding System (FACS) mit seinen Action Units (AUs) erläutert. Anschließend werden in Kapitel 3 verwandte Arbeiten vorgestellt, die eine Gesichtsmuskelanalyse betreiben. Die Umsetzung des gewählten Choe und Ko [4] Ansatzes wird in Kapitel 4 vorgestellt, ehe anschließend in Kapitel 5 die Ergebnisse diskutiert werden. Das letzte Kapitel 6 beinhaltet das Fazit sowie einen Ausblick zur Arbeit.

2 Facial Action Coding System

Der menschliche Gesichtsausdruck bildet nicht nur die Grundlage zur persönlichen Kommunikation [2], sondern liefert Informationen zur Emotion, sozialen Interaktion sowie Persönlichkeit und Absicht [5]. Emotionen (z.B. Angst und Ekel) können sich in der Form des Gesichtsausdrucks stark ähneln. Einen Gesichtsausdruck, der sich aus der Aktivität verschiedener Muskelpartien zusammensetzt, einer bestimmten und bereits bekannten Emotion zuzuordnen, ist daher schwierig [6]. Zur Lösung dieses Problems wurde Ende der 70er Jahre das FACS veröffentlicht [7]. Es beruht auf menschlichen Wahrnehmungen und wurde über die Jahre zum Standard für die Gesichts-, Emotions- und Schmerzerkennung [1].

Im FACS wird ein Gesichtsausdruck in seine anatomisch basierte Gesichtsbewegungen zerlegt. Dafür wird jede Kontraktion eines oder mehrerer Gesichtsmuskel einer bestimmten AU zugeordnet. Insgesamt wurden 46 AUs definiert, welche die Gesichtsmuskel betreffen. Außerdem wurden über 7.000 AU-Kombinationen im Alltag beobachtet [1]. Abbildung 1 zeigt sechs AUs, die der oberen oder unteren Gesichtshälfte zugewiesen sind. So sehr sich manche AUs gegenseitig beeinflussen und häufiger in Kombination auftreten (z.B. AU1 + AU4), so stark können sich auch einzelne AUs aufgrund der Anatomie im Gesicht gegenseitig ausschließen [1].

Das Erlernen der AU Codierung bedarf mehrerer 100 Stunden Training. Um eine Minute Videomaterial zu codieren benötigt ein erfahrener Codierer über eine Stunde [1]. Diese zeitintensiven Arbeiten sowie die bei der manuellen Codierung fehlende Objektivität macht das FACS teuer und fehleranfällig, weshalb in der Vergangenheit automatische Codierungssysteme entwickelt wurden [8].

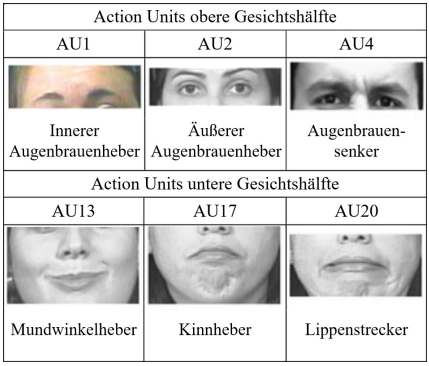


Abbildung 1: Beispiele für Action Units der oberen und unteren Gesichtshälften [5].

3 Verwandte Arbeiten

Die jahrelange Forschung im Gebiet der Gesichtsanalyse brachten unterschiedliche Verfahren hervor, mit denen die AUs des FACS gemessen bzw. geschätzt werden. Während es heutzutage immer mehr auf 3D-Daten basierende Arbeiten und Ansätze gibt [1], wurden in der Vergangenheit eine Reihe von 2D-basierten Ansätzen entwickelt. Bevilacqua und seine Kollegen [9] bewerteten die Gesichtsausdrücke mit Hilfe von Polygonbereiche, die aus den 2D-Daten erstellt wurden. Diese Bereiche weisen je nach ihrer Größe

auf eine schwache oder starke Ausprägung einer AU hin.
Einen bis dahin völlig unerforschten Ansatz entwickelten Jarlier et. al [10], die unter Verwendung der Temperaturänderung in einem Gesichtsausdruck Muster für AUs bestimmten. Sie zeigten, dass sich die thermischen Schwankungen zu einer spezifischen AU zuteilen lassen.

4 Umsetzung

Die in der Umsetzung durchgeführten Aufnahmen wurden mit einem Probanden durchgeführt. Weitere Probanden wurden nicht aufgezeichnet, da dies zur Evaluierung des Ansatzes keinen Mehrwert beitragen würde. Für tiefergehende Arbeiten ist die Aufnahme weiterer Probanden jedoch unumgänglich, um absolute Subjektivität ausschließen zu können.

4.1 Datenerhebung

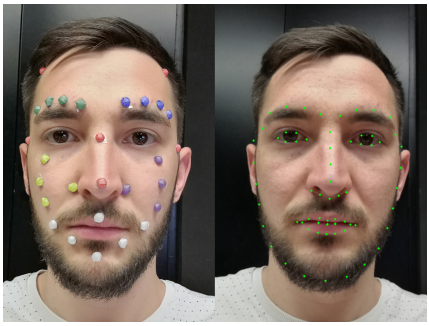


Abbildung 2: Links: 28 nach ihrer Gruppenzugehörigkeit eingefärbte Gesichtsmarker. Rot: Referenzmarker, grün: rechte Augenbraue, blau: linke Augenbraue, gelb: rechte Wange, violett: linke Wange, weiß: Mund. Rechts: Mit OpenCV und dlib[1] erfasste 68 Landmarken.

3D-Daten: Die 3D Daten wurden mit dem Vicon Motion Capture System aufgenommen. Die Anordnung der 28 3D-Marker ist im linken Teil der Abbildung 2 dargestellt. Für eine bessere Abgrenzung der einzelnen Gesichtsbereiche wurden die Marker bei der Aufzeichnung in entsprechende Gruppen eingeteilt. Diese Gruppen sind Augenbraue links/rechts, Backe links/rechts, Mund sowie Referenzmarker.

2D-Daten: Die 2D-Daten wurden mithilfe der OpenSource Frameworks OpenCV und dlib[1] unter Python aufgezeichnet. Für die Gesichtserkennung nutzt OpenCV den von Viola und Jones [11] entwickelten Ansatz. Dieser wurde bis heute am häufigsten eingesetzt [1], da die Erkennung des Gesichts sehr robust ist. Innerhalb des erkannten Gesichtsbereichs werden anschließend die Landmarken, unter Verwendung des von Kazemi und Sullivan [12] veröffentlichten Entscheidungsbaumansatzes, bestimmt. Die Anordnung der 68 Landmarken sind im rechten Teil der Abbildung 2 dargestellt. Gut ersichtlich ist die Gruppierung der Landmarken für die unterschiedlichen Gesichtsbereiche (Augen, Augenbrauen, Mund, Nase und Gesichtskontur um Kiefer und Kinn).
In dieser Ausarbeitung wurden vier Basiselemente mit 2D- und 3D-Aufnahmen bestimmt, die in Abbildung 3 zu sehen sind und jeweils die stärkste Intensität einer AU darstellen. Neben dem neutralen Basiselement wurden Basiselemente mit hochgezogener linker Augenbraue (entspricht AU2 links), sowie angezogenem linken (entspricht AU13 links) und rechten Mundwinkel (entspricht AU13 rechts) aufgenommen. Weitere AUs lassen sich für künftige Analysen nach Belieben hinzufügen. Nach der Datenerhebung wurden die Daten der Basiselemente zum neutralen Gesicht ausgerichtet und zusätzlich von Hand bereinigt. Durch die Bereinigung werden jene Marker und Landmarken, welche die AUs nicht beeinflussen, auf die Werte des neutralen Gesichtsausdrucks gesetzt.

[1] Weitere Informationen: dlib – http://www.dlib.net/ und OpenCV – http://www.opencv.org/.

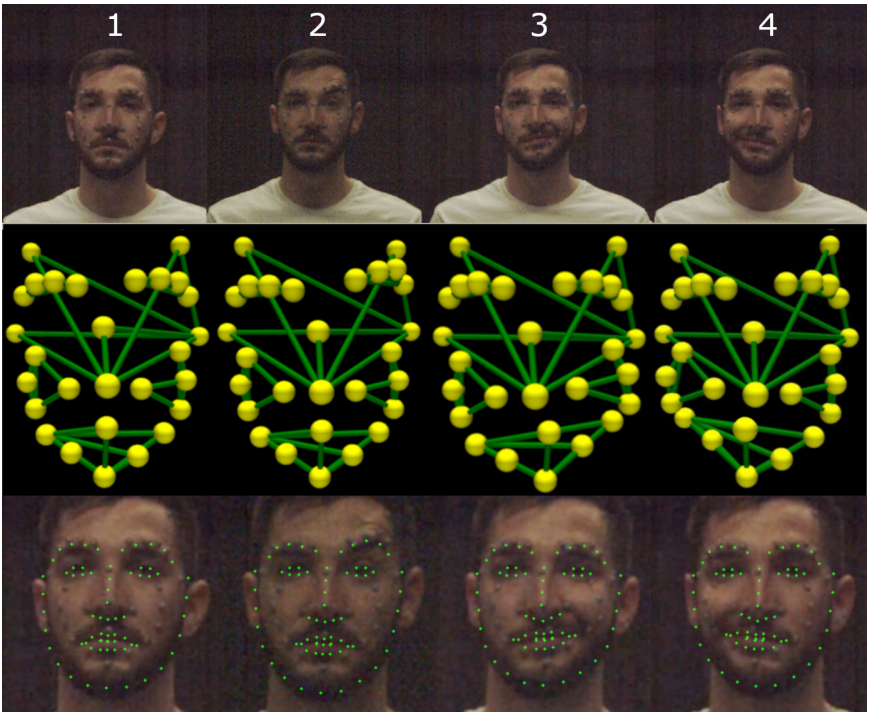


Abbildung 3: Aufgenommene Gesichtsausdrücke als Basiselemente mit realen (oben) Markern, virtuellen Markern (mitte) und eingezeichneten Landmarken (unten) dargestellt. Gesichtsausdrücke: 1 - neutral, 2 - Augenbraue links, 3 - Mundwinkel links, 4 - Mundwinkel rechts.

4.2 Alignment

Während in einem dreidimensionalen Raum die Abstände der Marker zueinander immer gleich bleiben und somit die Marker stets das gleiche Zentrum besitzen, können die Abstände der Landmarken zueinander im zweidimensionalen Raum durch die Distanz zur aufzeichnenden Kamera beeinflusst werden. Für das Alignment der 3D-Daten ist eine euklidische Transformation ausreichend. Für die Berechnung der Rotation wird die Singular Value Decomposition (SVD) nach [13] verwendet. Um keine Beeinflussung aufgrund unterschiedlicher Abstände zur Kamera zu erhalten, muss für das Alignment der 2D-Daten neben der Transformation t und Rotation R auch die Skalierung s berücksichtigt werden. Ziel des Alignments ist es, die verschiedenen Momentaufnahmen in immer die gleiche Position zu transformieren, um eine gemeinsame Ausgangslage für die anschließenden Berechnungen der Muskelaktivierungen zu schaffen. Für die Berechnung der Rotation und Translation wurden die in Abbildung 2 rot eingefärbten Referenzmarker verwendet. Die Gleichung (1) beschreibt das Alignment für die 2D- bzw. 3D-Daten. Hierfür werden die n Marker des aktuellen Gesichtsausdrucks P_j^P

durch die Rotation, Skalierung und Transformation in die Position der Marker des neutralen Gesichtsausdrucks P_j^M gebracht. Weitere Variabeln und ihre Bedeutungen sind zur besseren Übersicht in Tabelle 1 zusammengefasst und gelten auch für die nachfolgenden Gleichungen.

$$P_j^M = sR\,P_j^P + t, (j = 1, 2, ..., n) \quad (1)$$

Da ein Gesichtsausdruck aus mehreren einzelnen, unterschiedlich stark getätigten Basiselementen besteht, lässt sich der Gesichtsausdruck E mit Gleichung (2) beschreiben. Hierbei steht x_i für die Muskelaktivität jedes Basiselements und nimmt je einen Wert zwischen [0..1] an. Multipliziert mit den Differenzen der Basiselementen e_i und der Addition des neutralen Ausdrucks E_0, ergibt sich hierbei der Gesichtsausdruck E.

$$E = E_0 + \sum_{i=1}^{m} x_i e_i \quad (2)$$

4.3 Berechnung der Muskelaktivierung

Die in diesem Kapitel verwendeten Gleichungen zur Berechnung der Muskelaktivierung bzw. AUs stammen aus dem vorgeschlagenen Algorithmus von Choe und Ko [4] und wurden für diese Ausarbeitung an die drei zu berechnenden AUs angepasst. Durch die Minimierung der Variabeln $\hat{d}_j$ und d_j in Gleichung (3) wird die Berechnung der Muskelaktivierung beschrieben. Die Variable d_j steht hierbei für die Distanz vom aktuellen Gesichtsausdruck zum neutralen Gesichtsausdruck. Sie lässt sich mit der Multiplikation der Muskelaktivierung x_i und der Distanz von P_j vom Basiselement E_i zu P_j vom neutralen Gesichtsausdruck E_0 berechnen.

$$\sum_{j=1}^{n} |\hat{d}_j - d_j|^2 = \sum_{j=1}^{n} |\hat{d}_j - \sum_{j=1}^{m} x_i d_{ij}|^2 \quad (3)$$

Tabelle 1: Variablen und deren Bedeutung, die in den Gleichungen verwendet werden.

Variable	Bedeutung
E_0	neutraler Gesichtsausdruck
E_i	Basiselement – enstpricht dem stärksten Ausdruck einer AU
e_i	Differenz vom Basiselement E_i zum neutralen Gesichtsausdruck E_0
m	Anzahl der Basiselemente
x_i	Muskelaktivierung – zwischen [0..1]
P_j^M	Position der Marker im neutralen Gesichtsausdruck
P_j^P	Position der Marker im aktuellen Gesichtsausdruck
d_{ij}	Distanz von P_j vom Basiselement E_i zu P_j vom neutralen Gesichtsausdruck E_0
d_j	Distanz von P_j vom aktuellen Gesichtsausdruck zu P_j vom neutralen Gesichtsausdruck E_0

Da die Muskelaktivierungen von Vektor x zwischen [0..1] liegen sollen, kann zur Lösung der Gleichung (3) die eingeschränkte quadratische Optimierung (*constrained quadratic programming*) in Gleichung (4) verwendet werden. In dieser Arbeit ist das Ergebnis der Vektor $x = [x_1, x_2, x_3]^T$ mit den enthaltenen Muskelaktivierungen der drei untersuchten AUs. Liegt ein Wert im Vektor x unterhalb von 0, muss der neutrale Ausdruck nachgebessert bzw. neu aufgenommen werden, da eine stetig leichte Muskelkontraktion nicht auszuschließen ist. Liegt ein Wert über 1, ist der untersuchte Ausdruck stärker als das Basiselement. Hier bietet es sich an,

das Basiselement mit dem untersuchten Ausdruck zu ersetzen. Erst wenn die Ergebnisse der Muskelkontraktionen zwischen [0..1] liegen, wurden in den Basiselementen jeweils die stärksten Ausdrücke gefunden.

$$\frac{1}{2}x^T Qx - x^T c$$
$$0 \leq x_i \leq 1 \; (i = 1, 2, ..., m) \tag{4}$$

In der Matrix Q in Gleichung (5) werden die Distanzen der Basiselemente, d.h. die Entfernung vom stärksten zum neutralen Ausdruck der einzelnen AUs, miteinander multipliziert.

$$Q = 2\begin{pmatrix} \sum_{j=1}^{n} |d_{1j}|^2 & \sum_{j=1}^{n} d_{1j}d_{2j} & \sum_{j=1}^{n} d_{1j}d_{3j} \\ \sum_{j=1}^{n} d_{2j}d_{1j} & \sum_{j=1}^{n} |d_{2j}|^2 & \sum_{j=1}^{n} d_{2j}d_{3j} \\ \sum_{j=1}^{n} d_{3j}d_{1j} & \sum_{j=1}^{n} d_{3j}d_{2j} & \sum_{j=1}^{n} |d_{3j}|^2 \end{pmatrix} \tag{5}$$

Der Vektor c in Gleichung (6) beinhaltet ebenfalls die Distanzen der Basiselemente zum neutralen Ausdruck, jedoch werden diese mit der zum aktuellen Ausdruck gemessenen Entfernung multipliziert.

$$c = 2\begin{pmatrix} \sum_{j=1}^{n} \hat{d}_j d_{1j} \\ \sum_{j=1}^{n} \hat{d}_j d_{2j} \\ \sum_{j=1}^{n} \hat{d}_j d_{3j} \end{pmatrix} \tag{6}$$

Die Lösung des Vektors x wird im nächsten Kapitel besprochen.

5 Diskussion

Die in Kapitel 4 durchgeführten Berechnungen sind sowohl für die hochauflösenden 3D-Daten, als auch für die mit dlib und OpenCV erfassten 2D-Daten geeignet. Abbildung 4 zeigt die Ergebnisse der Gesichtsausdrücke aus Abbildung 3, aufgeteilt nach den AUs innerhalb ihrer Datenform 3D (Vicon) bzw. 2D (OpenCV).

Es zeigte sich, dass die Ergebnisse der Vicon Daten genauer und stabiler sind und keine sprunghaften Ausreißer zwischen den einzelnen Frames beinhalten, wie es bei den OpenCV Daten des Öfteren der Fall war. Gründe hierfür waren vor allem die starke Abhängigkeit der Bildqualität und -beleuchtung, sowie weitere Störfaktoren – bspw. die Gesichtsbehaarung. Außerdem benötigt der 2D-Ansatz eine frontale Aufnahme des Gesichts, damit das Gesicht sowie die Landmarken korrekt und ohne größere Fehler erfasst werden können. Die Notwendigkeit des relativ kontrollierten Szenarios mit einer frontal aufgezeichneten Kopfhaltung haben z.B. auch Martinez [1] in ihrer Arbeit bemängelt, da auf diese Weise die Natürlichkeit des Ausdrucks verloren gehen kann. Da das Motion Capture System in einer bestimmten Szene von mehreren Quadratmetern die Marker erfassen kann, stellt dies ein wesentlicher Vorteil gegenüber den mit OpenCV aufgenommenen 2D-Daten dar.

Ein Problem, das in weiteren Arbeiten angegangen werden muss, ist die Ungenauigkeit der Berechnung einer nicht aktivierten AU. Bei beiden Datenformen zeigt sich, dass im neutralen Gesichtsausdruck (1 in Abbildung 3) jede AU eine Aktivierung aufweist. Idealerweise würden hier die Ergebnisse für jede AU bei 0 liegen. Dennoch haben hier die 3D-Daten einen entscheidenden Vorteil, da auch leichte Intensitäten einer AU besser wahrgenommen und berechnet wurden, was ebenfalls von Savran et. al [14] als Ergebnis festgehalten wurde. Die deutlich bessere Wahrnehmung leichter Aktivität, kann mit den 125 Frames pro Sekunde begründet werden.

Der negative Wert der AU13R in Ausdruck 2 aus Abbildung 4 der OpenCV Daten wird sehr wahrscheinlich auch aufgrund von schlechten Bild- und Lichtverhältnissen und nicht wegen einem schlecht aufgenommenen neutralen Gesichtsausdruck aufgetreten sein. Die

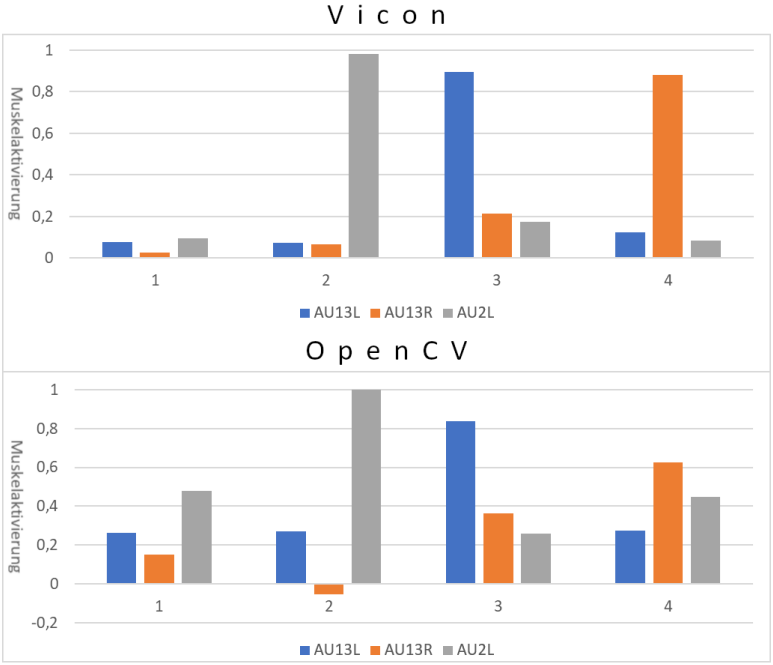


Abbildung 4: Ergebnisse der Muskelaktivitätsberechnungen aufgeteilt nach Gesichtsausdrücken und den Datenerhebungen von Vicon (oben) und OpenCV (unten). AU13L: Mundwinkelheber links, AU13R: Mundwinkelheber rechts, AU2L: äußerer Augenbrauenheber links.

restlichen Ergebnisse, gerade aus den Berechnungen der 3D-Daten aus Vicon, zeigen keine weiteren großen Auffälligkeiten.

Die Aufnahme der Basiselemente war bei den 2D- und 3D-Daten gleichermaßen einfach, vorausgesetzt die Bildqualität und -beleuchtung im bildbasierten Vorgehen waren ausreichend genug. Ein schnell auftretendes Problem wird mit hoher Wahrscheinlichkeit sein, dass die 2D-Landmarkendetektion innerhalb des Gesichtes begrenzt ist. Wird die Gesichtsaktivität auf den tatsächlichen Muskel zurückgeführt, können dabei wichtige Informationen verloren gehen. Hier bietet der markerbasierte Ansatz einen wesentlichen Vorteil, da sich diese Informationen gezielt berücksichtigen lassen. Der Verlauf der Muskeln ist in Abbildung 5 zu sehen, wobei schnell ersichtlich wird, dass nicht jeder Muskel durch die Landmarken repräsentiert werden kann, in etwa, wenn es sich um Muskeln an den Wangen handelt. Zusätzlich wird ein 2D-basiertes System Schwierigkeiten haben, bestimmte AU Aktivitäten zu erkennen. So wird es nahezu unmöglich sein, die Verformung der Unterlippe aus AU17 (Abbildung 1) zu erkennen, wenn die Kamera nur frontal auf den Probanden gerichtet ist. Hier können wiederum 3D Marker eingesetzt werden, die auch in der Tiefe wichtige Daten liefern. Ein ähnliches Problem bemerkten auch Sandbach et. al [16], da in ihrer Arbeit die AU18 bei

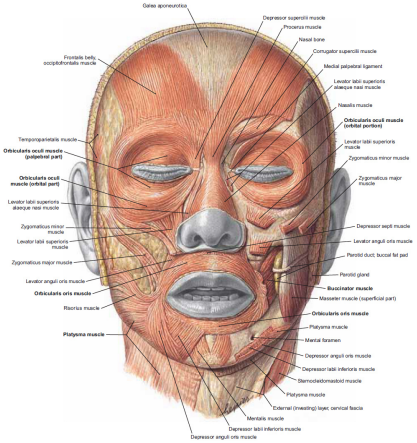


Abbildung 5: Anatomie von Gesichtsmuskeln [15]. Einige der schematisch dargestellten Gesichtsmuskeln sind auf beiden Seiten vorhanden.

einer frontalen Draufsicht teilweise schwer von AU10, AU17 und AU24 zu unterscheiden waren.

Interessant wäre die Kombination eines markerbasierten mit einem markerlosen System. So können Marker gezielt dort platziert werden, wo die markerlosen Systeme keine Gesichtsaktivität messen können oder die einzelnen AUs schwer voneinander unterscheidbar sind. Den Fokus auf fusionierte 2D- und 3D-Modalitäten zu lenken wurde bereits vor einigen Jahren vorgeschlagen [14].

Als weiteres Problem stellte sich der Rückgang der einzelnen AU Intensitäten heraus, wenn diese gleichzeitig getätigt werden. Dies zeigt ein Vergleich in der Abbildung 6. Zu sehen sind die Ergebnisse der Berechnungen aus den 2D- und 3D-Daten, wobei die Daten aus dem gleichen Frame generiert wurden. Die Ergebnisse der markerbasierten Berechnungen sind, verglichen mit weiteren berechneten Aufnahmen, wesentlich genauer als die Ergebnisse der nichtmarkerbasierten Berechnungen. Gründe hierfür können ein schlechtes 2D-Alignment oder eine ungenaue Landmarkdetektion bei gleichzeitig auftretenden Ausdrücken sein.

Die Entscheidung vorerst nur beispielhaft drei AUs zu untersuchen erwies sich als richtig, da so schneller erste Ergebnisse erzielt wurden. Nichtsdestotrotz müssen in weiteren Untersuchungen zusätzliche AUs in die Arbeit aufgenommen werden. Zur Überprüfung der AU Ergebnisse sollten anatomische Gegebenheiten berücksichtigt werden, da wie in Kapitel 2 bereits genannt, manche AUs sich gegenseitig ausschließen.

6 Fazit und Ausblick

Die Implementierung des Ansatzes von Choe und Ko [4] funktioniert sowohl für die 2D- als auch für die 3D-Daten. Die einfache Erweiterbarkeit sowie der vergleichbare niedrige Rechenaufwand zeigt, dass das System weiter vertieft werden kann. Hierfür müssen zusätzliche Basiselemente für die Berechnung der AUs aufgenommen und ihr Zusammenspiel miteinander analysiert werden. Im direkten Vergleich liefern die 3D-Daten gerade bei schwach ausgeprägten AUs bessere Ergebnisse als die 2D-Daten, weshalb sich in zukünfigten Untersuchungen auf das Motion Capture System konzentriert wird.

Das Erzeugen von 2D-Daten ist zwar schneller und kostengünstiger, jedoch liefert das relativ überwachte Szenario – frontale Aufnahme ohne Kopfdrehen – vermutlich weniger natürliche Gesichtsausdrücke als in einem Motion Capure System, in welchem sich die aufgezeichnete Person frei bewegen kann. Außerdem bringen die 3D-Daten genauere Ergebnisse, gerade wenn es um das Unterscheiden von AUs oder um AUs mit schwacher Aktivierung geht.

Der Trend für die kommende Jahre wird sehr wahrscheinlich zu fusionierten 2D- und 3D-Modalitäten gehen, die auch die zeitlichen

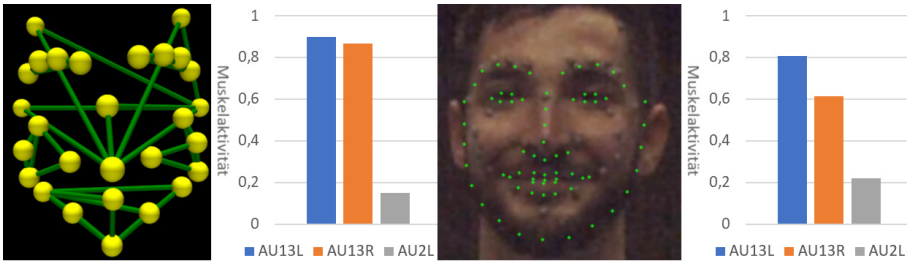


Abbildung 6: a) 28 virtuelle Marker beim Gesichtsausdruck *Grinsen* mit den AU Muskelaktivitäten. b) 68 Landmarken beim Gesichtsausdruck *Grinsen* mit den AU Muskelaktivitäten. AU13L: Mundwinkelheber links, AU13R: Mundwinkelheber rechts, AU2L: äußerer Augenbrauenheber links.

Aspekte einer Gesichtsaufnahme berücksichtigen werden, um genauere Vorhersagen über den Ausdruck machen zu können.
In weiteren Untersuchungen werden weitere Probanden mit dem Motion Capture System aufgenommen, da die Ausprägungen der Gesichtsausdrücke eines jeden Menschen unterschiedlich sind. Außerdem wird in den zukünftigen Untersuchungen vorerst auf die Nutzung von 2D-Daten verzichtet, da diese deutlich zu häufig ungenügende Ergebnisse lieferten.

Literatur

[1] B. Martinez, M. F. Valstar, B. Jiang und M. Pantic, "Automatic Analysis of Facial Actions: A Survey", *IEEE Transactions on Affective Computing*, S. 1, 2017.

[2] M. H. Alkawaz, D. Mohamad, A. H. Basori und T. Saba, "Blend Shape Interpolation and FACS for Realistic Avatar", *3D Research*, Jg. 6, Nr. 1, S. 2799, 2015.

[3] M. Breidt, H. H. Bülthoff und C. Curio, "Robust semantic analysis by synthesis of 3D facial motion", in *Face and Gesture 2011*, IEEE, 21.03.2011 - 25.03.2011, S. 713–719.

[4] B. Choe und H.-S. Ko, "Analysis and synthesis of facial expressions with hand-generated muscle actuation basis", in *ACM SIGGRAPH 2005 Courses on - SIGGRAPH '05*, J. Fujii, Hrsg., New York, New York, USA: ACM Press, 2005, S. 15.

[5] F. De la Torre und J. F. Cohn, "Facial Expression Analysis", in *Visual Analysis of Humans*, T. B. Moeslund, A. Hilton, V. Krüger und L. Sigal, Hrsg., Bd. 33, London: Springer London, 2011, S. 377–409.

[6] A. Pampouchidou, P. Simos, K. Marias, F. Meriaudeau, F. Yang, M. Pediaditis und M. Tsiknakis, "Automatic Assessment of Depression Based on Visual Cues: A Systematic Review", *IEEE Transactions on Affective Computing*, S. 1, 2017.

[7] P. Ekman und W. V. Friesen, "Facial Action Coding System: A Technique for the Measurement of Facial Movement", 1978.

[8] J. Hamm, C. G. Kohler, R. C. Gur und R. Verma, "Automated Facial Action Coding System for dynamic analysis of facial expressions in neuropsychiatric disorders", *Journal of neuroscience methods*, Jg. 200, Nr. 2, S. 237–256, 2011.

[9] V. Bevilacqua, D. D'Ambruoso, G. Mandolino und M. Suma, "A new tool to support diagnosis of neurological disorders by means of facial expressions", in *2011 IEEE International Symposium on Medical Measurements and Applications*, IEEE, 30.05.2011 - 31.05.2011, S. 544–549.

[10] S. Jarlier, D. Grandjean, S. Delplanque, K. N'Diaye, I. Cayeux, M. I. Velazco, D. Sander, P. Vuilleumier und K. R. Scherer, "Thermal Analysis of Facial Muscles Contractions", *IEEE Transactions on Affective Computing*, Jg. 2, Nr. 1, S. 2–9, 2011.

[11] P. Viola und M. Jones, "Robust real-time face detection", in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, IEEE Comput. Soc, 7-14 July 2001, S. 747.

[12] V. Kazemi und J. Sullivan, "One millisecond face alignment with an ensemble of regression trees", in *2014 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2014, S. 1867–1874.

[13] K. S. Arun, T. S. Huang und S. D. Blostein, "Least-Squares Fitting of Two 3-D Point Sets", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Jg. PAMI-9, Nr. 5, S. 698–700, 1987.

[14] A. Savran, B. Sankur und M. Taha Bilge, "Comparative evaluation of 3D vs. 2D modality for automatic detection of facial action units", *Pattern Recognition*, Jg. 45, Nr. 2, S. 767–782, 2012.

[15] C. D. Clemente, *Anatomy: A regional atlas of the human body*, 6th ed. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins Health, 2011.

[16] G. Sandbach, S. Zafeiriou, M. Pantic und L. Yin, "Static and dynamic 3D facial expression recognition: A comprehensive survey", *Image and Vision Computing*, Jg. 30, Nr. 10, S. 683–697, 2012.

Trennung der Dekodierung und Detektion von Markern zur Positionsüberwachung eines Patientenpositionierroboters in Echtzeit

Hannah Büchner
Hochschule Reutlingen
Hannah.Buechner@Student.Reutlingen-University.de

Abstract

Die Strahlentherapie ist eine zentrale Säule für die Behandlung von Krebs. Dabei werden effiziente Strahlendosen und die optimale Zellzerstörung von Krebszellen mittels exakter Positionierung des Patienten vor den Strahlenausgang angestrebt. Für die hoch präzise Positionierung und Überwachung von Patienten werden Roboter eingesetzt, da sie den Vorteil einer hohen Reproduzierbarkeit aufweisen. Für die optimale Ausrichtung und zur Sicherheit der Patienten muss vor und während der Strahlenbehandlung auch bei unvorhergesehenen Bewegungen, ausgelöst durch beispielsweise ein Husten oder Niesen des Patienten, schnell reagiert werden. Deshalb müssen die Positionen und Bewegungen des Positionierroboters in Echtzeit kontrolliert werden. Ein solches robotergestütztes Patientenpositioniersystem wird im Rahmen dieser Arbeit untersucht. Es wird ein Konzept zur Trennung der Dekodierung und Detektion von Markern für die Positionsverfolgung entwickelt. Mit der Trennung soll eine Zeitreduzierung bei der Positionsanalyse des Roboters erreicht werden, um eine Echtzeittauglichkeit für das System zu erreichen. Die Ergebnisse zeigen, dass durch diese Trennung eine Reduzierung der Zeit erreicht werden kann.

CCS Concepts

• **Computer systems organization → Embedded systems**; *External interfaces for robotics*;

Keywords

Dekodierung, Detektion, Kodierte Marker, Patientenpositioniersystem, PPS, Roboter, Echtzeitüberwachung

1 Einleitung

Durch die stetige technische Entwicklung hat sich die Strahlentherapie als eine zentrale Säule der Krebstherapie etabliert [1]. Bei der Strahlentherapie werden Krebszellen mit Hilfe von Strahlung zerstört. In der Strahlentherapie existieren verschiedene Arten der Behandlung. Eine Form ist die Ionentherapie, bei der Protonen oder andere Ionen für die Bestrahlung verwendet werden. Die verwendeten Teilchen geben ihre Energie erst frei, wenn sie durch das Eindringen in den Körper auf eine gewisse Geschwindigkeit abgebremst werden. Bei der Ionentherapie wird somit umliegendes gesundes Gewebe besser geschont als bei anderen Formen, da die Behandlung nur lokal im Zentrum des zu behandelnden Gewebes erfolgt. Diese Therapieform eignet sich vor allem bei Tumoren, die gegenüber anderen Behandlungsformen

Betreuer Hochschule:	Prof. Dr.-Ing. Oliver Burgert Hochschule Reutlingen Oliver.Burgert@Reutlingen-University.de
Betreuer Firma	Tina Winter-Emden BEC GmbH T.Winter- Emden@b-e-c.de

Wissenschaftliche Vertiefungskonferenz
27. November 2019, Hochschule Reutlingen
Copyright 2019 Hannah Büchner

unempfindlich, schwer erreichbar oder nahe an Risikostrukturen gelegen sind.[1] Bei dieser Form der Therapie ist eine exakte Positionierung des Patienten direkt vor den Strahlenausgang notwendig, um eine optimale Zellzerstörung zu erreichen. Neben den Anforderungen an die Genauigkeit der Patientenpositionierung werden auch die Anforderungen an die abgegebene Dosis der Strahlung strenger [2]. Um die Dosis so effizient wie möglich zu nutzen, muss das zu bestrahlende Gewebe direkt in das Isozentrum des Strahls gerichtet werden [1]. Dafür ist eine hoch präzise Patientenpositionierung notwendig, die im Idealfall eine exakte Nachbildung des Aufbaus bei der Bildgebung für die Planung der Strahlentherapie darstellt [2]. Die Verwendung von Robotern ermöglicht solch eine Patientenpositionierung mit einer sehr hohen Genauigkeit [1]. Der Einsatz von diesen Positionierrobotern hat vor allem den Vorteil der hohen Reproduzierbarkeit. [2]

Die Positionen und Bewegungen des eingesetzten Roboters müssen dabei kontrolliert bzw. überwacht werden. Diese Kontrolle muss in Echtzeit möglich sein, um beispielsweise eine optimale Ausrichtung des Positionierroboters vor der Behandlung zu gewährleisten. Zusätzlich müssen während der Behandlung die Bewegungen des Patienten überwacht werden. Außerplanmäßige Bewegungen, wie beispielsweise ein Husten oder Niesen des Patienten, sollten schnellstmöglich zur Deaktivierung der Strahlenquelle führen, sodass so wenig gesundes Gewebe wie möglich beschädigt wird.

Im Rahmen dieser Arbeit wird ein robotergestütztes Patientenpositioniersystem untersucht, welches in der Strahlentherapie eingesetzt wird. Die Positionsgenauigkeit des Systems sowie dessen Bewegungen werden aktuell mittels kodierter Marker überwacht. Das dafür verwendete Trackingsystem soll durch ein weiterentwickeltes System ersetzt werden. Dafür wird ein Konzept entwickelt, welches eine Trennung der Dekodierung und Detektion der Marker für die Positionsverfolgung vorsieht. Unter Erhalt der hohen Präzision soll dieses Konzept die Verarbeitungszeit der Markeridentifikation verkürzen, sodass die Echtzeitfähigkeit der Reaktionen bei der Positions- und Stillstandsüberwachung verbessert werden kann.

Während in Kapitel 2 das markerbasierte Tracking sowie aktuelle Forschungsarbeiten in diesem Bereich vorgestellt werden, beschreibt Kapitel 3 das Konzept zur Trennung der Dekodierung und Detektion. Dabei wird der Ablauf des vorhandenen Prototyps sowie dessen Erweiterung erläutert. In Kapitel 4 werden die Ergebnisse der Trennung von Dekodierung und Detektion vorgestellt. Das Kapitel 5 diskutiert diese Ergebnisse, zieht ein Fazit und gibt einen Ausblick für künftige Arbeiten.

2 Stand der Technik

In diesem Kapitel wird ein Einblick in den Stand der Technik in Bezug auf das Tracking von Markern gegeben. Außerdem wird das Patienten-Positioniersystem *exacure* vorgestellt.

2.1 Tracking von Markern

Um die Position und Orientierung von Objekten über ein optisches Kamerasystem zu erfassen und über die Zeit hinweg zu verfolgen, gibt es unterschiedliche Ansätze. Dies kann sowohl markerlos sowie durch die künstliche Kennzeichnung von Objekten umgesetzt werden. Bei einem markerbasierten Ansatz werden an dem Objekt Marker angebracht, die in dem Kamerabild einfacher zu detektieren sind [3]. Dieser Ansatz benötigt im Vergleich zu einem markerlosen Ansatz zusätzliche Vorbereitungen der Objekte. Allerdings ist der markerbasierte Ansatz robuster und bietet durch eine einfachere Detektion im Bild eine

[1] https://www.krebsgesellschaft.de/onko-internetportal/basis-informationen-krebs/therapieformen/strahlentherapie-bei-krebs.html [Zuletzt geöffnet am 23.09.2019]

höhere Genauigkeit. Als künstliche Markierungen existieren bereits verschiedene Arten, die in der Robotik und in Augmented Reality (AR)-Anwendungen zum Einsatz kommen [4]. Neben aktiven, selbst leuchtenden LED-Markern existieren auch retroreflektierende Marker, welche von einer externen Quelle beleuchtet werden müssen sowie passive Marker, die zusätzliche Informationen kodieren können [5].

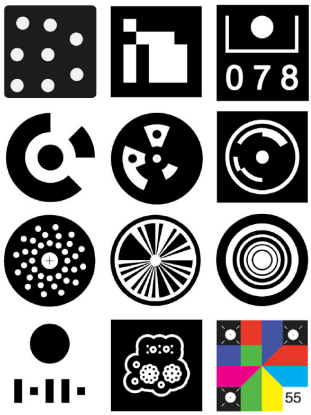


Abbildung 1: Verschiedene passive, kodierte Marker [6]

Die Abbildung 1 zeigt verschiedene passive Marker, die zusätzliche Informationen kodieren können. So kann in einem Marker eine eindeutige Identifikation hinterlegt werden, die wiederum beispielsweise einer 3D-Koordinate zugeordnet werden kann. Jeder Marker besteht hierbei meist aus zwei Komponenten. Einerseits aus dem Identifikationsteil, d.h. dem kodierten Teil und andererseits aus einem charakteristischen Merkmal für die einfachere Detektion. [6] [7]
In der Literatur werden außerhalb des medizinischen Kontexts quadratische Marker verwendet, um AR-Objekte im Kamerabild der realen Umgebung zu platzieren und diese zu verfolgen [8]. Auch im medizinischen Kontext wurden diese Marker verwendet. Beispielsweise wurden bereits Marker für robotergestützte, laparoskopische Eingriffe genutzt. Die Marker wurden dabei am Kopf des Operateurs befestigt, von einer Kamera identifiziert und zur Steuerung eines laparoskopischen Roboterarms genutzt [9]. Retroreflektierende Marker finden sich am häufigsten im klinischen Umfeld. Dort werden diese Marker meist für das Tracking von Operationsinstrumenten verwendet. Durch die Identifizierung können diese Instrumente in der Bildgebung visualisiert werden und dem Operateur nicht nur einen besseren Überblick verschaffen sondern auch beispielsweise die Positioniergenauigkeit verbessern. Mittels retroreflektierender Marker konnte so die Nadelsteuerung bei dem perkutanen Nierenzugang robotergestützt vorgenommen werden [10]. Auch bei der Patientenpositionierung werden optische Verfahren für die Positionsüberwachung verwendet. In [2] wurde die Patientenposition und die Bewegung der Patientencouch während der Behandlung mit einem externen Laser-Tracker und sechs retroreflektierenden Markern überwacht.
Das im Folgenden beschriebene System arbeitet mit solchen kodierten Markern für die Positionsüberwachung.

2.2 Exacure-System

Das in dieser Arbeit vorgestellte Konzept wurde für das Patienten-Positioniersystem *exacure*[2] entwickelt. Dieses System wird in der Strahlentherapie angewandt. Exacure besteht dabei aus einem Industrieroboter, welcher an der Decke montiert ist. Das medizinproduktzertifizierte Positionierungssystem besteht aus unterschiedlichen Modulen. Mit dem *exacontrol*-Modul kann das Komplettsystem mit allen Modulen verwaltet werden. Unter dem *examove*-Modul wird das eigentliche Positioniersystem basierend auf der

[2] https://www.exacure.com/information [Zuletzt geöffnet am 22.09.2019]

Abbildung 2: exacure-System mit den Modulen examove, exacouch, und exatrack[4]

KUKA-Robotertechnologie[3] verstanden. Das *exacouch*-Modul wird an dem Positioniersystem befestigt dient als Patientenliege für die Behandlung. Über das *exatrack-Modul* wird das optische Trackingsystem angeschlossen, sodass eine hochpräzise Positionierung bei der Behandlung ermöglicht wird. Für die Positionsüberwachung mittels optischem Kamerasystem werden an der Patientenliege (exacouch) angebrachte photogrammetrische kodierte Marker identifiziert. Dabei ist in diesen Markern eine Identifikationsnummer in Form eines Musters aus unterschiedlich angeordneten Kreisen hinterlegt, welches zu einer realen 3D-Koordinate zugeordnet werden kann. Die Abbildung 2 zeigt das exacure-System mit den Modulen examove, exacouch und exatrack.

Das exacure-System erreicht mit dem vorhanden Kamerasystem zur optischen Überwachung eine Positionsgenauigkeit von unter 0.5 mm und 0.1°. Es befindet sich unter anderem bei MedAustron in Österreich im Einsatz [1].

[3]https://www.kuka.com/de-de/branchen/healthcare/kuka-medical-robotics [Zuletzt geöffnet am 22.09.2019]

Die Ziffer in der oberen linken Ecke ist die ID des Pharmacodes. Der schwarze Kreis mit einer weißen Ellipse ist das charakteristische Merkmal, welches für die Erkennung verwendet wird. Die Abfolge verschieden breiter Balken auf der rechten Seite kodiert die Identifikationsnummer (77) des Codes, der eine reale 3D-Koordinate zugeordnet wird.

Abbildung 3: Beispiel eines Pharmacodes

3 Konzept

Das Konzept zur Trennung der Dekodierungs- und Detektionsphase von Markern für die Positionsüberwachung basiert auf einem Prototyp, der mit Hilfe der Matrox Imaging Library (MIL) entwickelt wurde. Die MIL bietet eine Reihe an Funktionalität für die Entwicklung von Anwendungen im Bereich der Bildverarbeitung und Bildanalyse. Sowohl die Bilderfassungs-, Verarbeitungs- als auch Analyseverfahren sind in Hinsicht auf Genauigkeit, Robustheit und Geschwindigkeit optimiert, um Anforderungen anspruchsvoller Anwendungen gerecht zu werden.[5]

Mit dem bereits existierenden Prototyp kann die drei-dimensionale (3D) Position und Ausrichtung eines Objekts bestimmt werden, indem auf dem Objekt angebrachte Referenzpunkte identifiziert werden. Diese Referenzpunkte sind als 3D-Koordinaten in sogenannten Pharmacodes kodiert. Abbildung 3 zeigt den Aufbau eines Pharmacodes. Eine detaillierte Erklärung des Dekodierablaufs ist in Kapitel 3.1 beschrieben, während in Kapitel 3.2 die Erweiterung des Prototyps durch die Trennung der Dekodierungs- und Detektionsphase erläutert wird.

[4]https://www.exacure.com/modules [Zuletzt geöffnet am 31.08.2019]

[5]https://www.matrox.com/imaging/de/products/software/mil/ [Zuletzt geöffnet am 14.09.2019]

3.1 Phase der Dekodierung

Um die Pharmacodes zu dekodieren und deren 3D-Koordinate für die Objektlokalisierung zu verwenden, wird das kalibrierte Kamerabild in einem ersten Schritt durch ein adaptives Threshold-Verfahren von einem Grauwertbild in ein Binärbild konvertiert. Bei dem adaptiven Threshold-Verfahren hängt der Wert für die Binarisierung für jeden Pixel von seiner Nachbarschaft ab. Der minimale sowie maximale Wert der Nachbarschaft eines Pixels wird über morphologische Operationen (Dilatation und Erosion) ermittelt und daraus ein Mittelwert gebildet. Dieser Mittelwert entscheidet, welcher Pixel als Vordergrund und welcher als Hintergrund eingestuft wird.

In einem weiteren Schritt werden jene verbundene Regionen von Pixeln (sogenannte Blobs) in dem binarisierten Bild gesucht. Dabei werden mit Hilfe der MIL nur jene Blobs selektiert, die den zuvor definierten Kriterien entsprechen. Zu diesen Kriterien zählen unter anderem eine gewisse Mindestgröße sowie eine möglichst kreisförmige Kontur. Außerdem muss die gefundene Region von Vordergrund-Pixeln eine Region von Hintergrund-Pixeln umschließen. Dadurch wird gewährleistet, dass jene Blobs ausgewählt werden, die einen weiteren Blob enthalten. Für diese Blobs werden anschließend die Mittelpunkte sowie deren Rechteck mit der kleinsten möglichen Fläche (BoundingBox) im Pixel-Koordinatensystem berechnet und abgespeichert.

Innerhalb dieser BoundingBox wird für jeden gefundenen Blob darauffolgend nach der eingeschlossenen, zusammenhängenden Region mit einer zuvor definierten Mindestgröße gesucht. Für jeden inneren Blob wird ebenfalls der Mittelpunkt im Pixel-Koordinatensystem berechnet und zwischengespeichert.

Da die Pharmacodes beliebig auf dem Objekt platziert werden können, ist es notwendig herauszufinden, ob sich der Code links oder rechts von dem gefundenen Blob befindet. Hierfür werden die Mittelpunkte jedes inneren (weißen) Blobs und des zugehörigen äußeren (schwarzen) Blobs verwendet. Da der innere Blob nicht zentral im äußeren Blob liegt, kann dadurch die Richtung bestimmt werden. Der Pharmacode wird somit in die Richtung vom Mittelpunkt des inneren Blobs zum Mittelpunkt des äußeren Blobs gesucht.

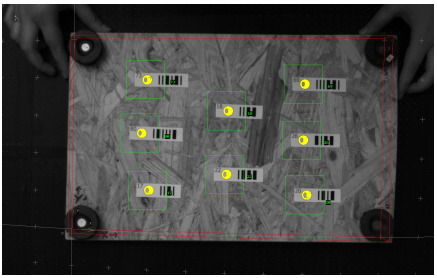
Die MIL verfügt über eine Reihe an Funktionalitäten, um ein-dimensionale (1D) und zwei-dimensionale (2D) Codes zu identifizieren. Dadurch ist auch die Dekodierung von Pharmacodes möglich. In dem Prototyp wird der Pharmacode mit Hilfe der MIL gelesen und die kodierte 3D-Koordinate im relativen Koordinatensystem, d.h. dem Patientencouch-Koordinatensystem, extrahiert.

Diese 3D-Koordinaten werden mit den Mittelpunkten der inneren Blobs im Pixel-Koordinatensystem für die Objektlokalisierung miteinander verglichen. Für die Lokalisierung müssen mindestens vier Pharmacodes identifiziert und deren Koordinaten extrahiert werden. Mit Hilfe der MIL kann das Kamerabild durch die Pixel- und zugehörigen relativen Koordinaten neu ausgerichtet werden. Dabei werden die Translation und Rotation zwischen der Kamera und dem Objekt berechnet.

Wenn das Objekt erfolgreich lokalisiert wurde, kann ein Polygon für eine visuelle Darstellung in das Kamerabild gezeichnet werden. Dafür werden die 3D-Koordinaten des Objekts vom relativen Koordinatensystem in das Pixelkoordinatensystem transformiert. Abbildung 4 zeigt die Dekodierung der Pharmacodes in einem Frame.

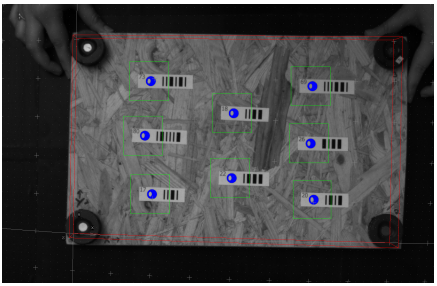
3.2 Phase der Detektion

Da der Vorgang der Dekodierung sehr zeitaufwändig und rechenintensiv ist, wird der Prototyp erweitert und eine Trennung der Dekodierung und Detektion vorgenommen. Der Ablauf für die Dekodierung ist gleich dem in Kapitel 3.1 beschriebenen Konzept. Dieser



Die gelben Kreise visualisieren die Blobs, die den Kriterien entsprechen. Die grüne Ziffer zeigt die ID eines Pharmacodes nach einer erfolgreichen Dekodierung. Die grünen Vierecke veranschaulichen die Region (Suchbox), in der die anschließende Detektion durchgeführt wird. Das rote Polygon zeigt die Lokalisierung des Objekts.

Abbildung 4: Visuelle Darstellung der Dekodierung



Die blauen Kreise visualisieren die Blobs, die den Kriterien entsprechen und innerhalb der durch die grünen Vierecke dargestellten Suchregionen gefunden wurden. Das rote Polygon zeigt die Lokalisierung des Objekts.

Abbildung 5: Visuelle Darstellung der Detektion

Vorgang wird auf das erste Frame angewendet. Zusätzlich zu dieser ersten Dekodierung wird an der Stelle, an der die 3D-Koordinate des Pharmacodes erfolgreich extrahiert wurde, eine dynamische Region um den Mittelpunkt des inneren Blobs berechnet. Die Region um jeden Blob ist entsprechend so groß gewählt, dass in jede Richtung mindestens zwei Blobs abgedeckt werden würden. Während bei der Dekodierung die erste Blobsuche auf dem gesamten Frame durchgeführt wird, wird bei der Detektion die Blobsuche für jeden Pharmacode auf die dynamisch berechnete Region (Suchbox) beschränkt. Dadurch kann gewährleistet werden, dass bei einer Bewegung des Objekts über die Frames hinweg der gefundene Blob tatsächlich zu dem zuvor dekodierten Pharmacode gehört. Durch die Suche in einer verkleinerten Region, sowie die Trennung der einmaligen Dekodierung und anschließenden fortlaufenden Detektion kann die Verarbeitungszeit reduziert werden. Abbildung 5 zeigt die Detektion der Pharmacodes in einem Frame.
Sofern das Objekt über die Zeit hinweg bewegt wird, so kann der Fall eintreten, dass während der Detektion ein Blob und damit ein dekodierter Pharmacode sowie dessen 3D-Koordinate verloren geht, d.h. nicht mehr gefunden wird. Um jedoch eine möglichste präzise Lokalisierung des Objekts zu gewährleisten, ist es wichtig, möglichst viele Pharmacodes und deren Koordinaten zu identifizieren. Im Falle eines Verlusts, sollte daher eine Neu-Dekodierung durchgeführt werden. Im Folgenden werden drei unterschiedliche Ansätze beschrieben, wie im Falle eines Verlusts effektiv vorgegangen werden sollte. Diese Ansätze wurden im Rahmen des Prototyps evaluiert. Die zugehörigen Ergebnisse sind in Tabelle 1 aufgeführt.

3.2.1 Ansatz 1 - Neu-Dekodierung nach einer bestimmten Anzahl an Frames: Der erste Ansatz dient der Prävention. Dabei wird nach einer bestimmten Anzahl an Frames eine Neu-Dekodierung durchgeführt, um einen möglichen Verlust von Blobs vorzubeugen. Für die Evaluation wurde alle zehn Frames eine Neu-Dekodierung vorgenommen. Die Zahl wurde an die Framerate des Bildmaterials angepasst.

3.2.2 Ansatz 2 - Neu-Dekodierung sobald weniger als vier Blobs gefunden wurden: Der

zweite Ansatz beschäftigt sich mit der Mindestanzahl an dekodierten Pharmacodes, um eine erfolgreiche Objektlokalisierung zu erhalten. Für die Lokalisierung werden mindestens vier Pharmacodes benötigt. Drei Punkte für die Bestimmung der Translation und Rotation und einen vierten Punkt, um die Genauigkeit zu erhöhen. Mit dem vierten Punkte kann die Standardabweichung als Gütekriterium berechnet werden. Mit dem Ansatz wird über die Frames hinweg gezählt, wie viele Pharmacodes bei der initialen Dekodierung identifiziert wurden. Bei jeder Detektion wird erneut gezählt, wie viele Pharmacodes bzw. deren Blobs erfolgreich gefunden wurden. Sobald die Zahl der detektierten Blobs kleiner als die Mindestanzahl an notwendigen Blobs ist, wird eine Neu-Dekodierung durchgeführt.

3.2.3 Ansatz 3 - Neu-Dekodierung sobald ein Blob verloren wurde: Bei dem dritten Ansatz wird bei jeder Detektion geprüft, ob die Anzahl an gefundenen Blobs mit der Anzahl an dekodierten Pharmacodes übereinstimmt. Sobald diese Bedingung nicht mehr zutrifft und dementsprechend ein Blob verloren wurde, wird eine Neu-Dekodierung durchgeführt.

Um eine Entscheidung zu treffen, welcher Ansatz der effektivste ist und damit die geringste Anzahl an Neu-Dekodierungen benötigt und trotzdem die Objektlokalisierung präzise und erfolgreich ist, wurden die Ansätze auf drei Videos sowie dem Livebild der Kamera angewandt und die Anzahl an Neu-Dekodierungen gezählt. Außerdem wurde das Verhalten der Objektlokalisierung und der Visualisierung des Polygons beobachtet. Tabelle 1 zeigt die Ergebnisse für das jeweilige Bildmaterial, sowie die Gesamtanzahl der Frames und die Anzahl der Neu-Dekodierungen. Unter Original wird das Beispiel-Video des Prototyps verstanden. Dieses Video besteht aus 176 Frames, hat eine Gesamtdauer von 15 Sekunden, eine Auflösung von 1280 x 960 Pixel und wurde mit einer Framerate von 11,6 fps aufgenommen. Video 1 und 2 sind eigens erstellte Videos mit einer Framerate von 10 fps und einer Auflösung von 2592 x 2048 Pixeln. Das erste Video enthält 182 Frames und dauert 18 Sekunden, das zweite Video umfasst 406 Frames und erstreckt sich über eine Dauer von 40 Sekunden.
Die Ergebnisse zeigen deutlich, dass der erste Ansatz in den meisten Fällen die höchste Anzahl an Neu-Dekodierungen aufweist. Zusätzlich war die Objektlokalisierung und Visualisierung weniger präzise als in den anderen Ansätzen. Der zweite Ansatz hat in drei von vier Fällen die geringste Anzahl ein Neu-Dekodierungen, allerdings war die Objektlokalisierung nicht mehr genau. Je weniger Codes gefunden wurden, desto schlechter wurde das Objekt lokalisiert. Der dritte Ansatz ist ein Mittel zwischen dem ersten und zweiten Ansatz. Die Anzahl der Neu-Dekodierung ist geringer als im ersten Ansatz, jedoch in den meisten Fällen höher als beim zweiten Ansatz. Die Objektlokalisierung war jedoch in allen Videos und auch im Kamera-Livebild am genausten. Daher wurde der dritte Ansatz für das weitere Vorgehen gewählt.

3.3 Rahmenbedingungen

Die Trennung der Dekodierung und Detektion von Markern an einem Patientenpositionier-Roboter erfordert die Verifizierung der Zugehörigkeit von dekodiertem Pharmacode und gesuchtem Blob. Diese Verifizierung wird durch einige Rahmenbedingungen ermöglicht. Dadurch, dass die berechnete Suchbox für den Blob bei der Detektion auf den Koordinaten des dekodierten Pharmacodes basiert, kann davon ausgegangen werden, dass der gefundene Blob tatsächlich zu dem dekodierten Pharmacode gehört. Weiterhin gilt die Norm für das in Kapitel 2.2 beschriebene Patientenpositioniersystem:
Die Norm *IEC 60601-2-64: Besondere Anforderungen an die Grundsicherheit und grundlegende Leistung von medizinischen elektrischen*

Tabelle 1: Ergebnisse der unterschiedlichen Ansätze für das Vorgehen bei dem Verlust eines Blobs (Markers)

Ansatz	Bildmaterial	Anzahl der Gesamtframes	Anzahl der notwendigen Detektion
1	Original	176	18
2	Original	176	16
3	Original	176	33
1	Video 1	182	19
2	Video 1	182	3
3	Video 1	182	14
1	Video 2	406	41
2	Video 2	406	4
3	Video 2	406	25
1	Kamera	298	30
2	Kamera	234	23
3	Kamera	224	12

Geräten mit Lichtionenstrahl gibt bei einer linearen Bewegung eine Geschwindigkeitsbregrenzung vor. Dabei darf keine Geschwindigkeit schneller als 100 $\frac{mm}{s}$ sein. [11]
Das Patientenpositioniersystem mit einem Roboter von KUKA gibt einen Sensortakt von 4 ms vor.
Die verwendete Kamera (CMOS-Flächenkamera Norite N-5A100[6]) mit dem Objektiv (KOWA LM16SC[7]) besitzt einen horizontalen Blickwinkel von ca. 34,15°und einen vertikalen Blickwinkel von ca. 42,48°. Mit einem Sensortakt von 4 ms benötigt die Kamera eine Framerate von 250 fps. Diese Framerate wird von der Kamera unterstützt. Bei einem minimalen Abstand von 80 cm und einem maximalen Abstand von 120 cm zwischen Objekt und Kamera sowie der Framerate von 250 fps unterscheidet sich eine Bewegung eines Objekts um wenige Pixel im Bild (ca. 1-3 Pixel).

[6]https://www.adimec.com/cameras/machine-vision-cameras/norite-series/ [Zuletzt geöffnet am 13.09.2019]
[7]https://lenses.kowa-usa.com/6mp-plus-sc-series/1240-lm16sc.html [Zuletzt geöffnet am 13.09.2019]

Durch diese Rahmenbedingungen, den geringen Unterschied einer Bewegung zwischen zwei Frames und der dynamischen Berechnung der Suchbox ist es somit gewährleistet, dass der detektierte Blob zu dem zuvor dekodierten Pharmacode gehört.

4 Ergebnisse

Um eine Aussage treffen zu können, ob die Trennung der Dekodierungs- und Detektionsphase zu einer Reduzierung der Verarbeitungszeit pro Frame führt, wurden Messreichen mit unterschiedlichem Bildmaterial aufgezeichnet. Neben den in Kapitel 3.2.3 eigens aufgenommenen Videos mit einer Framerate von 10 fps wurde zusätzlich die Framerate der verwendeten Kamera auf 125 fps erhöht und ebenfalls ein Video aufgenommen. Das Video 3 hat eine geringere Auflösung von 1296 auf 1024 Pixeln und beinhaltet 2612 Frames mit einer Gesamtdauer von 20 Sekunden. In den aufgenommenen Videos wurde darauf geachtet, dass die Bewegungen des Objekt möglichst den realen Bedingungen entsprechen. Somit wurde das Objekt bei einer Framerate von 125 fps ca. 5 $\frac{cm}{s}$ bewegt.
Die gemessenen Zeiten beschreiben die CPU-Zeit des Rechners, die bei dem Ablauf der

Algorithmen für die Dekodierung bzw. Detektion vergangen ist. Für die Zeitmessungen wurden keine Visualisierungen außer dem Objektpolygon gezeichnet, da diese zusätzliche Zeit in Anspruch nehmen. Die angegebenen Zeiten zeigen die Verarbeitungszeit pro Frame im Durchschnitt.
Für die Zeitmessungen der Videos wurde ein Intel (R) Core (TM) i7-4710MQ , CPU @ 2.50 GHz Prozessor und ein Arbeitsspeicher von 8 GB RAM verwendet. Die Zeitmessungen des Livebilds der Kamera wurden auf einem Rechner mit einem Intel (R) Core (TM) i5-9600K, CPU @ 3.70 GHz Prozessor und einem Arbeitsspeicher von 16 GB RAM durchgeführt. Die Tabellen 2 und 3 zeigen die Ergebnisse der Zeitmessungen. Es werden sowohl das betrachtete Bildmaterial, als auch die Anzahl der Gesamtframes, die jeweilige Anzahl der dekodierten und detektierten Frames sowie die Verarbeitungsdauer im Durchschnitt pro Frame für die Dekodierung und Detektion in Millisekunden dargestellt.
Die Zeiten für die Dekodierung variieren zwischen 14.85 und 32.48 Millisekunden. Die Zeiten für die Detektion dagegen schwanken zwischen 2.43 und 9.59 Millisekunden.

5 Fazit

In den Ergebnissen in Kapitel 4 ist zu sehen, dass sich die Zeiten für die Dekodierung und Detektion stark voneinander unterscheiden. Die durchschnittliche Zeit für die Detektion ist in den meisten Fällen ein Drittel der Zeit, die für die Dekodierung vergeht. Durch diese Reduzierung der Zeit kann auch von einer Zeitminimierung des Gesamtablaufs, dem Verfolgen der Pharmacodes über die Zeit hinweg, ausgegangen werden.
Durch die in Kapitel 3.3 beschriebene Geschwindigkeit sowie dem Sensortakt des Roboters wird von einer Framerate der Kamera von 250 fps ausgegangen. Dies konnte im Rahmen des Prototyps jedoch nicht getestet werden. Daher kann keine Aussage darüber getroffen werden, ob das vorgestellte Konzept die Anforderungen der Echtzeitüberwachung erfüllt. Jedoch zeigt das Konzept, dass eine Zeitminimierung möglich und die Echtzeitüberwachung dadurch realisierbar ist.
Um die Zeit weiter zu reduzieren könnten zusätzliche Messungen auf einem leistungsfähigeren Rechner durchgeführt werden. Um nicht nur bei der Detektion Zeit einzusparen, könnten Recherchen bezüglich anderer kodierter Marker getätigt werden. Die Verwendung von einem anderen Dekodieralgorithmus könnte unter Umständen ebenfalls zu einer Einsparung der Verarbeitungszeit führen.
Die Trennung der Dekodierungs- und Detektionsphase ermöglicht weiterhin die Parallelisierung des Suchalgorithmus. Während im aktuellen Prototyp bei der Detektion die Blobs nacheinander in der dafür definierten Region gesucht werden, könnte dies aufgeteilt und parallel abgearbeitet werden. Dies lässt vermuten, eine weitere Einsparung der Zeit zu erreichen.
Es konnte somit gezeigt werden, dass die Trennung der Dekodierung und Detektion von Markern erfolgreich für die Positionsüberwachung eines Patientenpositionierroboters verwendet werden kann. Eine Aussage über die Verwendung in Echtzeit ist nicht möglich, da hierfür weitere Testmessungen durchgeführt und evaluiert werden müssen. Darüber hinaus sollten außerdem auch Messungen bezüglich der Positionsgenauigkeit durchgeführt werden.

Literatur

[1] H. Choi, J.-O. Park, S. Y. Ko, S. Park, S. Cho, W.-G. Jung, Y. K. Park und J. S. Kang, "Development of a robotic patient positioning system with a wide beam-angle range for fixed-beam particle therapy", *Journal of the Korean Physical Society*, Jg. 69, Nr. 8, S. 1271–1285, Okt. 2016. DOI: 10.3938/jkps.69.1271.

[2] O. Nairz, M. Winter, P. Heeg und O. Jäkel, "Accuracy of robotic patient positioners used in ion beam therapy", *Radiation*

Tabelle 2: Ergebnisse mit einer Framerate von 10 fps

Bildmaterial	Anzahl der Gesamtframes	Anzahl dekodierter Frames	Anzahl detektierter Frames	Dekodierung (in ms)	Detektion (in ms)
Original	176	33	143	14.85	2.43
Video 1	182	14	168	30.07	9.59
Video 2	406	25	381	32.48	9.56
Kamera	301	2	299	28.50	5.82

Tabelle 3: Eregbnisse mit einer Framerate von 125 fps

Bildmaterial	Anzahl der Gesamtframes	Anzahl dekodierter Frames	Anzahl detektierter Frames	Dekodierung (in ms)	Detektion (in ms)
Video 3	2612	97	2515	18.02	3.74
Kamera	1001	2	999	15.50	2.66

Oncology, Jg. 8, Mai 2013. DOI: 10.1186/1748-717X-8-124.

[3] A. Köpfle, "Ein modulares optisches Trackingsystem für medizintechnische Anwendungen: integrierte Datenflussarchitektur in Hard- und Software und Applikationsframework", Deutsch, Diss., Mannheim, 2012. Adresse: https://madoc.bib.uni-mannheim.de/34781/.

[4] A. Sagitov, K. Shabalina, R. Lavrenov und E. Magid, "Comparing fiducial marker systems in the presence of occlusion", in *2017 International Conference on Mechanical, System and Control Engineering (ICMSC)*, Mai 2017, S. 377–382. DOI: 10.1109/ICMSC.2017.7959505.

[5] T. Luhmann, S. Robson, S. Kyle und J. Boehm, *Close-Range Photogrammetry and 3D Imaging*, 2. Auflage. Berlin: Walter de Gruyter, 2014, ISBN: 978-3-11-030269-1.

[6] S. Tushev, B. Sukhovilov und E. Sartasov, "Architecture of industrial close-range photogrammetric system with multi- functional coded targets", in *2017 2nd International Ural Conference on Measurements (UralCon)*, Okt. 2017, S. 435–442. DOI: 10.1109/URALCON.2017.8120748.

[7] L. Calvet, P. Gurdjos, C. Griwodz und S. Gasparini, "Detection and Accurate Localization of Circular Fiducials under Highly Challenging Conditions", in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Juni 2016, S. 562–570. DOI: 10.1109/CVPR.2016.67.

[8] Y. Gao, H. Wang und X. Bian, "Marker tracking for video-based augmented reality", in *2016 International Conference on Machine Learning and Cybernetics (ICMLC)*, Bd. 2, Juli 2016, S. 928–932. DOI: 10.1109/ICMLC.2016.7873011.

[9] M. Minamoto, M. Sato, T. Kanno und K. Kawashima, "Tele-Operation of Robot by Image Processing of Markers Attached to Operator's Head", in *2018 IEEE International Conference on Mechatronics and Automation (ICMA)*, Aug. 2018, S. 2414–2419. DOI: 10.1109/ICMA.2018.8484620.

[10] D. Zhang, Z. Li, K. Chen, J. Xiong, X. Zhang und L. Wang, "An optical tracker based robot registration and servoing method for ultrasound guided percutaneous renal access", *BioMedical Engineering OnLine*, Jg. 12, Mai 2013. DOI: 10.1186/1475-925X-12-47.

[11] "Medical electrical equipment – Part 2-64: Particular requirements for the basic safety and essential performance of light ion beam medical electrical equipment", Norm 60601 Part 2-64, 2014.

Perspektivische Transformation zur Kalibrierung von Verkehrsüberwachungskameras

David Salb
Hochschule Reutlingen
David_Julian.Salb@Student.Reutlingen-University.de

Abstract

Um eine möglichst umfassende Datengrundlage zur Analyse von Verkehrsinformationen zu gewährleisten, ist die Geschwindigkeitsmessung einzelner Fahrzeuge, in unterschiedlichen Spuren, aus dem Bildmaterial öffentlicher Verkehrsüberwachungskameras, erforderlich. Diese Art von Bildverarbeitung setzt ein kalibriertes Kamerasystem voraus, um die perspektivische Verzerrung des Kamerabildes auszugleichen. Im Zuge dieser Arbeit wird ein Kalibrierungsverfahren vorgestellt, welches es dem Anwender ermöglicht, eine typische Verkehrsüberwachungszene durch Eingabe weniger Merkmale zu kalibrieren. Eine abschließende Evaluation bestätigt die Eignung der Methode.

CCS Concepts

• **Computing methodologies** → **Camera calibration**; • **Human-centered computing** → Interaction design;

Keywords

Kamerakalibirierung, Verkehrsüberwachung, Perspektivische Transformation, Bildverarbeitung, OpenCV

Betreuer Hochschule: Prof. Dr. Curio
Hochschule Reutlingen
Cristobal.Curio@Reutlingen-University.de

Wissenschaftliche Vertiefungskonferenz
27. November 2019, Hochschule Reutlingen
Copyright 2019 David Salb

1 Einleitung

Die Analyse von Bildmaterial zur Erkennung der Verkehrsdichte stellt einen zentralen Anwendungsfall für die Verkehrsüberwachung dar. So kann die Entstehung von Staus besser verstanden und vorhergesagt werden. Existierende Ansätze zur Staumessung reichen vom Vergleich des aktuellen Überwachungsbildes mit einem verkehrsarmen Referenzbild [1] über das Zählen von bewegten Objekten [2] bis hin zur Geschwindigkeitserkennung von einzeln erkannten und getrackten Fahrzeugen [3]–[5].

Verfahren der letztgenannten Kategorie setzen ein kalibriertes Kamerasystem voraus, da sie für die Geschwindigkeitsberechnung eine Wegmessung benötigen.

Diese ist nur dann möglich, wenn die erkannten Fahrzeug-Merkmale einem räumlichen Kontext zugeordnet werden können. Da jedoch ein Bildsensor die Projektion des dreidimensionalen Raums abbildet und die so erzeugte Projektionsebene in den meisten Fällen nicht parallel zur Straßenebene liegt, unterliegt jedes erkannte Merkmal einer perspektivischen Verzerrung. So wird ein Fahrzeug mit konstanter Geschwindigkeit, welches sich von der Kamera entfernt, in der Projektion langsamer umso näher es sich Richtung Horizont bewegt. Um also die reale Geschwindigkeit zu erfassen, gilt es diesen perspektivischen Effekt auszugleichen.

Um den realen, zurückgelegten Weg eines gegebenen Merkmals messen zu können, muss das Bezugssystem zudem skaliert werden. Das Fahrzeug ließe sich sonst nicht von einem

Spielzeugauto unterscheiden. Um diese Größeninformation zu erhalten, muss eine bekannte Strecke eingegeben [6] oder automatisch erkannt werden [7].
Im Zuge dieser Arbeit wird ein Kalibrierungsverfahren vorgestellt, welches es dem Anwender ermöglicht, eine typische Verkehrsüberwachungsszene durch Eingabe weniger Merkmale zu kalibrieren (Kapitel 2). Für die Skalierung bietet das System einen Eingabedialog, welcher es ermöglicht, die Abstände zu markanten Punkten zu referenzieren und zu bemaßen (Kapitel 2). Abschließend wird der vorgestellte Proof of Concept-Ansatz mittels aufgenommenen Realdaten evaluiert (Kapitel 3). Das System dient als Grundlage für eine, im Rahmen einer weiterführenden Masterthesis durchgeführte, Staumessung.

2 Kalibrierungsmethodik

Um die in Abbildung 1 dargestellte perspektivische Verzerrung ausgleichen zu können, muss die Rückprojektion der Bildebene auf die Straßenebene mathematisch beschrieben werden. Dies wird mittels einer perspektivischen Transformation erreicht. Für diese Transformation werden Punktkorrespondenzen aus dem Bild und dem Zielkoordinatensystem genutzt, um eine $3x3$ Transformationsmatrix (M) zu generieren[1]. Mithilfe dieser Matrix und der in Gleichung 1 dargestellten Formel, lässt sich ein beliebiger Punkt des Ursprungsbildes in ein entzerrtes Zielkoordinatensystem transformieren.

$$x_{dst} = \frac{M_{1,1} * x_{src} + M_{1,2} * y_{src} + M_{1,3}}{M_{3,1} * x_{src} + M_{3,2} * y_{src} + M_{3,3}}$$
$$y_{dst} = \frac{M_{2,1} * x_{src} + M_{2,2} * y_{src} + M_{2,3}}{M_{3,1} * x_{src} + M_{3,2} * y_{src} + M_{3,3}} \quad (1)$$

Wendet man die perspektivische Transformation auf das gesamte Bild, also für jedes Pixel des Ursprungsrechtecks, an, wird die Straßenebene in die Vogelperspektive transformiert[2]; dargestellt in Abb. 4.
Für die Erzeugung dieser Punktkorrespondenzen ist ein Straßenabschnitt gut geeignet, da er ein Rechteck und somit eine einfache geometrische Form beschreibt. Jeder der vier Eckpunte des Straßenabschnitts stellt eine Punktkorrespondenz zu dem perspektivisch entzerrten Zielkoordinatensytem dar. Die Konstruktion dieser Punktkorrespondenzen in der Straßenebene wird in Kapitel 2.1 erklärt. Um auf die reale Geschwindigkeit schließen zu können muss das Zielkoordinatensytem noch für die Wegmessung skaliert werden. Zu diesem Zweck kann der Benutzer mittels eines Bemaßungsdialogs (Kapitel 2.2) eine reale, bekannte Strecke referenzieren.

2.1 Perspektivisches Rechteck konstruieren

In einem typischen Verkehrsüberwachungsbild lassen sich aus der Szene die in Abbildung 1 beschriebenen Fluchtpunkte ableiten. Der erste Fluchtpunkt **(FP1)** ergibt sich aus dem Schnittpunkt der paralell zur Fahrtrichtung verlaufenden Linien (Abb. 2). Der zweite Fluchtpunkt **(FP2)** konstruiert sich analog aus Kanten, welche orthogonal zur Fahrbahnrichtung sowie paralell zur Fahrbahnebene liegen - beispielsweise Straßenschilder, Fahrzeugkanten oder Haltelinien (Abb. 2).
Das System bietet einen Eingabedialog für die Fluchtpunkt-Identifikation. Dieser ermöglicht es, mehrere, für den jeweiligen Fluchtpunkt ausschlaggebende Linien, zu markieren; in Abb. 5 für den ersten Fluchtpunkt dargestellt. Das System errechnet anschließend den Fluchtpunkt aus dem Median der Schnittpunkte aller eingegebenen Linien. Sind beide Fluchtpunkte erfasst, lässt sich ein Polygon konstruieren, welches der projizierten Lage

[1] **OpenCV, getPerspectiveTransform**
https://docs.opencv.org/2.4/modules/imgproc/doc/geometric_transformations.html#getperspectivetransform
zuletzt geöffnet am 09. Okt. 2019

[2] **OpenCV, warpPerspective**
https://docs.opencv.org/2.4/modules/imgproc/doc/geometric_transformations.html#warpperspective
zuletzt geöffnet am 09. Okt. 2019

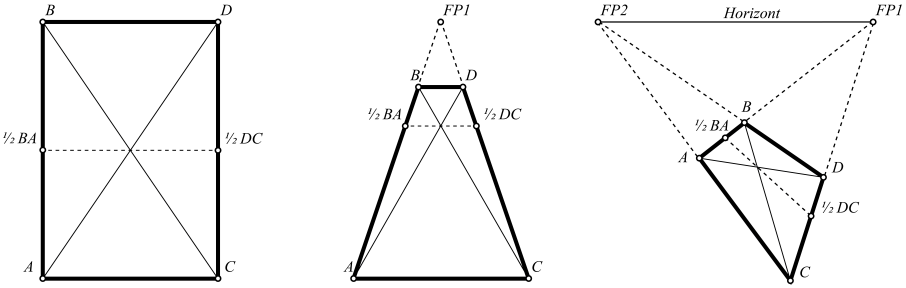


Abbildung 1: Perspektivische Verzerrung eines Rechtecks | **Links:** Ein Rechteck aus der Draufsicht, ohne perspektivische Verzerrung. Die Mitte des Rechtecks bildet der Schnittpunkt aus den Verbindungsgeraden der gegenüberliegenden Punkte **A** & **D** sowie **B** & **C**. **Mitte:** Das Rechteck wird entgegen eines Fluchtpunktes (**FP1**) verzerrt. Die Platzierung der gestrichelten Mittellinie bleibt proportional erhalten. **Rechts:** Das Prinzip gilt auch für Szenen, in denen der zweite Fluchtpunkt (**FP2**) nicht im Unendlichen liegt, sich das Rechteck also in zwei Fluchtachsen verzerrt.

Abbildung 2: Fluchtlinien | **Durchgezogen:** Linien, welche sich im 1. Fluchtpunkt schneiden. **Gestrichelt:** Linien, welche sich im 2. Fluchtpunkt schneiden.

Abbildung 3: Perspektivisches Rechteck im Ursprungsbild | Mit Hilfe der erkannten Fluchtpunkte kann ein Rechteck zur Gewinnung von Punktkorrespondenzen konstruiert werden

eines Rechtecks in einer horizontalen Ebene entspricht **(Abb. 3)**. Dieses Rechteck lässt sich durch Eingabe von zwei Punkten konstruieren, welche in der Projektion diagonal gegenüber liegen. Dafür wird zunächst ein Punkt am Fahrbahnrand eingegeben **(A)**. Im nächsten Schritt wird ein zweiter Punkt gewählt, welcher am gegenüberliegenden Fahrbahnrand und am anderen Ende des Kalibrierungsbereichs liegt **(D)**. Die übrigen Eckpunkte des Polygons (**B** und **C**) sind durch die Schnittpunkte der verlängerten Verbindungsgeraden von **A** und **B** zu jeweils **FP1** und **FP2** gegeben:

$$B = Schnittpunkt(FP_1A, FP_2D) \quad (2)$$

$$C = Schnittpunkt(FP_2A, FP_1D) \quad (3)$$

2.2 Zielkkordinatensystem skalieren

Um das Zielkoordinatensystem zu skalieren und so eine tatsächliche Wegmessung für die gesamte Straßenebene zu erreichen, müssen bekannte Strecken im Kamerabild vom Anwender erkannt und referenziert werden. Um

Abbildung 4: Kamerabild, transformiert in die Vogelperspektive | Mittels Gleichung 1 lässt sich jedes Pixel in das Zielkoordinatensystem transformieren, die Straßenebene wird in die Vogelperspektive transformiert

Abbildung 5: Eingabeverfahren zur Fluchtpunktfindung | Es werden mindestens zwei Fluchtlinien markiert. Der Fluchtpunkt (weißes Kreuz) ergibt sich aus dem Median der Schnittpunkte (schwarze Punkte) aller eingegebenen Fluchtlinien (gestrichelte Linien).

die Proportionen des gesamten Zielkoordinatensystems erfassen zu können, muss die reale Distanz von mindestens einer X-, als auch einer Y-Komponente im Zielkoordinatensytem bekannt sein.
Zu diesem Zweck bietet das System eine Eingabefunktion, welche es ermöglicht, im perspektvischen Rechteck mehrere Bemaßungslinien zu positionieren und deren Abmessungen einzugeben. Dabei wird mit der Längeneingabe des gemessenen Straßenabschnitts die Y-Komponente des Zielkoordinatensystems skaliert und mit Eingabe der Fahrbahnbreite die X-Komponente skaliert.
Um den Eingabefehler zu reduzieren, ist die Bemaßung mehrerer Distanzen möglich. In diesem Fall wird der Mittelwert für die X- und Y-Komponenten herangezogen.

2.3 Geschwindigkeitsberechnung

Mit der erzeugten Transformationsmatrix lässt sich mittels Gleichung 1 ein im Ursprungsbild erkanntes Merkmal (M) in das Zielkoordinatensystem transformieren (m).
Der Weg w den ein erkanntes Merkmal M zwischen zwei Videobildern zurücklegt, ergibt sich aus dem euklidischen Abstand des perspektivisch transformierten Merkmals m in den aufeinanderfolgenden Videobildern, m_{T-1} und m_T. Der zuvor errechnete Skalierungsfaktor s_x, s_y geht folgendermaßen in die Distanzrechnung ein:

$$w(m_{T-1}, m_T) = \sqrt{((m_{T,x} - m_{T-1,x})s_x)^2 + ((m_{T,y} - m_{T-1,y})s_y)^2} \quad (4)$$

Um die resultierende Geschwindigkeit v des Merkmals zwischen zwei Bildern (in Metern pro Sekunde $[\frac{m}{s}]$) zu errechnen, wird der zurückgelegte Weg w (in Metern [m]) mit der Bildwiederholungsfrequenz $f[\frac{1}{s}]$ multipliziert:

$$v(m_{T-1}, m_T) = w(m_{T-1}, m_T)f \quad (5)$$

3 Experimentelle Evaluation

Um die Kalibrierungsmethodik zu testen, wurde eine Kantenextraktion nach Shi & Tomasi [8] in Verbindung mit einem Lucas Kanade Merkmalstracker [9] implementiert[3]. Die Transformation (Gleichung 1) und Errechnung der Ortsänderung (Gleichungen 4 - 5) erfolgten dabei wie beschrieben.

[3] **OpenCV, Tutorial Optical Flow**
https://docs.opencv.org/3.4/d4/dee/tutorial_optical_flow.html
zuletzt geöffnet am 09. Okt. 2019

3.1 Testbedingungen

Für eine Validierung des Systems wurden die Live-Videofeeds öffentlich zugänglicher Verkehrskameras genutzt. Diese wurden von einem Erprobungsfahrzeug durchfahren, welches einen GPS-Monitor zur Messung der Kontrollgeschwindigkeit an Bord hatte. Die Test-Route umfasste fünf Kameras, von denen jedoch eine kurzfristig ihren Blickwinkel auf eine andere Straße lenkte und eine andere durch starken Wind zu sehr wackelte. Es verblieben drei Kameraaufnahmen, deren Material sich eignete, um eine Validierung durchzuführen.

Jede der drei Kameras wies die für die US-Verkehrsüberwachung verbreitete Auflösung von 480 x 320 Pixel und eine Bildwiederholungsfrequenz von 15 fps auf.

Aufgrund der höheren Pixeldichte wurde das Messfenster im unteren Bildbereich gewählt. Zudem wurde auf einen möglichst geraden Streckenverlauf geachtet. Testgegenstand war die Genauigkeit der Geschwindigkeitsmessung sowie die Diskrepanz abhängig von der Fahrspur. Aus diesem Grund wurde im ersten Test die zur Kamera nächstgelegene Spur gewählt, für die zweite und dritte Kamera wurde eine mittige Fahrbahnspur verwendet. Das transformierte Bild wurde mithilfe der Fahrbahnmarkierungen bemaßt (Abb. 8). Der Markierungsabstand wurde mittels der Google-Earth Linealfunktion am Testort gemessen (14,3 Meter; vom Ende des einen -, zum Ende des nächsten Sprutrennstreifens).

3.2 Ergebnisse

Abbildung 6 zeigt das Testfahrzeug mit erkanntem Merkmal (X). In Abbildung 7 sind das in die Vogelperspektive transformierte Video-Frame sowie die transformierten Merkmale dargestellt. Neben dem Testfahrzeug lassen sich hier auch die erkannten Merkmale von einem anderen Fahrzeug (*) sowie eine Fehlmessung (+) erkennen. Die errechnete Geschwindigkeit ist über ~25 Frames aufgezeichnet und für jedes Merkmal in Abbildung 9 aufgetragen.

Abbildung 6: Erkannte Merkmale in der Verkehrsszene | **X**: Im Testfahrzeug erkanntens Merkmal. *: Fremdes Fahrzeug. +: fehlerhafte Erkennung.

Die Kontrollgeschwindigkeit aus dem GPS-Log beträgt für diesen Test 85 km/h, mittels gestrichelter Linie dargestellt. Die ermittelte Geschwindigkeit des Testfahrzeugs ist als dicke Linie dargestellt und schwankt zwischen 80 und 100 km/h.

Für die anderen beiden Tests beträgt die Kontrollgeschwindigkeit 82 km/h und 95 km/h. Gemessen wurden hier 80 ± 7 km/h (Abb. 10) und 95 ± 15 km/h (Abb. 11). Die Ergebnisse sind reproduzierbar.

3.3 Diskussion und Ausblick

Die hohe Fluktuation der gemessenen Geschwindigkeiten ist auf die relativ simple Merkmalserkennung (Kantenextraktion [8]) in Verbindung mit einer niedrigen Videoauflösung von 480 x 320 Pixel zurückzuführen. Die Erfahrung zeigt, dass die Kantenerkennung, insbesondere für dunkle Fahrzeuge, mangels kontrastreicher Kanten, anfälliger für sprunghafte Verschiebungen des erkannten Merkmals sind. Zudem werden die Feature mithilfe dieser Methode selten in direkter Nähe der Fahrbahnebene (Reifen, Schatten) sondern meist an Fensterkanten oder Ähnlichem erkannt. Da das System jedoch *nur* für die Straßenebene kalibiriert ist, wird die Geschwindigkeit

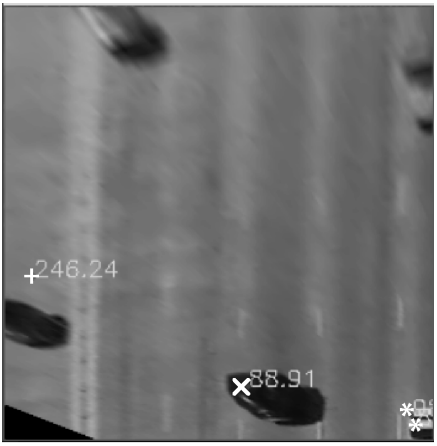


Abbildung 7: Perspektivisch transformierte Merkmale | **X**: Im Testfahrzeug erkanntes Merkmal. *: Fremdes Fahrzeug (unten rechts). +: Fehlerhafte Erkennung.

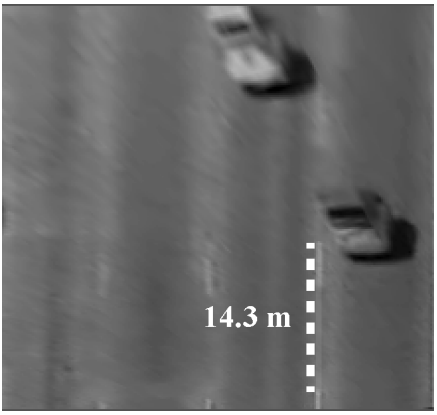


Abbildung 8: Bemaßungseingabe im transformierten Bild | Es wird eine bekannte Strecke markiert und eingegeben. In diesem Beispiel wird der Abstand der Fahrbahnmarkierungen (14.3 m) genutzt.

des Merkmals umso stärker verfälscht, je höher es über der Straßenebene liegt. Der Effekt, den die Bildwiederholungsfrequenz auf das Erkennungsverhalten der Merkmale hat, bleibt noch zu testen.

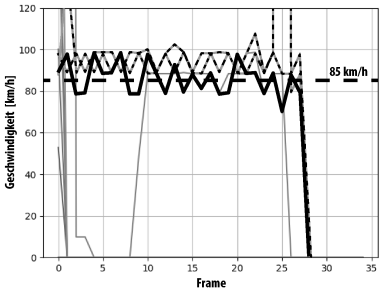


Abbildung 9: Kontrollgeschwindigkeit bei 85km/h | Geschwindigkeitsverlauf der getrackten Merkmale. Fett gestrichelt; die Kontrollgeschwindigkeit von 85 km/h. Fett; das erkannte Testfahrzeug mit ca 90 km/h (Entspricht X aus Abb. 6 und 7). Die grau und schwarz gestrichelten Linien gehören zu einem einzigen fremden Fahrzeug und liegen in einem Band von ca. 10 km/h. Die fehlerhafte Erkennung ist in grau dargestellt.

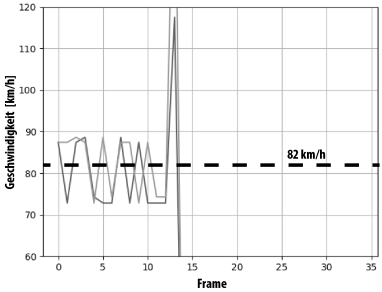


Abbildung 10: Kontrollgeschwindigkeit bei 82km/h | Beide Linien gehören zum Testfahrzeug. Die gemessene Geschwindigkeit beträgt ~80 ± 7 km/h. Die Geschwindigkeiten konnten nur über ~12 Frames verfolgt werden, bevor die Merkmale das Messfenster verließen.

Der Fokus dieser Arbeit liegt in der Kalibrierungsmethodik. Für diese bestätigen die Daten eine Nutzbarkeit des Verfahrens zur Kalibrierung der meisten Straßenszenen. Die Grenzen dieser Methode liegen bei starken Kurven oder stark gewölbten Straßen. In solchen Fällen gestaltet sich die Konstruktion

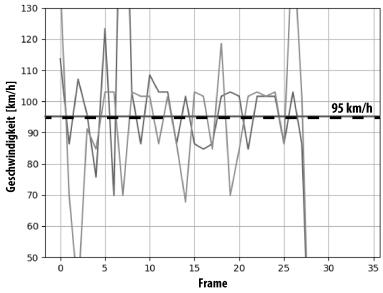


Abbildung 11: Kontrollgeschwindigkeit bei 95km/h | Beide Linien, gehören zum Testfahrzeug. Sieht man von stärkeren Ausreißern ab, bewegt sich die erkannte Geschwindigkeit bei 95 ± 15 km/h.

von Punktkorrespondenzen für die perspektivische Transformation schwierig, da sich die Fluchtpunkte für verschiedene Streckenabschnitte zu stark unterscheiden. Eine Lösung für dieses Problem ist ein kleineres Messfenster. Dies setzt jedoch ebenfalls eine robustere Merkmalserkennung voraus, da Fluktuationen über den geringeren Messzeitraum weniger gut zu filtern sind.

Aus der Vorbereitung und Implementierung der hier behandelten Kalibrierungsmethodik, ergeben sich bereits eine Reihe wichtiger Einblicke und Überlegungen für die weitere Bearbeitung und Verbesserung der Merkmalserkennung. Diese wird im Rahmen einer weiterführenden Masterthesis zur Staumessung eingehender untersucht.

Literatur

[1] D. Sitaram und N. Padmanabha, "Still image processing techniques for intelligent traffic monitoring", in *2015 Third International Conference on Image Information Processing (ICIIP)*, Dez. 2015, S. 252–255. DOI: 10.1109/ICIIP.2015.7414775.

[2] N. Abbas, M. Tayyab und M. Qadri, "Real Time Traffic Density Count using Image Processing", *International Journal of Computer Applications*, Jg. 83, S. 16–19, Dez. 2013. DOI: 10.5120/14476-2736.

[3] M. Dubská, A. Herout und J. Sochor, "Automatic Camera Calibration for Traffic Understanding.", in *BMVC*, Bd. 4, 2014, S. 8.

[4] P. Giannakeris, V. Kaltsa, K. Avgerinakis, A. Briassouli, S. Vrochidis und Y. Kompatsiaris, "Speed Estimation and Abnormality Detection from Surveillance Cameras", *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, S. 93–936, 2018.

[5] T. N. Schoepflin und D. J. Dailey, "Dynamic camera calibration of roadside traffic management cameras for vehicle speed estimation", *IEEE Transactions on Intelligent Transportation Systems*, Jg. 4, Nr. 2, S. 90–98, Juni 2003, ISSN: 1524-9050. DOI: 10.1109/TITS.2003.821213.

[6] V. Do, L. Nghiem, N. Pham Thi und N. Pham Ngoc, "A simple camera calibration method for vehicle velocity estimation", in *2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, Juni 2015, S. 1–5. DOI: 10.1109/ECTICon.2015.7207027.

[7] J. Sochor, R. Juránek und A. Herout, "Traffic surveillance camera calibration by 3D model bounding box alignment for accurate vehicle speed measurement", *Computer Vision and Image Understanding*, Jg. 161, S. 87–98, 2017, ISSN: 1077-3142. DOI: https://doi.org/10.1016/j.cviu.2017.05.015. Adresse: http://www.sciencedirect.com/science/article/pii/S1077314217301108.

[8] J. Shi und C. Tomasi, "Good features to track", *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, S. 593–600, 1994.

[9] J.-Y. Bouguet, "Pyramidal implementation of the lucas kanade feature tracker", 2000.

Vom Monolith zu Microservices - Best Practices und Herausforderungen am Beispiel des Webkonferenzsystems Accelerator

Tobias Lampprecht
Hochschule Reutlingen
Tobias.Lampprecht@Student.Reutlingen-University.de

Abstract

Im Rahmen dieser Arbeit wird ein Transformationsprozess für den Wechsel von einer monolithischen Softwarearchitektur zu einer Microservice-Architektur erarbeitet. Hierfür werden zunächst Vorteile einer Microservice-Architektur dargestellt und mit den jeweiligen Herausforderungen während einer Transformation gegenübergestellt. Anschließend folgt eine Sammlung von Best Practices für diese Transformation. Zuletzt werden einige ausgewählte Methoden an einem konkreten Beispiel durchgeführt.

CCS Concepts

• **Software and its engineering → Software architectures**; **Software design engineering**;

Keywords

Microservice-Architektur, Software-Monolith, Transformationsprozess

1 Einleitung

Microservice-Architekturen gewinnen seit einigen Jahren immer mehr an Beliebtheit [1]. Dafür sind unter anderem agile Entwicklungsmethoden, der Zwang nach skalierbaren Anwendungen und immer kürzer werdende Releasezyklen verantwortlich. Auch bekannte Internetunternehmen wie beispielsweise Amazon[1] , Netflix[2] und Zalando[3] haben dazu beigetragen, dass sich Microservices zu einem Trend entwickelt haben. Viele dieser Unternehmen beschreiben in Fachartikeln, Büchern und Blogs, welche Vorteile sich durch den Einsatz von Microservices ergeben haben. Daher ist es naheliegend, dass auch viele kleinere Unternehmen die zahlreichen Vorteile von Microservices für sich nutzen wollen. In einer monolithischen Softwarearchitektur sind Funktionalitäten als eine gemeinsame Einheit implementiert, eine Trennung in Teilsysteme ist nicht vorgesehen. Im Gegensatz dazu wird in einer Microservice-Architketur ein Teilsystem unabhängig von anderen Funktionalitäten bereitgestellt. Microservices eignen sich unter anderem für Szenarien, in denen einzelne Komponenten unabhängig voneinander skalieren müssen oder neue Features schnell bereitgestellt werden sollen [2]. Dennoch wäre es fatal ohne Überlegungen alle Systeme auf eine Microservice-Architektur umzustellen, schließlich muss eine neue Lösung den

Betreuer Hochschule: Prof. Dr.-Ing. Peter Hertkorn
Hochschule Reutlingen
Peter.Hertkorn@Reutlingen-University.de
Wissenschaftliche Vertiefungskonferenz
27. November 2019, Hochschule Reutlingen
Copyright 2019 Tobias Lampprecht

[1] http://highscalability.com/amazon-architecture (Abgerufen am: 25.04.19)
[2] https://www.nginx.com/blog/microservices-at-netflix-architectural-best-practices/ (Abgerufen am: 21.03.19)
[3] https://www.infoq.com/news/2016/02/Monolith-Microservices-Zalando (Abgerufen am: 25.04.19)

jeweiligen Anforderungen gerecht werden. Außerdem ist der Wechsel von einer monolithischen Architektur zu einer Microservice-Architektur kein trivialer Prozess und mit einigen Herausforderungen, sowohl während der Transformation als auch beim Einsatz von Microservices, verbunden.

1.1 Zielsetzung

Das Ziel dieser Arbeit ist es, Best Practices für die Transformation von monolithischen Architekturen zu Microservice-Architekturen zu beschreiben. Dies schließt die meisten Fälle zur Einführung von Microservices ein, da üblicherweise bereits ein monolithisches System existiert, welches aktuellen Anforderungen eventuell nicht mehr gerecht wird. Hierzu werden die wesentlichen Aspekte aus der Literatur sowie aus Erfahrungsberichten von erfolgreichen Softwarearchitekten zusammengetragen.
Zudem werden Herausforderungen der Transformation den möglichen Vorteilen im jeweiligen Kontext gegenübergestellt. Eine mögliche Transformation von Accelerator[4] mit Hilfe der Best Practices wird am Ende dieser Arbeit konzipiert.

1.2 Motivation

Viele Projekte wurden früher anhand des Wasserfallmodells durchgeführt, obwohl sich in der Praxis gezeigt hat, dass dieser Ansatz selten zu einem befriedigenden Ergebnis führte [3]. Zu Beginn werden Anforderungen definiert, die Schritt für Schritt an die jeweiligen Fachabteilungen zur Architektur, Implementierung und Bereitstellung weitergegeben werden. Das lineare Durchlaufen der einzelnen Phasen kann, je nach Anforderungen, viel Zeit in Anspruch nehmen. Somit kann auch die Dauer von der Anforderungsanalyse bis zur Auslieferung der Software relativ hoch sein. Bei immer kürzer werdenden Innovationszyklen und sich ständig ändernder Anforderungen kann es dazu führen, dass anfangs definierte Anforderungen nach einer langen Entwicklungszeit nicht mehr zeitgemäß sind. Um sich am Markt durchzusetzen, ist es notwendig, flexibel und schnell auf äußere Einflüsse und technologische Innovationen reagieren zu können. Agile Unternehmensstrukturen und Entwicklungsmethoden sind notwendig, damit proaktiv Anpassungen und Veränderungen vorgenommen werden können. In der agilen Entwicklung werden kleine Teams bevorzugt, sodass fachlich unabhängige Komponenten auf die Teams aufgeteilt werden können.
Continuous Delivery [4] bietet die Möglichkeit, kleine Features schnell bereitzustellen. Hierzu ist auch eine enge Zusammenarbeit zwischen Entwicklern und dem Betrieb notwendig - idealerweise durch die DevOps-Philosophie [5]. Da jedoch ein monolithisches System für ein solches Szenario nicht die perfekte Architektur bietet [6], sollte eine dafür ausgerichtete Architektur gefunden werden. Bei einer Microservice-Architektur ist das Continuous Delivery Teil der Philosophie, weshalb sich diese Architektur in solchen Szenarien anbietet [4].

1.3 Strukturierung der Arbeit

Im aktuellen Kapitel wurde in die Thematik eingeführt und eine Zielsetzung formuliert. Kapitel 2 beinhaltet die zu bewältigenden Herausforderungen einer Transformation zu Microservices. In Kapitel 3 folgt eine Sammlung von Best Practices, die sich für die Transformation bewährt haben. In Kapitel 4 werden die gesammelten Best Practices am Webkonferenzsystem Accelerator konzeptionell umgesetzt. Abschließend enthält Kapitel 5 eine kurze Zusammenfassung der Arbeit sowie einen Ausblick.

2 Herausforderungen

Jede Veränderung am System bringt auch einige Herausforderungen mit sich. Die Vorteile von Microservice-Architekturen sind sehr

[4] https://accelerator.reutlingen-university.de

vielversprechend. Dennoch kann es durchaus sein, dass Nachteile die Vorteile überwiegen [7]. Bevor die Entscheidung für einen Wechsel zu einer Microservice-Architketur getroffen wird, sollte dies sorgfältig geprüft und geplant werden, damit sichergestellt werden kann, dass die neue technische Architektur die Bedürfnisse des Unternehmens berücksichtigt[5] und eine Verbesserung der Ausgangssituation entsteht.
Um abwägen zu können, ob ein Wechsel wirtschaftlich ist, werden zunächst ganz allgemein die sich üblicherweise ergebenden Vorteilen den Herausforderungen von Microservice-Architekturen gegenübergestellt.

Komplexität eines verteilten Systems
Während in einer monolithischen Architektur jegliche Kommunikation zwischen mehreren Komponenten systemintern abläuft, ist bei einer Microservice-Architektur eine Kommunikation zu anderen Microservices über ein Netzwerk notwendig. Dies bedeutet, dass eine Microservice-Architektur ein verteiltes System ist und auch dessen Komplexität besitzt [8]. Die Netzwerklatenz beeinflusst die Antwortzeiten von Microservices negativ und sollte daher insbesondere bei zeitkritischen Anwendungen berücksichtigt werden.
Dennoch kann die strikte Trennung durch verteilte Systeme auch Vorteile bringen. Die Komplexität eines einzelnen Services wird geringer und kann relativ einfach geändert werden [8]. Die Services sind voneinander unabhängig, wodurch eine lose Kopplung der Systemfunktionalitäten erreicht wird. Außerdem kann eine verteilte Systemarchitektur auch Auswirkungen auf die Stabilität des Systems haben. In einer monolithischen Architektur kann ein Bug in einem Teil des Systems das gesamte System lahmlegen[5]. In verteilten Systemen bleiben Störungen lokal, sodass keine anderen Services davon betroffen sind [8]. Trotzdem muss das System mit dem Ausfall einzelner Services umgehen können. Die strikte Trennung bewirkt zudem eine bessere Wiederverwendbarkeit der Services.

Erhöhter Ressourcenbedarf durch Isolation
Microservices können unabhängig voneinander ausgeliefert werden. Um dies zu realisieren, besitzt jeder Microservice seine eigene Laufzeitumgebung, welche jeweils etwas Overhead benötigt[6] . Mit steigender Anzahl an Microservices steigt auch der benötigte Overhead, um diese bereitzustellen.
Vorteile, die sich durch die Isolation ergeben, sind die getrennte Testbarkeit von Funktionalitäten der Microservices sowie, dass viele Änderungen in kurzer Zeit ermöglicht werden [2].

Größe eines Microservices
Wie der Name "Microservice" schon verrät, sollte ein Service möglichst klein sein. Trotzdem existiert keine allgemeingültige Größe für einen Microservice. Die Herausforderung besteht darin, im jeweiligen Kontext die richtige Größe zu ermitteln. Eberhard Wolff [2] gibt hierzu einige Ratschläge, die helfen, die richtige Größe für die jeweiligen Bedürfnisse zu identifizieren:

- **Teamgröße:** Ein Microservice muss von einem Team entwickelt werden können. Dies schränkt die Größe eines Microservices ein.
- **Modularisierung und Ersetzbarkeit:** Module und Microservices sollten nur so groß sein, dass ein Entwickler sie verstehen kann. Weiterentwicklung oder Ersetzen der Microservices ist dadurch möglich.
- **Infrastruktur:** Die Infrastruktur muss genügend Ressourcen für alle Microservices besitzen. Daher sollte die Größe nicht

[5] https://techbeacon.com/app-dev-testing/monolith-microservices-horror-stories-best-practices (Abgerufen am: 20.03.19)

[6] http://microservices.io/patterns/microservices.html (Abgerufen am: 22.03.19)

zu klein gewählt werden, da mit jedem Microservice Overhead verbunden ist.
- **Kommunikation:** Mit steigender Anzahl an Microservices steigt auch die Kommunikation zwischen diesen. Dies schränkt die untere Größe von Microservices weiter ein.

3 Best Practices

In den nachfolgenden Unterkapiteln folgt eine Sammlung von Best Practices für die Transformation zu einer Microservice-Architektur sowie für die Erstellung von Microservices. Die Reihenfolge der Kapitel stimmt grundsätzlich mit der Vorgehensweise des Transformationsprozesses überein.

3.1 Die aktuelle Architektur kennen

Im ersten Schritt ist es notwendig, sich mit der aktuellen monolithischen Architektur vertraut zu machen. Hierfür können insbesondere Unified Modeling Language (UML) Diagramme verwendet werden, um einen Überblick über das Gesamtsystem zu bekommen. Es ist möglich, dass die aktuelle Architektur bereits aus vielen einzelnen Komponenten besteht. Für die spätere Zerlegung der Architektur ist es hilfreich, die Domäne jeder Komponente zu verstehen. Warum das wichtig ist, wird in Unterabschnitt 3.3 deutlich. Auch kann es in der späteren Entwicklung der Microservices hilfreich sein, den aktuellen Technologie-Stack zu verstehen, um eventuell vorhandene Bibliotheken einfach migrieren zu können [9].

3.2 Abhängigkeitsinformationen herausfinden

Ein wesentlicher Schritt im Transformationsprozess ist es, das Zusammenspiel der bestehenden Komponenten zu analysieren, um herauszufinden, welche Komponenten die meisten Aufrufe anderer Komponenten besitzen. Systemkomponenten, die lose mit dem Rest des Systems gekoppelt sind, lassen sich einfacher herauslösen, als Komponenten, die eng miteinander verbunden sind.

3.3 Modellierung des Systems mittels Domain-Driven Design

Domain-Driven Design (DDD) ist ein von Eric Evans [10] vorgestellter Ansatz, bei dem e Zusammenhänge modelliert und eine klare Abgrenzung der Domänen erreicht wird. Voneinander unabhängige fachliche Zusammenhänge sollten jeweils durch eigenständige Microservices bereitgestellt werden, daher spielt die Modellierung eine wichtige Rolle. Da DDD ein sehr umfangreiches Thema ist, werden nachfolgend die wesentlichen Konzepte von DDD nach [11] knapp vermittelt. Anschließend folgt eine Einordnung zu Microservices.

Im ersten Schritt wird ein Domänenmodell erstellt, welches alle Geschäftsobjekte und deren Beziehungen abbildet. Das Domänenmodell kann anschließend in Sub-Domains aufgeteilt werden. Wesentlicher Bestandteil von DDD ist eine allgemeingültige Sprache (Ubiquitous Language), welche sowohl von den Domänenexperten als auch der eigentlichen Software verwendet werden. Jede Sub-Domain besitzt bestimmte Verantwortlichkeiten der Domäne. Eine Abgrenzung hilft dabei, dass sich die Sub-Domain auf eine bestimmte Sache konzentriert. Diese Grenze wird als Bounded Context (Kontextgrenzen) bezeichnet. In jedem Bounded Context kann eine eigene Ubiquitous Language gültig sein.

Da bei Microservices eine zentrale Herausforderung die Definition der Grenzen ist, können die DDD-Grundsätze helfen, die Größe und den richtigen Schnitt der Services zu erreichen. So kann beispielsweise ein Bounded Context ein Microservice darstellen, der für eine bestimmte Funktionalität verantwortlich ist.

3.4 Zerlegung des Monolithen

Während des Transformationsprozesses steht die Entscheidung an, wie die monolithische Architektur in Komponenten unterteilt werden kann und welche Prinzipien sich am besten eignen, um den Monolithen zu zerlegen. Dies ist keine einfache Aufgabe, dennoch existieren für diese Zerlegung in Microservices einige Möglichkeiten, die nachfolgend vorgestellt werden.

Das Entwurfsprinzip "Pattern of Change" wird von Lewis und Fowler[7] auf Microservices bezogen. Dieses besagt, dass Dinge, die sich gleichzeitig ändern, sich im selben Microservice befinden sollen. So ist es naheliegend, dass Microservices zusammengeführt werden sollten, wenn diese ständig gemeinsam geändert werden. Beim Zerlegen des Monolithen sollte also darauf geachtet werden, dass zusammenhängende Funktionalitäten gemeinsam in einen Microservice migriert werden müssen.

Balalaie et al. [9] beschreiben das rekonstruierende Entwurfsmuster "decompose by data ownership", also die Zerlegung eines Monolithen basierend auf dem Eigentum an Daten. Zusammenhängende Datensätze werden zu einer Einheit zusammengefasst. Jede Einheit wird dann zusammen mit der dazugehörigen Geschäftslogik in einen Microservice gepackt. Das Entwurfsmuster ist insbesondere dann geeignet, wenn die Domäne des Systems nicht komplex ist und Daten einfach zu Einheiten gruppiert werden können.

Das US-amerikanische E-Commerce Unternehmen Best Buy hat ihre Plattform mit Hilfe des "Strangler Pattern" [12] erfolgreich auf eine Microservice-Architektur umgestellt[8] . Dabei handelt es sich um ein Entwurfsmuster, bei dem inkrementell Teile des Altsystems durch neue Module ersetzt werden. Im Fall von Best Buy wurden Funktionalitäten extrahiert und als Microservices entwickelt. Das Altsystem wird dadurch nach und nach durch Microservices ersetzt.

Eine häufig angewendete Strategie ist es, Funktionalitäten zu extrahieren, die ziemlich lose mit dem Rest des Systems gekoppelt sind, oder Komponenten, die von unkritischen Systemfunktionen verwendet werden[9] . Hierzu gehört auch das Prinzip, dass ein Ausfall eines Microservices maximal den Ausfall einer Teilfunktion zur Folge haben darf[10] . Folglich sollten lose gekoppelte Komponenten extrahiert werden, sodass Microservices möglichst unabhängig voneinander ihre Funktionalität erfüllen können.

Alternativ zu den bisher genannten Prinzipien kann auch als erstes eine Trennung der Schichten realisiert werden, wie in [6], [13] vorgeschlagen wird. So kann insbesondere das User Interface (UI) als eigenständiger Microservice entwickelt werden. Sinnvoll ist es jedoch meist nur, wenn sowieso Änderungen an der UI bevorstehen, da die Änderungen dann mit der Entwicklung eines ersten Microservices kombiniert werden kann. Eine Trennung nach Schichten sollte jedoch nur ein Schritt zum Ziel sein. Die UI kann später weiter geschnitten und anderen Microservices zugeordnet werden.

Eine weitere Möglichkeit ist eine Zerlegung des Monolithen nach Use Cases[11] . Durch solch eine Zerlegung kann ein Microservice für eine spezielle Aufgabe zuständig sein und

[7] https://martinfowler.com/articles/microservices.html (Abgerufen am: 27.03.19)

[8] https://blog.runscope.com/posts/monolith-microservices-transforming-real-world-ecommerce-platform-using-strangler-pattern (Abgerufen am: 24.04.19)

[9] https://www.skillbyte.de/vom-monolithen-zu-microservices-eine-architektur-strategie/ (Abgerufen am: 19.03.19)

[10] https://www.innoq.com/de/articles/2017/12/microservices-herausforderungen/ (Abgerufen am: 02.04.19)

[11] http://microservices.io/patterns/microservices.html (Abgerufen am: 22.03.19)

idealerweise keine Interaktion zu anderen Microservices benötigen.
Angrenzend an dieses Prinzip ist auch eine Validierung des vorhergehenden Serviceschnittes mittels Use Cases möglich[12] . Hierbei wird zuerst mit einem geeigneten Prinzip ein grober Schnitt des Monolithen vorgenommen und anschließend durch die Kern-Use Cases geprüft, ob diese möglichst innerhalb eines Serviceschnittes abgearbeitet werden können.

3.5 Erstellung von Microservices

Im vorherigen Abschnitt wurden einige Prinzipien und Entwurfsmuster zur Zerlegung eines Monolithen in Microservices vorgestellt. Im Gegensatz dazu fokussiert sich dieser Abschnitt auf allgemeine Prinzipien zur Erstellung von Microservices.

Ein Ansatz der von Lewis und Fowler[13] beschrieben wird und von der britischen Tageszeitung The Guardian[14] angewendet wird ist es, neue Funktionen als Microservices zu entwickeln, welche mit dem Monolith kommunizieren. Dieser Ansatz ist besonders für temporär notwendige Funktionen praktisch, die nach Beendigung des Gebrauchs einfach abgeschaltet werden können.

Das Netflix-Entwicklerteam gibt den Ratschlag für jeden Microservice einen eigenen Datenspeicher zu verwenden[15]. Richardson [14] benennt dies als das "Database per service"-Muster. Dieses Muster sichert das Prinzip der losen Kopplung. Eine Datenbank kann ohne Effekte auf weitere Services verändert werden. Notwendige Daten für andere Microservices können dann beispielsweise über eine API bereitgestellt werden. Für Betriebsvorgänge, die mehrere Services durchlaufen müssen, kann das "Saga pattern" [15] eine zufriedenstellende Lösung sein. Dieses Muster ermöglicht verteilte Transaktionen zwischen den Microservices durch eine Sequenz von lokalen Transaktionen.

3.6 Kommunikation zwischen Monolith und Microservices

Typischerweise existiert während des Wechselprozesses ein hybrides System, indem der Monolith mit den bereits umgesetzten Microservices kommuniziert. Es werden schrittweise Microservices aus dem Monolithen extrahiert, bis dieser schließlich abgelöst werden kann. Ein API-Gateway [16] kann diese Aufgabe übernehmen. Es kann auch dazu verwendet werden, Benutzern mit unterschiedlichen Endgeräten, den passenden Service bereitzustellen.

3.7 Containerisierung und Clustering

Mit einer steigenden Anzahl an Microservices eines Systems, kann es viel Zeit in Anspruch nehmen, alle Microservices gleichzeitig in Betrieb zu nehmen. Ein wesentlicher Aspekt bei der Transformation zu Microservices ist daher eine automatisierte Bereitstellung der Microservices zu etablieren. Ein System, das dafür meist zum Einsatz kommt ist Kubernetes[16].

4 Beispiele am Webkonferenzsystem Accelerator

In diesem Kapitel wird eine mögliche Transformation von Accelerator zu einer Microservice-Architektur anhand des Transformationsprozesses aus Abschnitt 3 konzipiert. Im Rahmen dieser Arbeit wird jedoch auf die Modellierung des Systems mittels DDD verzichtet, da zum einen der Umfang von DDD

[12] https://www.innoq.com/de/articles/2017/12/microservices-herausforderungen/ (Abgerufen am: 02.04.19)
[13] https://martinfowler.com/articles/microservices.html (Abgerufen am: 27.03.19)
[14] https://www.theguardian.com/
[15] https://www.nginx.com/blog/microservices-at-netflix-architectural-best-practices/ (Abgerufen am: 21.03.19)
[16] https://kubernetes.io

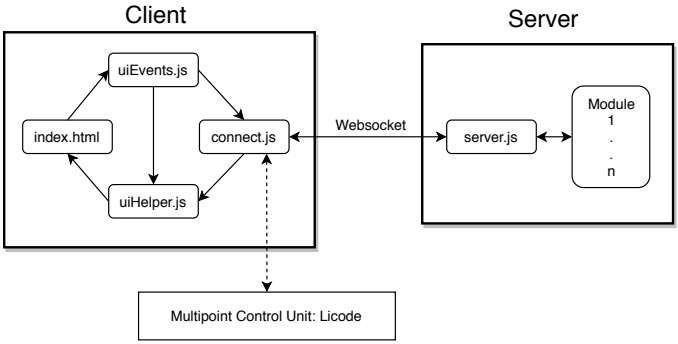


Abbildung 1: Systemarchitektur Accelerator

den Rahmen der Arbeit überschreiten könnte und zum anderen die Größe und Komplexität des vorliegenden Systems übersichtlich ist. Aus diesen Gründen werden in Unterabschnitt 4.2 die Prozessschritte *Abhängigkeitsinformationen herausfinden* und *Modellierung des Systems* zusammengefasst und eine Modellierung nach Komponentenklassen vorgenommen.

4.1 Die aktuelle Architektur kennen

Im ersten Schritt der Transformation wird ein Überblick über die Ausgangssituation verschafft. Die aktuelle Architektur von Accelerator besteht aus einer Node.js-Anwendung, einer Multipoint Control Unit (MCU) und einer Präsentationsschicht, bestehend aus einer HTML-Datei (Hypertext Markup Language) und einigen JavaScript-Dateien. In Abbildung 1 ist diese Architektur dargestellt.
Die Node.js-Anwendung besteht insbesondere aus dem Webframework Express.js[17] und verwendet für die Echtzeit-Kommunikation zwischen Benutzern die Bibliothek Socket.io[18]. Als MCU verwendet Accelerator Licode[19], welches auf der WebRTC-Technologie basiert. Das User Interface besteht im Wesentlichen aus einer index.html, die dynamisch den passenden Inhalt bereitstellt. Mehrere JavaScript-Dateien sind für die Verbindung zum Backend sowie für verschiedene Events zuständig.

4.2 Abhängigkeitsinformationen herausfinden und Modellierung des Systems

Alle Funktionalitäten von Accelerator können in die folgenden Komponentenklassen eingeteilt werden:

- Kommunikationskomponenten
  - Audio
  - Video
  - Chat
- Kollaborationskomponenten
  - Whiteboard
  - Etherpad
  - Filesharing
  - Benutzerelemente
- Moderationskomponenten
  - Präsentation
  - Screenshare
  - Zeiger
  - Medienwiedergabe
- Datenverwaltungskomponenten
  - Räume
  - Benutzer
- Eigene Komponenten
  - Notizen
  - Chatbot

[17] https://expressjs.com
[18] https://socket.io
[19] http://lynckia.com/licode/

Desweiteren kann eine Einteilung nach Funktionalitäten erfolgen, die innerhalb und außerhalb eines Raumes zur Verfügung stehen. Die Datenverwaltungskomponenten, zu denen die Raum- und Benutzerverwaltung gehört, ist die einzige Komponentenklasse, die außerhalb eines Raumes besteht. Die Abhängigkeiten zu den anderen Komponentenklassen sind daher gering, sodass sich diese Funktionalitäten als ersten Microservice gut eignen. Die weiteren Prozessschritte werden daher anhand des Microservices *Datenverwaltungskomponenten* durchgeführt. Ein weiterer Microservice könnte anhand den Komponenten *Notizen* und *Chatbot* erstellt werden. Dies sind Komponenten die keinerlei Synchronisation zwischen mehreren Benutzern benötigen, sondern lediglich jeweils einem Benutzer zur Verfügung stehen.

4.3 Zerlegung des Monolithen

Für die Datenverwaltungskomponenten kann das Entwurfsmuster "decompose by data ownership" angewendet werden, da eine Trennung nach dem Eigentum an Daten erfolgt. Die Geschäftslogik, die aus dem Monolith heraus gelöst werden muss, sind die Funktionalitäten:

- Raum erstellen,
- Raum löschen,
- alle Räume anzeigen,
- Raum betreten und
- Benutzername festlegen.

Das User Interface dieser Funktionalitäten ist bereits als zwei eigenständige Komponenten im Monolith implementiert und kann dadurch einfach herausgeschnitten werden.

4.4 Erstellen von Microservices

Die Grundlage für den Microservice *Datenverwaltungskomponenten* (siehe Abbildung 2) ist eine Node.js-Anwnedung mit Express und

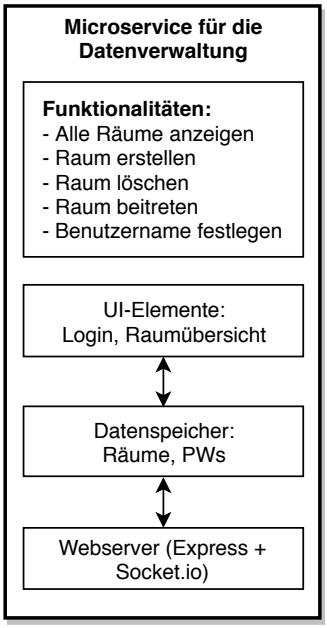


Abbildung 2: Microservice für die Datenverwaltung

Socket.io, sodass die notwendigen Funktionalitäten aus dem Monolith übernommen werden können. Im Datenspeicher sind Informationen über alle verfügbaren Räume gespeichert sowie ein optionales Passwort zu jedem Raum. Der Benutzername wurde im Monolith Clientseitig gespeichert, das auch für den Microservice übernommen werden kann. Der Microservice besitzt die UI-Elemente *Login* und *Raumübersicht*. Die Funktionalitäten sind das Erstellen, Löschen und Beitreten eines Raumes, alle Räume Anzeigen und einen Benutzernamen festlegen. Das Entwurfsmuster "Database per Service" wird von diesem Microservice angewendet, da dieser seinen eigenen Datenspeicher besitzt. In Zukunft kann es sinnvoll sein, den Benutzername ebenfalls innerhalb des Microservices zu speichern.

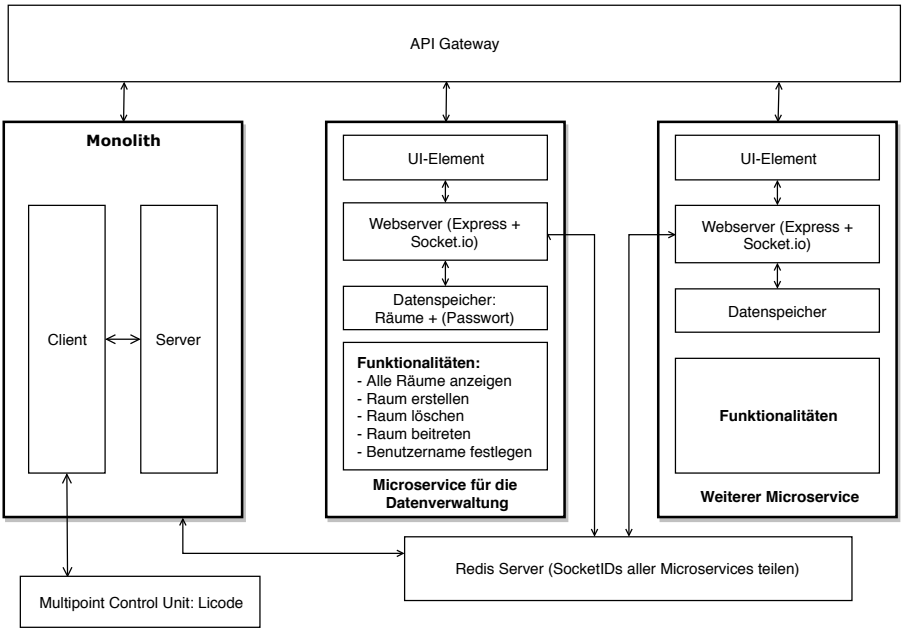


Abbildung 3: Mögliches Zusammenspiel zwischen Monolith und Microservices

4.5 Kommunikation zwischen Monolith und Microservice

Für eine schrittweise Transformation von Accelerator zu einer Microservice-Architektur bietet sich das *Strangler-Pattern* an. Es kann zunächst ein Microservice für die Datenverwaltungskomponenten erstellt werden, der mit dem Monolith kommuniziert. Ein API-Gateway (Application Programming Interface) kann hierbei als Schnittstelle für Daten oder Funktionalitäten dienen, die übergreifend benötigt werden. Außerdem kann es notwendig sein, einen Benutzer über alle Services hinweg anhand seiner Session (Socket ID) zu identifizieren. Dazu kann ein Redis-Server eingesetzt werden, der alle Sessions verwaltet. In Abbildung 3 ist ein mögliches Zusammenspiel der einzelnen Services und des Monolithen dargestellt.

5 Zusammenfassung und Ausblick

In dieser Arbeit wurden Vorteile von Microservice-Architekturen den sich möglicherweise ergebenden Herausforderungen bei der Transformation von einer monolithischen Architektur zu einer Microservice-Architektur gegenübergestellt. Die Vorteile sind unter anderem, eine schnelle Auslieferung von neuen Funktionalitäten, eine getrennte Testbarkeit und die Skalierbarkeit von einzelnen Services. Dem gegenüber stehen jedoch die Herausforderungen wie bspw. die Komplexität eines verteilten Systems oder der perfekte Serviceschnitt mit einer passenden Größe des Microservices.

Es wurden Best Practices für die Transformation zu einer Microservice-Architektur sowie

für die Erstellung von Microservices zusammengetragen, mit dessen Hilfe eine Transformation gelingen kann. Zuletzt wurden anhand des Transformationsprozesses die Datenverwaltungskomponenten von Accelerator als erster möglicher Microservice konzeptionell erarbeitet.
In Zukunft können weitere Microservices mit Hilfe des Transformationsprozesses realisiert, bereitgestellt und betrieben werden, sodass das monolithische System schrittweise abgelöst werden kann.

Literatur

[1] M. Fowler und J. Lewis, "Microservices: Nur Ein Weiteres Konzept in Der Softwarearchitektur Oder Mehr", *Objektspektrum*, S. 14–20, Jan. 2015.

[2] E. Wolff, *Microservices - Grundlagen flexibler Softwarearchitekturen*, 2. Heidelberg: dpunkt Verlag, 2018, ISBN: 978-3-960-88414-9.

[3] O. Vogel, Hrsg., *Software-Architektur: Grundlagen - Konzepte - Praxis*, de, 2. Aufl. Heidelberg: Spektrum Akad. Verl, 2009, OCLC: 255278881, ISBN: 978-3-8274-1933-0.

[4] E. Wolff, *Continuous Delivery - Der pragmatische Einstieg*, 2. Heidelberg: dpunkt Verlag, 2016, ISBN: 978-3-864-91931-2.

[5] D. Harrison und K. Lively, *Achieving DevOps - A Novel About Delivering the Best of Agile, DevOps, and Microservices*, 1st ed. New York: Apress, 2019, ISBN: 978-1-484-24388-6.

[6] E. Wolff, *Das Microservices-Praxisbuch: Grundlagen, Konzepte und Rezepte (German Edition)*, 1. Heidelberg: dpunkt Verlag, 2018, ISBN: 978-3-86490-526-1.

[7] J. Soldani, D. A. Tamburri und W.-J. Van Den Heuvel, "The Pains and Gains of Microservices: A Systematic Grey Literature Review", en, *Journal of Systems and Software*, Jg. 146, S. 215–232, Dez. 2018, ISSN: 01641212. DOI: 10.1016/j.jss.2018.09.082.

[8] D. Takai, *Architektur für Websysteme*. Carl Hanser Verlag GmbH & Co. KG, Sep. 2017, ISBN: 978-3-446-45056-1. DOI: 10.3139/9783446452480.

[9] A. Balalaie, A. Heydarnoori, P. Jamshidi, D. A. Tamburri und T. Lynn, "Microservices migration patterns: Microservices migration patterns", en, *Software: Practice and Experience*, Juli 2018, ISSN: 00380644. DOI: 10.1002/spe.2608.

[10] E. Evans, *Domain-Driven Design: Tackling Complexity in the Heart of Software*. Addison-Wesley Professional, 2003, ISBN: 0321125215.

[11] V. Vernon, *Domain-Driven Design kompakt*. Heidelberg: dpunkt.verlag, 2017, ISBN: 978-3-86490-439-4.

[12] N. Rajagopalan und K. Gowda, "Journey from Monolith Application to Microservices", en, *IEEE India Info. Vol. 14 No. 2*, Jg. 14, Nr. 2, S. 5, 2019.

[13] S. Sunil und Sarita, "Transform Monolith into Microservices using Docker", in *2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA)*, Aug. 2017, S. 1–5. DOI: 10.1109/ICCUBEA.2017.8463820.

[14] C. Richardson, *Microservices Patterns: With examples in Java*. Manning Publications, 2018, ISBN: 1617294543.

[15] M. Štefanko, O. Chaloupka und B. Rossi, "The Saga Pattern in a Reactive Microservices Environment:" en, in *Proceedings of the 14th International Conference on Software Technologies*, Prague, Czech Republic: SCITEPRESS - Science and Technology Publications, 2019, S. 483–490, ISBN: 978-989-758-379-7. DOI: 10.5220/0007918704830490.

[16] F. Montesi und J. Weber, *Circuit Breakers, Discovery, and API Gateways in Microservices*, 2016. eprint: arXiv:1609.05830.

Analyse einer Dokumentenverarbeitung mit Microsoft Office im Krankenhausinformationssystem

Maximilian Malik
Hochschule Reutlingen
Maximilian.Malik@Student.Reutlingen-University.DE

Abstract

Die Modernisierung langjährig existierender Teillösungen in Krankenhausinformationssystemen (KISs) wird oft vernachlässigt. Ehemals innovative Lösungen können sich so im Laufe der Zeit zu Problemen und Hindernissen entwickeln. Eine solche Teillösung ist die Verwaltung von klinischen Dokumenten. Mit zunehmender Vernetzung und zunehmendem Austausch zwischen medizinischen Institutionen werden hierfür moderne Lösungen benötigt. Die Modernisierung bestehender Lösungen kann dabei oft nur unter Berücksichtigung der vorhanden Spezifikationen vorgenommen werden. Im Rahmen dieser Arbeit wird eine Systemanalyse für ein bestehendes Dokumentenverarbeitungssystem eines KIS durchgeführt. Durch Anwenden eines Reverse Engineering Prozesses mit dem Ziel der nachträglichen Dokumentation werden die Spezifikationen der bestehenden Lösung erhoben. Basierend auf den erhobenen Spezifikationen und Erkenntnissen werden dem KIS-Hersteller Handlungsempfehlungen für das weitere Vorgehen bei der Modernisierung der Dokumentenverarbeitung gegeben.

CCS Concepts

• **Information systems → Information storage systems**; • **Software and its engineering** → *Software reverse engineering*; *Documentation*;

Keywords

Krankenhausinformationssystem, Dokumentenmanagementsystem, Microsoft Office, Component Object Model, Reverse Engineering

1 Einleitung

Dieser Artikel ist in Zusammenarbeit mit einem Hersteller für Krankenhausinformationsysteme (KISs) entstanden und behandelt sensible Inhalte. Aus Datenschutzgründen sind der KIS-Hersteller sowie das betreffende KIS in dieser Arbeit anonymisiert und werden im Folgenden als KIS-Hersteller bzw. Referenz-KIS bezeichnet.

Das Referenz-KIS bietet softwareseitig umfangreiche und vielfältige Unterstützung für das Patienten- und Fallmanagement für Schweizer Hospitäler. Ein Aspekt dieser Unterstützung ist das benutzerorientierte Verwalten von patienten- oder fallbezogenen klinischen Dokumenten. Die softwareunterstützte Dokumentenverwaltung beinhaltet die Verarbeitung von Dokumenten durch die Anwendungen *Microsoft Word* (Word) und *Microsoft Excel* (Excel). Über direkt von Microsoft bereitgestellte programmierbare Schnittstellen (APIs) können diese systemseitig gesteuert werden. Mittels dieser APIs werden im Referenz-KIS klinische Dokumente automatisch verarbeitet oder dem Benutzer zur Ansicht und Bearbeitung bereitgestellt. Die Lösungen zur Dokumentenverwaltung und

Betreuer Hochschule: Prof. Dr. Rer. Medic. Christian Thies
Hochschule Reutlingen
Christian.Thies@Reutlingen-University.DE

Wissenschaftliche Vertiefungskonferenz
27. November 2019, Hochschule Reutlingen
Copyright 2019 Maximilian Malik

-verarbeitung ist seit vielen Jahren in das Referenz-KIS integriert und wird bis heute aktiv verwendet. Durch den langjährigen Einsatz ist die vorhandene Lösung zwar bewährt, entspricht heute jedoch nicht mehr den eigenen Anforderungen des KIS-Herstellers.

1.1 Problemstellung und Ziele

Die verwendeten APIs zur Steuerung von Word und Excel basieren auf der mittlerweile obsoleten Component Object Model (COM) Technologie. COM ist seit vielen Jahren durch das .NET Framework abgelöst, ist mit diesem jedoch uneingeschränkt kompatibel[1]. Mit dem generellen Technologiefortschritt haben sich im Laufe der Zeit Nachteile der COM-basierten Lösung zur Dokumentenverarbeitung ausgeprägt und entwickelt. Im Zuge der kontinuierlichen Modernisierung des Referenz-KIS ist es deshalb wünschenswert, diese Lösung unter Berücksichtigung aller damit verbundenen Aspekte zu ersetzen. Allerdings sind keine Dokumentationen oder Spezifikationen zur bestehenden Lösung vorhanden, wodurch die Ersetzung dieser nicht ohne Weiteres möglich ist.

Ziel dieser Arbeit ist, das Referenz-KIS hinsichtlich der bestehenden, COM-basierten Lösung zur Verarbeitung klinischer Dokumente zu analysieren. Die funktionellen Spezifikationen und Probleme dieser Lösung werden dabei erhoben und dokumentiert. Die zu erhebenden Spezifikationen sollen in einer künftigen Arbeit als Grundlage für die Ersetzung der bestehenden Lösung dienen. Hierfür wird im Anschluss an die Analyse eine Handlungsempfehlung für das weitere Vorgehen formuliert werden.

[1] https://docs.microsoft.com/de-de/office/client-developer/outlook/pia/introduction-to-interoperability-between-com-and-net (Zuletzt geprüft am 22.08.2019)

1.2 Aufbau der Arbeit

In Kapitel 2 werden zunächst die COM-Technologie sowie die von Microsoft bereitgestellten APIs zur Steuerung von Word und Excel näher beschrieben. Anschließend werden in Kapitel 3 ein Prozess zur Redokumentation für die Erhebung der funktionellen Spezifikationen der COM-basierten Lösung sowie dessen Ergebnisse beschrieben. In Kapitel 4 werden diese Ergebnisse sowie weiterführende Arbeiten diskutiert. Abschließend wird in einem Fazit der Arbeit eine Handlungsempfehlung für den KIS-Hersteller für das weitere Vorgehen formuliert.

2 Grundlagen

Im Folgenden werden die COM-Technologie und die darauf aufbauenden Microsoft interop Bibliotheken beschrieben.

2.1 Component Object Model

COM ist ein von Microsoft entwickeltes System, um Interoperabilität auf binärer Ebene zwischen objektorientierter Software verteilt und plattformunabhängig zu ermöglichen [1]. COM wurde erstmals 1993 veröffentlicht [2]. Die Softwareinteroperabilität ist in COM als Client-Server Modell implementiert und basiert auf Zeigern und Interfaces. In COM beschreiben Interfaces eine Gruppe von Funktionen unabhängig ihrer Implementierung [1]. Diese Funktionen stellen Dienste für Clients dar [1]. Um diese Dienste nutzen zu können, erhalten Clients sogenannte Schnittstellenzeiger (engl. interface pointer), die auf eine Instanz des entsprechenden Interfaces verweisen. Über Schnittstellenzeiger können alle Dienste verwenden werden, die das Interface definiert.

COM Interfaces werden von Komponenten implementiert. Eine COM Komponente ist ein Stück Software, das ein oder mehrere COM Interfaces implementiert. Die Interfaces und ihre Implementierungen sind dabei bewusst getrennt. Auf die Daten der COM Komponenten kann deshalb nur über die davon implementierten Interfaces zugegriffen werden [1].

Für die Verwendung eines Dienstes verwaltet das COM System den zeigerbasierten Zugriff auf die entsprechende Komponente.
Aufgrund des zeigerbasierten Ansatzes muss COM-basierte Software mit einer Programmiersprache entwickelt werden, die Zeigerstrukturen erzeugen kann und entweder implizit oder explizit Funktionen über Zeiger aufrufen kann [1]. COM stellt keine weiteren Anforderungen, definiert allerdings Spezifikationen an die Programmierung und an die Modellierung von COM Objekten, um Interkationen zwischen unabhängigen Softwarekomponenten auf binärer Ebene zu ermöglichen. Das macht COM zu einem Standard für binäre Interoperabilität [1]. Binär deshalb, weil der COM Standard erst nach der Übersetzung des Codes in binären Maschinencode (Kompilierung) greift. Die Entwicklung und Verwendung von COM Software ist daher unabhängig von Struktur und Implementierungsdetails und weitgehend unabhängig von Sprache [1].

2.2 Die COM Architektur

Abbildung 1 zeigt eine Übersicht der COM Architektur. Die einzelnen Komponenten und Informationsflüsse des Diagramms werden zum Verständnis der Abbildung im Folgenden erläutert.
COM Interfaces definieren und beschreiben Funktionen, die für Clients zur Nutzung bereitgestellt werden. Zur Funktionsbeschreibung gehören Kontrakt (Inputs, Outputs und Typen), erwartetes Verhalten und Verantwortlichkeiten (Funktionalitäten) ihrer Implementierung [1]. Das COM System stellt ein Interface (*IUnkown*) mit grundlegenden Funktionen für Polymorphie und zur Verwaltung der Lebensdauer einer Interfaceinstanz bereit [1]. Alle COM Interfaces müssen direkt oder indirekt von IUnkown abgeleitet werden [1][2]. Zusätzlich muss für jedes COM Interface ein *Globally Unique Identifier* (GUID), auch *Universally Unique IDentifier* (UUID) genannt [3], erzeugt werden. Diese werden als *Unique Interface ID* (IID) bezeichnet [1]. COM Interfaces werden zur Kapselung der Implementierung von der Nutzung von Softwarekomponenten verwendet. Tatsächliches Objektverhalten wird so verborgen und erlaubt Details einer Implementierung zu verändern oder auszutauschen, ohne dass Clientanwendungen neu kompiliert werden müssen.
COM Interfaces können nicht direkt instanziiert oder genutzt werden. Sie müssen als **COM Komponenten** implementiert werden [1][2], die wiederum instanziiert und verwendet werden können. COM Komponenten implementieren mindestens ein COM Interface [1]. Für Komponenten muss ebenfalls eine GUID erzeugt werden. Diese wird als *Class ID* (CLSID) bezeichnet. CLSIDs werden zur eindeutigen Identifikation, zur Lokalisierung installierter Komponenten und zur Anforderung bestimmter Implementierungen von Diensten verwendet [1]. Wird ein Dienst mittels CLSID angefordert, wird die entsprechende Komponente von einem COM Server bereitgestellt.
COM Server sind Binärdateien, die COM Komponenten beinhalten. Sie können im selben Prozess (in-process) oder außerhalb des Prozesses (out-of-process) verwendet werden, in dem ein Dienst aufgerufen wird. Zur in-process Nutzung werden COM Server als *dynamic linked library* (DLL) bereitgestellt [1]. Für die out-of-process Nutzung werden COM Server als *executable file* (EXE) bereitgestellt [1]. In beiden Fällen wird der Aufbau einer direkten Verbindung zwischen Clients und Instanzen von Komponeten vom COM System geregelt. Die Art der Bereitstellung der COM Server ist für Clients daher irrelevant [1].
COM Clients sind jegliche Art von Code, die Dienste einer COM Komponente über ein Interface nutzen. Über die CLSID werden von der COM Library eine Instanz der dazugehörigen Komponente angefordert [1]. Die COM

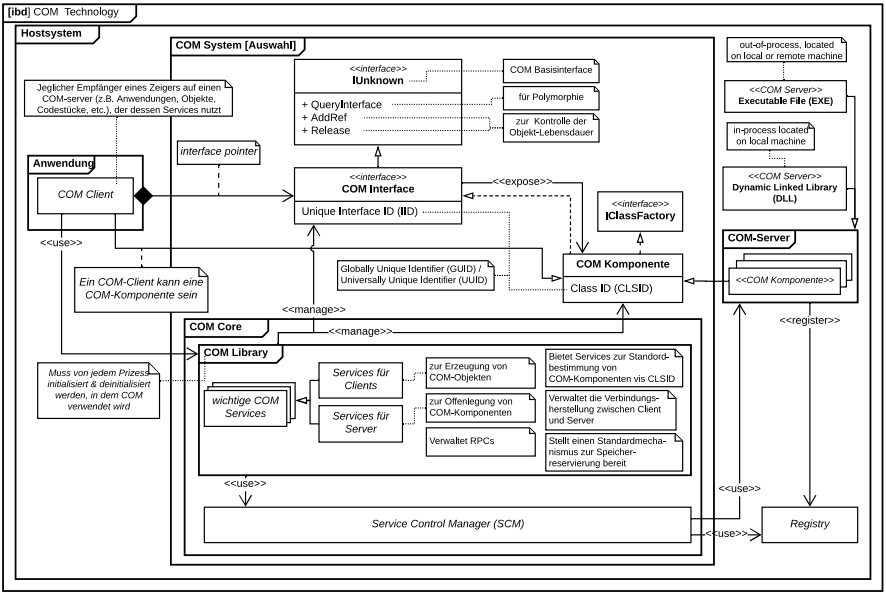


Abbildung 1: Übersicht der COM Architektur

Library liefert daraufhin einen Schnittstellenzeiger, der auf die Instanz (COM Objekt) verweist. Clients sind dann für die Verwaltung der Lebensdauer der COM Objekte verantwortlich. Solang das COM Objekt erhalten wird, können dessen Dienste über ein implementiertes Interface verwendet werden.

Die **COM Library** beinhaltet wichtige Dienste, um die Entwicklung von COM-basierter Software zu unterstützen [1]. Dazu gehören Dienste zur Offenlegung der Komponenten eines COM Servers für die Nutzung, Dienste zur Erzeugung von COM Objekten, Dienste zur Lokalisierung von Komponenten, Dienste für die transparente Nutzung von *remote procedure calls* (RPC) [4] und ein Standardmechanismus zur Reservierung von Speicher, der zwischen Objekten, die in unterschiedlichen Prozessen existieren, ausgetauscht werden soll [1].

Um den Standort von Komponenten bzw. ihrer Server zu bestimmen, wird der **Service Control Manager** (SCM) mit einer CLSID befragt [1]. Der SCM übernimmt die tatsächliche Lokalisierung des COM Servers. Dabei wird die Registry des Systems nach der angefragten CLSID durchsucht. Ist der zur CLSID korrespondierende COM Server gefunden, wird eine Instanz der entsprechenden COM Komponente darin angefordert [1]. Wird das Objekt erfolgreich erzeugt, wird ein darauf verweisender Zeiger zum Client zurückgeliefert. Dort wird das referenzierte Objekt durch ein Interface maskiert. Solche Zeiger werden deshalb in COM als Schnittstellenzeiger bezeichnet. Über Schnittstellenzeiger kommunizieren Clients direkt mit dem dahinterstehenden Objekt [1]. Dabei können nur die Dienste verwendet werden, die in dem Interface definiert sind, das zur Maskierung verwendet wird [1]. Sollen Dienste eines anderen Interfaces verwendet werden, dass in derselben Komponente implementiert ist, muss ein neues COM Objekt angefordert werden.

In der **Registry** werden die Standorte der auf dem System installierten COM Server gespeichert. Die Registry ist eine Datenbank für die Konfiguration von Hardware, Software und Benutzern eines Systems. Wird COM-basierte Software installiert, muss diese auf dem System registriert werden. Dazu werden Einträge in der Registry hinterlegt, die den Standort, unterstützte CLSIDs und weitere Informationen enthalten [1].

2.3 Microsoft Office Interop

Der KIS-Hersteller verwendet für die Steuerung der Anwendungen und Dateien für Word und Excel die von Microsoft bereitgestellten primären interop-Assemblies (PIAs):

- Microsoft.Office.Interop.Word.dll
- Microsoft.Office.Interop.Excel.dll

Assemblies sind auslieferbare Softwarekomponenten des .NET Frameworks, in denen Versionskontrolle, Wiederverwendung, Aktivierungsumfang und Sicherheitsberechtigungen vereint sind. PIAs sind eine spezielle Art von Assemblies, die als Proxies, sogenannte *Runtime Callable Wrappers* (RCWs), für COM Komponenten dienen. RCWs stellen die Kompatibilität von .NET zu COM Komponenten her und waren bis zur .NET Version 4.0 in Form von PIAs notwendig, um COM Komponenten in .NET Anwendungen verwenden zu können.
PIAs bieten alle APIs der darunterliegenden COM Komponenten für die Softwareentwicklung mit dem .NET Framework an. Die APIs der oben genannten PIAs, im Folgenden Office-PIAs genannt, ermöglichen die Interaktion mit der jeweiligen Anwendung und Dokumenten zu Word und Excel. Im Referenz-KIS werden nur wenige APIs der Office-PIAs benötigt. Aufgrund dessen wurden eigene Wrapper-Klassen für die PIAs implementiert, die innerhalb des Referenz-KIS verwendet werden. Die tatsächliche Steuerung von Word und Excel basiert jedoch auf COM-basierter Softwarekomponenten.

3 Systemanalyse

Für die COM-basierte Lösung zur Verarbeitung klinischer Dokumente sind keine Spezifikationen oder Dokumentationen vorhanden. Zur Bestimmung des Funktionsumfangs der Lösung wird deshalb ein Prozess zur Redokumentation angewandt. Redokumentation ist eine Form von Reverse Engineering zur nachträglichen Dokumentation eines Systems [5]. Verschiedene Prozesse zur Redokumentation ganzer Softwaresysteme sind in [5], [6] und [7] beschrieben. Die beiden letzteren Arbeiten zielen spezifisch auf die Rekonstruktion ganzer Softwaresystemarchitekturen ab, wohingegen der Ansatz in [5] allgemeiner beschrieben ist. Dieser wurde deshalb ausgewählt und zur Durchführung für die vorliegende Problemstellung angepasst.

3.1 Redokumentation-Prozess

Zur Vorbereitung der Redokumentation wird der aktuelle Stand der bestehenden Lösung zur Dokumentenverarbeitung erfasst. Dabei gilt es, die dafür relevanten Komponenten des Referenz-KIS sowie die dazu verfügbaren Wissensträger zu ermitteln und bezüglich ihres Vertrauensniveaus (VNs) zu bewerten. Das VN dient hier als Maß für die Integrität der Informationen oder des Wissens, die aus einer Komponente bzw. aus einem Wissensträger extrahiert werden können. Basierend auf den Ergebnissen der Vorbereitung wird ein Plan für das weitere Vorgehen der Redokumentation erarbeitet und anschließend durchgeführt.

3.1.1 Systemübersicht Im Referenz-KIS sind zur Kapselung der Office-PIAs und der Dokumentenverarbeitung zusätzliche Klassen implementiert. Diese Klassen lassen sich wie folgt unterteilen:

a) Klassen, die als Wrapper der Office-PIAs dienen. Diese Klassen werden durch den Namensraum *N2* repräsentiert und sind Bestandteil einer Bibliothek im Referenz-KIS.

b) Klassen, die den Namensraum *N2* verwenden, um alle innerhalb des Referenz-KIS notwendigen Funktionalitäten zur Verarbeitung von klinischen Dokumenten mittels Word und Excel zu implementieren. Diese Klassen werden im Folgenden als funktionelle Office-Klassen (FOCs) bezeichnet. FOCs werden durch den Namensraum *N1* repräsentiert.

Zur Erfassung der relevanten Systemkomponenten werden in den Quellcode-Dateien des gesamten Referenz-KIS Verweise auf die Namensräume *N1* und *N2* gesucht. Eine Systemkomponente gilt dann als relevant, wenn mindestens eine Referenz zu einem dieser Namensräume vorhanden ist. Zur Ermittlung aller potentiellen Referenzen zu den Namensräumen *N1* und *N2* wird die Anwendung *Total Commander*[2], ein Dateimanager für Windows, verwendet. Potentiell deshalb, weil Namensräume nicht eindeutig sind und die Referenzen daher evaluiert werden müssen. Durch die bereits vorhandene vollständige Trennung des Codes der einzelnen Komponenten des Referenz-KIS, können alle Ergebnisse eindeutig einer Systemkomponente zugewiesen werden. Für alle Systemkomponenten, die potentiell relevanten Quellcode aufweisen, werden die Suchergebnisse manuell inspiziert. Mögliche Referenzen zu einem der genannten Namensräume können so bestätigt oder wiederlegt werden. Die daraus resultierende Systemzusammenstellung wird in Abbildung 2 als internes Blockdiagramm (ibd) dargestellt. Dort sind die insgesamt 15 als relevant identifizierten Systemkomponenten, sowie deren Abhängigkeiten zur bestehenden Lösung für die Dokumentenverarbeitung sowie die Abhängigkeit dieser Lösung zur COM Technologie dargestellt.

3.1.2 Wissensträger und Vertrauensniveau

Da keine Dokumentationen für die vorhandene Lösung zur Verarbeitung klinischer Dokumente existieren, entfallen diese als mögliche Wissensträger. Allerdings arbeiten noch immer ehemalige Entwickler der Lösung beim KIS-Hersteller. Diese stellen die Hauptwissensträger dar. Weitere Wissensträger stellen die Implementierung der Lösung (Quellcode) sowie das Referenz-KIS als Anwendung dar. Tabelle 1 zeigt die identifizierten Wissensträger sowie deren VN. Das VN beschreibt, wie relevant oder geeignet die aus einem Wissensträger gewonnene Information für die Dokumentation ist. Das VN wird mit einem Punktesystem von 0 bis 5 bewertet, wobei 0 Punkte als nicht für Dokumentation geeignet (irrelevant) und 5 Punkte als maßgeblich (relevant) stehen. Zur Bewertung der VNs werden die betreffenden Personen bezüglich ihrer Selbsteinschätzung und der Einschätzung der anderen Wissensträger hinsichtlich der Vertrautheit mit der Lösung interviewed. Die Wertung des VNs einer Person ergibt sich aus dem Durchschnitt aller Schätzungen. Das VN der Implementierung und des Referenz-KIS werden durch eine Inspektion anhand ihrer Transparenz zur bestehenden Lösung bewertet.

3.2 Redokumentation-Planung

In Interviews mit den Personen-Wissensträgern werden bekannte Spezifikationen und Funktionen zur bestehenden Lösung gesammelt. Unabhängig davon werden die relevanten Komponenten (s. Abb. 2) des Referenz-KIS bezüglich der für die Benutzer angebotenen Dienste zur Verarbeitung klinischer Dokumente inspiziert. Außerdem werden durch eine Analyse des Quellcodes der Lösung deren Systemfunktionen erarbeitet.

Die Erhebung und Evaluation der Spezifikationen findet in drei Schritten statt: Im ersten Schritt werden die während der Inspektion des Referenz-KIS zu identifizierenden Dienste bezüglich ihrer Relevanz zur Lösung evaluiert. Hierfür werden die Ergebnisse dieser Inspektion mit denen der Interviews abgeglichen. Im zweiten Schritt werden die mittels

[2] https://www.ghisler.com (Zuletzt geprüft am 23.08.2019)

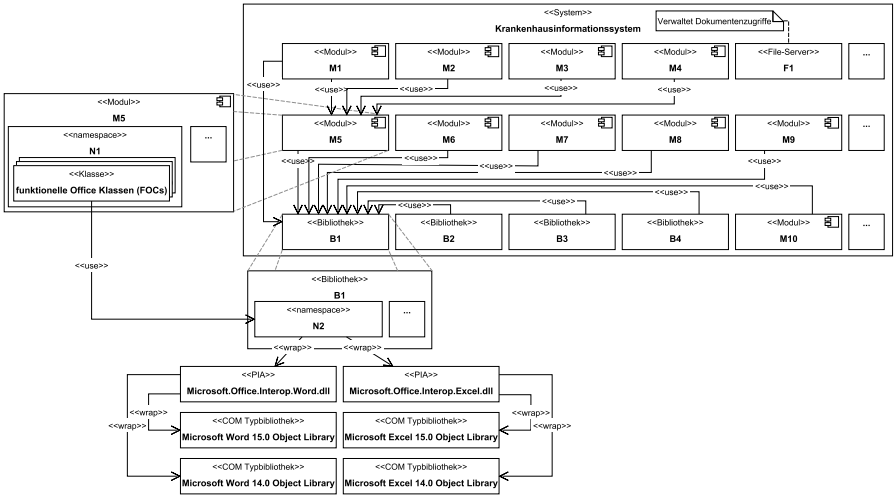


Abbildung 2: Übersicht der Systemkomponenten des Referenz-KIS mit Abhängigkeiten zur vorhandenen Lösung zur Verarbeitung klinischer Dokumente

Tabelle 1: Wissensträger zur vorhandenen Lösung zur Verarbeitung klinischer Dokumente

#	Wissenträger	Kategorie	Beschreibung	Vertrauensniveau
1	Entwickler 1	Person	Mitentwickler der COM-basierten Dokumentenverwaltung.	4/5
2	Entwickler 2	Person	Entwicklungsleiter und Mitentwickler der COM-basierten Dokumentenverarbeitung.	5/5
3	Entwickler 3	Person	Bugfixing für die COM-basierte Dokumentenverwaltung	1.5/5
4	Implementierung der Lösung	Quellcode	Der Code zur Lösung, insbesondere Beschreibungen und Kommentare darin.	4/5
5	Referenz-KIS	Anwendung	Bietet Dienste zur Dokumentenverwaltung für den Anwender an.	2.5/5

Quellcodeanalyse zu identifizierenden Systemfunktionen ebenfalls mit den Ergebnissen der Interviews abgeglichen. Die Übereinstimmungsrate der Ergebnisse wird als ein Indikator für die Vollständigkeit und Integrität der gesammelten Systemfunktionen und Dienste betrachtet.

Identifizierte Dienste lassen nicht auf die dahinterstehenden Systemfunktionen schließen. Aufgrund dessen müssen die einzelnen Dienste bezüglich ihrer Relevanz zur Lösung der Dokumentenverarbeitung evaluiert werden. Dazu werden im dritten Schritt die Dienste und Systemfunktionen aufeinander abgebildet. Mit dieser Zuordnung und einer Plausibilitätsprüfung wird evaluiert, welche Dienste für die Dokumentenverarbeitung relevant sind und ob alle zwingend benötigten Systemfunktionen für diese Dienste identifiziert wurden. Zusätzlich wird mit der Prüfung, ob für alle Systemfunktionen mindestens ein Dienst

abgebildet wird, ein Hinweis zur Vollständigkeit der identifizierten Dienste gegeben.

3.3 Redokumentation

Anquetil et al. beschreiben in ihrem Redokumentation Prozess vier Aktivitäten, um von der High-Level Dokumentation zur Low-Level Dokumentation zu gelangen [5]. Die erste der vier Aktivitäten, ist die Dokumentation einer High-Level Ansicht des Referenz-KIS [5]. Da nur Teile des Referenz-KIS für die Problemstellung relevant sind, wird hierfür die Teildarstellung aus Abbildung 2 als die High-Level Ansicht betrachtet. Die zweite Aktivität beschreibt die Dokumentation von Querverweisen (engl. *cross references*), durch die Analyse der Abhängigkeiten in einem System [5]. Diese Aktivität ist ebenfalls in Abbildung 2 in dem für die Problemstellung notwendigen Umfang umgesetzt. In der dritten Aktivität sollen die Kernelemente (Ziele, Komponenten und Funktionen) eines Systems und dessen Subsysteme dokumentiert werden [5]. Im Rahmen dieser Arbeit werden hierfür nur (Kern-)Elemente betrachtet, die für die Erhebung der Spezifikationen der Lösung zur Dokumentenverarbeitung relevant sind. Dabei werden die in Abbildung 2 dargestellten Komponenten des Referenz-KIS in Zusammenhang mit der vierten Aktivität inspiziert, jedoch nicht separat dokumentiert. Mit der vierten Aktivität sollen die einzelnen Elemente, die während der Planung identifiziert wurden, unabhängig analysiert und dokumentiert werden [5]. Dies wird entsprechend der Beschreibung der Spezifikationserhebung in Kapitel 3.2 durchgeführt. Das Ziel dabei ist die Identifizierung aller Dienste und Systemfunktionen, die in der vorhandenen Lösung zur Dokumentenverarbeitung enthalten sind.

Die Quellcodeanalyse ergab die Identifikation von 20 Systemfunktionen, die zur Verarbeitung der klinschen Dokumente verwendet werden. Mit der Inspektion der einzelnen Systemkomponenten konnten 19 Dienste

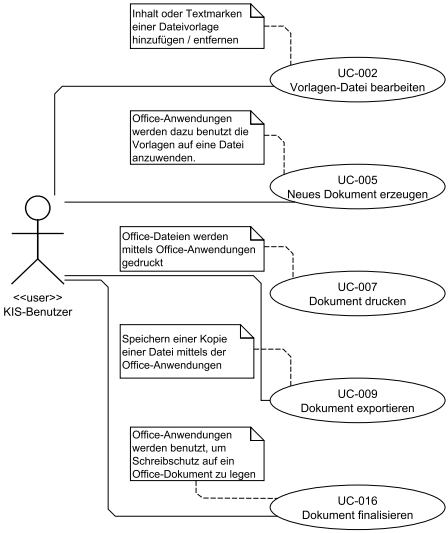


Abbildung 3: Identifizierte Dienste für KIS-Benutzer zur Verarbeitung klinischer Dokumente

rund um die Dokumentenverarbeitung identifiziert werden. Davon wurden anhand der erhobenen Systemfunktionen jedoch nur fünf Dienste als relevant für die Spezifikationserhebung in dieser Arbeit evaluiert (s. Kapitel 3.2). Diese fünf Dienste und die identifzierten Systemfunktionen sind als Anwendungsfälle in Abbildung 3 und Abbildung 4 dargestellt. Beim Abgleichen der Ergebnisse der Systemanalysen mit denen der Interviews wurde eine Übereinstimmungsrate von 89% (17/19) für die Dienste und 95% (19/20) für die Systemfunktionen erzielt. In beiden Fällen erweiterten die Systemanalysen das aus den Interviews gewonnene Wissen.

3.4 Bekannte Probleme

Während der Interviews mit den Personen-Wissensträgern wurden die bereits bekannten Probleme der bestehenden Lösung zur Dokumentenverarbeitung und -Verwaltung erhoben.

i Durch die Kombination der COM-Technologie und der Benutzerorientiertheit der Lösung wurde eine geteile

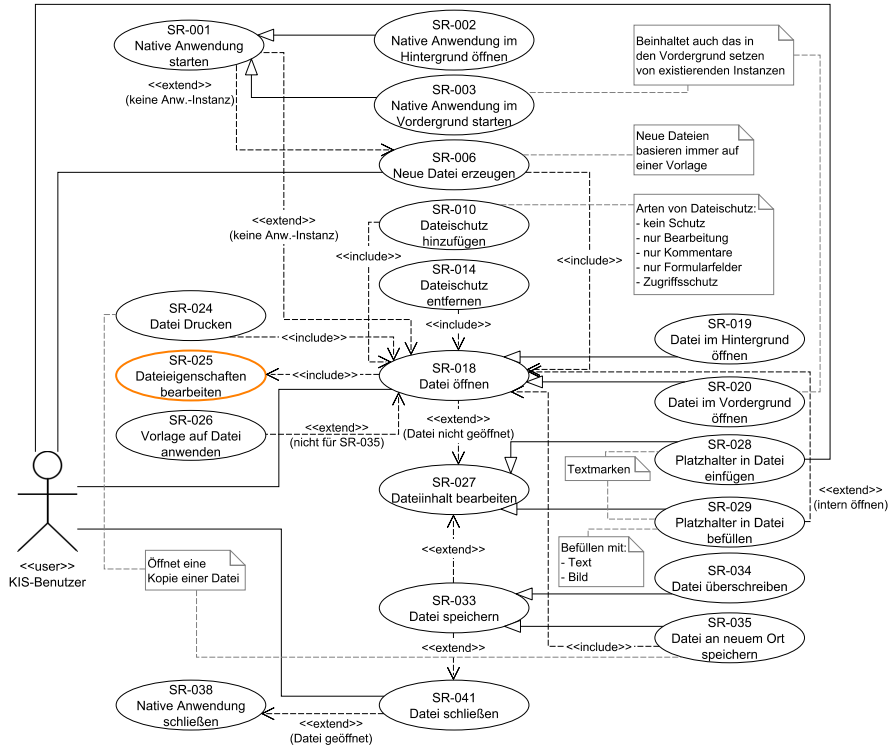


Abbildung 4: Systemfunktionen zur Verarbeitung klinischer Dokumente

Verantwortlichkeit für COM Objekte zwischen dem System und dem Benutzer geschaffen. Dadurch können Objekte vom Benutzer verworfen werden (engl. *disposed*), auf die das Referenz-KIS aktive Verweise hält. Dadurch ist die Lösung nur bedingt stabil und erfordert eine umfangreiche Fehlerabdeckung im Quellcode.

ii Nach Angaben der Entwickler ist die Verwendung der APIs von COM-basiertern Softwarekomponenten oder PIAs sehr unkomfortabel.

iii Microsoft rät davon ab, COM-basierte Office Automatisierung als zentralen Dienst umzusetzen[3]. Demnach kann die bestehende Lösung zur Dokumentenverarbeitung nicht zentralisiert werden.

iv Durch die Verteiltheit der Lösung zur Dokumentenverarbeitung sind für jede Client-Workstation des Referenz-KIS die Anwendungen Word und Excel notwendig. Die Lösung ist dadurch abhängig von Drittanbietersoftware, für die seitens des Kunden zusätzlich Lizenzen beschafft werden müssen.

v Die Verarbeitung der klinischen Dokumente ist vollständig dateibasiert. Aufgrund dessen ist zusätzlich zur COM-basierten Lösung im Referenz-KIS ein

[3] https://support.microsoft.com/de-de/help/257757/considerations-for-server-side-automation-of-office (Zuletzt geprüft am 10.09.2019)

Beisystem zur Verwaltung der Dokumente entwickelt worden. Dieses Beisystem wird im Wesentlichen durch ein *File Transfer Protocol* (FTP) Server und eigens entwickelten Dienst zur Dokumentenverwaltung dargestellt. Versions- und Paralletlitätskontrolle wären bislang ohne dieses Beisystem unmöglich.

vi Das zusätzliche Beisystem muss separat zu den Datenbanken des Referenz-KIS gesichert werden. Nach Angaben der Entwickler ist diese Sicherung weit aufwendiger als die einer Datenbank.

4 Diskussion und Ausblick

Die unterschiedlichen Ansätze zur Systemanalyse über das KIS als Anwendung und über den Quellcode des Systems haben sich lückenlos ergänzt. Die den Benutzern zur Verfügung stehenden Dienste und die internen Systemfunktionen zur Dokumentenverarbeitung konnten vollständig und plausibel aufeinander abgebildet werden. Zudem haben sich die Ergebnisse der Systemanalysen und den Interviews weitgehend gedeckt. Die erhobenen funktionellen Spezifikationen der im Referenz-KIS bestehenden Lösung zur Dokumentenverarbeitung können demnach als möglichst vollständig betrachtet werden.

In weiterführenden Arbeiten zur Modernisierung dieser Lösung ist darauf zu achten, dass diese Ergebnisse angemessen verwendet werden. Die Analyse zeigt den aktuellen Funktionsumfang der Lösung. Da keine Dokumentationen vorhanden sind, kann nicht nachvollzogen werden ob die Lösung entsprechend den Anforderungen der Benutzer über die Zeit angepasst oder erweitert wurde. Ein Indiz dafür, dass dies nicht der Fall ist, liefern die bekannten Probleme, die teilweise im Widerspruch zu den Lösungsspezifikationen stehen. Beispielsweise wird die Nicht-Zentralisierung der Lösung bemängelt, wohingegen diese stark benutzerorientiert und verteilt konzeptioniert ist. Die Spezifikationen sind als Maßstab oder zum Vergleich für moderne Lösungen daher ungeeignet.

Dem KIS-Hersteller wird bei der Modernisierung empfohlen, die Anpassung der gesamten Dokumentenverwaltung an den Stand der Technik zu betrachten. Im heutigen Gesundheitswesen sind der Wunsch und Drang nach Vernetzung stärker denn je. Indikatoren dafür sind die immer weiträumiger etablierten Standards zum Austausch medizinischer Daten. Beispiele hierfür sind die Arbeiten von Huang et al [8], von Muller et al [9] und Kim et al [10] zur besseren Austauschbarkeit medizinischer Dokumente. Vorreiter für Medizinstandards ist die Organisation *Health Level Seven International*[4] (HL7) mit den Standards HL7 Version 2 (V2), HL7 Version 3 (V3), HL7 Fast Health Interop Resources (FHIR), HL7 Clinical Document Architecture (CDA), Electronic Health Records (EHR) und weiteren. Diese Standards ermöglichen effiziente Kommunikation und Austausch zwischen medizinischen Institutionen und sind bereits in vielen Institutionen zum standardisierten Verwalten klinischer Dokumente eingesetzt, wie in den Arbeiten von Lin und Liou [11], Gilia et al [12] und Li et al [13] beschrieben.

5 Fazit

Die Systemanalyse zur Erhebung der Spezifikationen der bestehenden Lösung zur Dokumentenverarbeitung im Referenz-KIS war ein erster und wichtiger Schritt zur Modernisierung dieser Lösung. Dabei wurde auch eine Systemkarte der Integration dieser Lösung über die einzenen Systemkomponenten des Referenz-KIS gewonnen. Die Systemkarte zeigt alle Systempunkte die bei einer Modernisierung berücksichtigt werden müssen. Bevor allerdings ein Austausch oder eine Anpassung der Lösung vorgenommen wird, sollte eine angemessene Anforderungsanalyse der Benutzerbedürfnisse und der klinischen

[4] http://www.hl7.org/index.cfm (Zuletzt geprüft am 23.09.2019)

Prozesse durchgeführt werden. Zudem sollten Standardisierungsmöglichkeiten für das Dokumenten- und der damit verbundenen Datenmanagement trotz des damit verbundenen hohen Aufwands als mögliche Lösung betrachtet werden.

Literatur

[1] M. Corporation. (2018). Component Object Model (COM), Adresse: https://docs.microsoft.com/en-us/windows/win32/com/component-object-model--com--portal (besucht am 26. 08. 2019).

[2] D. Box, *Essential COM* -, 01. Aufl. Boston: Addison-Wesley Professional, 1998, ISBN: 978-0-201-63446-4.

[3] P. J. Leach, R. Salz und M. H. Mealling, *A Universally Unique IDentifier (UUID) URN Namespace*, RFC 4122, Juli 2005. DOI: 10.17487/RFC4122. Adresse: https://rfc-editor.org/rfc/rfc4122.txt (besucht am 26. 08. 2019).

[4] R. Thurlow, "RPC: Remote Procedure Call Protocol Specification Version 2", RFC Editor, Techn. Ber., Mai 2009, http://www.rfc-editor.org/rfc/rfc5531.txt. Adresse: http://www.rfc-editor.org/rfc/rfc5531.txt (besucht am 26. 08. 2019).

[5] N. Anquetil, K. M. de Oliveira, A. G. M. dos Santos, P. C. S. da Silva Jr., L. C. de Araujo Jr. und S. D. C. F. Vieira, "Software Re-Documentation Process and Tool", in *The 17th Conference on Advanced Information Systems Engineering (CAiSE '05), Porto, Portugal, 13-17 June, 2005, CAiSE Forum, Short Paper Proceedings*, 2005. Adresse: http://ceur-ws.org/Vol-161/FORUM%5C_16.pdf (besucht am 30. 08. 2019).

[6] C.-T. Su und D. Yeh, "Software architecture recovery and re-documentation tool of a Hospital Information System", in *2012 International Conference on Computer and Communication Engineering (ICCCE)*, 2012 International Conference on Computer und Communication Engineering (ICCCE), IEEE, Juli 2012. DOI: 10.1109/iccce.2012.6271169.

[7] G. CanforaHarman und M. D. Penta, "New Frontiers of Reverse Engineering", in *Future of Software Engineering (FOSE 07)*, Future of Software Engineering (FOSE 07), IEEE, Mai 2007. DOI: 10.1109/fose.2007.15.

[8] E.-W. Huang, T.-L. Tseng, M.-L. Chang, M.-L. Pan und D.-M. Liou, "Generating Standardized Clinical Documents for Medical Information Exchanges", *IT Professional*, Jg. 12, Nr. 2, S. 26–32, März 2010. DOI: 10.1109/mitp.2010.56.

[9] M. Muller, T. Frankewitsch, T. Ganslandt, T. Burkle und H. U. Prokosch, "The Clinical Document Architecture (CDA) enables electronic medical records to wireless mobile computing", *Stud Health Technol Inform*, Jg. 107, Nr. Pt 2, S. 1448–1452, 2004.

[10] H. Kim, T. Tran und H. Cho, "A Clinical Document Architecture (CDA) to Generate Clinical Documents within a Hospital Information System for E-Healthcare Services", in *The Sixth IEEE International Conference on Computer and Information Technology (CIT'06)*, IEEE, 2006. DOI: 10.1109/cit.2006.2.

[11] C.-J. Lin und D.-M. Liou, "Design and implement a generator with Clinical Document Architecture standard", in *2010 Second International Conference on Communication Systems, Networks and Applications*, IEEE, Juni 2010. DOI: 10.1109/iccsna.2010.5588765.

[12] G. Gilia, V. Mistretta, C. Muriana, T. Piazza und G. Vizzini, "Improving interoperability of clinical documents: A case study of LOINC mapping in analysis laboratories", in *2015 International Conference on Industrial Engineering and Systems Management (IESM)*, Okt. 2015, S. 1240–1244. DOI: 10.1109/IESM.2015.7380311.

[13] H. Li, H. Duan, X. Lu und Z. Huang, "A Clinical Document Repository for CDA Documents", in *2007 1st International Conference on Bioinformatics and Biomedical Engineering*, Juli 2007, S. 1084–1087. DOI: 10.1109/ICBBE.2007.280.

Konzeption und Durchführung einer Umfragereihe zur Erfassung von DevOps-Aspekten in Softwareentwicklungsprojekten

Georgia König
Hochschule Reutlingen
Georgia_Marie-Theres.Koenig@Student.Reutlingen-University.de

Abstract

Der DevOps-Ansatz und mit diesem assoziierte Inhalte und Praktiken finden immer mehr Anhänger, was Zahlen zur aktuellen DevOps-Adaptionsrate in Unternehmen implizieren. Die vorliegende Arbeit behandelt die Konzeption, Durchführung und Auswertung der DevOps@Novatec Umfrage. Dabei handelt es sich um eine Experteninterview-Reihe mit dem Ziel, die aktuelle Ausprägung von Fähigkeiten, die mit DevOps assoziiert werden in Softwareentwicklungsprojekten zu erfassen. Im Rahmen des Projektes wurden sowohl ein Fragebogen, als auch eine Vorgehensweise zur Auswertung der gesammelten Daten konzipiert und implementiert.

CCS Concepts

• **General and reference** → **Surveys and overviews**; *Evaluation*; • **Social and professional topics** → *Quality assurance*;

Keywords

DevOps, maturity models, continuous improvement, business process management

1 Einleitung

In den letzten Jahren ist die Anzahl der Softwareunternehmen, die mit DevOps assoziierte Praktiken adaptieren oder eine solche Adaption zumindest erwägen, stetig gestiegen. Diese Adaption ist häufig ausgelöst durch den Druck, den eine sich stetig wandelnde Geschäftsumgebung, in Bezug auf Wandel innerhalb des Wettbewerbs, auf diese Firmen ausübt [1]. Die damit verbundenen Anforderungen des Marktes drängen zu einer immer schnelleren time to market für Softwareentwicklungsprojekte. Dies ist ein Ziel, welches auch die DevOps-Bewegung verfolgt und zu dessen Erreichung diverse Werte, Methoden und Best Practices propagiert werden [2]. DevOps und die damit assoziierten Inhalte sollen die Barrieren zwischen Mitarbeitern innerhalb der Entwicklung und dem IT-Betrieb beseitigen und es somit erleichtern, Services kontinuierlich an Kunden zu liefern [1].
Das Unternehmen Puppet Labs veröffentlicht jährlich den 'State of DevOps Report'. Die Autoren dieses Whitepapers befragten 2017 rund 3.200 Fachleute zu ihrer täglichen Arbeit, darunter Führungskräfte, Entwickler, Tester und andere IT-Experten. Aus dem Bericht geht hervor, dass die Zahl der DevOps-Mitarbeiter stark gestiegen ist. Im Jahr 2014 arbeiteten 16 Prozent der Befragten im Themenfeld DevOps, 2017 waren es 27 Prozent. Der Bericht ermöglicht zudem die folgenden

Betreuer Hochschule:	Professor Hertkorn Hochschule Reutlingen Peter.Hertkorn@Reutlingen-University.de
Betreuer Firma	Dominik Pabst Novatec Consulting GmbH info@novatec-gmbh.de

Wissenschaftliche Vertiefungskonferenz
27. November 2019, Hochschule Reutlingen
Copyright 2019 Georgia König

Einblicke: IT -Unternehmen, die DevOps implementiert haben, sind in der Lage 46-mal häufiger Code Deployments bereit zu stellen. Ihre Änderungsdurchlaufzeit (lead time from commit to deploy) ist dabei 440-mal schneller im Vergleich mit Unternehmen ohne DevOps-Adaption. Ebenso verzeichnen Unternehmen, die DevOps einsetzen, eine 170 mal schnellere mittlere Zeit bis zur Wiederherstellung nach Ausfallzeiten (mean time to recover from downtime), sowie eine 5 mal niedrigere Änderungsausfallrate (change failure rate) [3].

1.1 Motivation

Vor dem geschilderten Hintergrund beschäftigt sich auch die Novatec Consulting GmbH mit DevOps-assoziierten Inhalten und wie diese in IT-Projekten umgesetzt werden. Dabei handelt es sich um ein unabhängiges IT-Consulting Unternehmen, welches Kunden verschiedenster Branchen bei der Durchführung von Vorhaben im IT-Umfeld berät und unterstützt. Hierfür ist es notwendig, einen Überblick über den aktuellen Stand bezüglich DevOps in den betreuten Softwareentwicklungsprojekten zu erlangen. Auf dessen Grundlage können angemessene Handlungsempfehlungen abgeleitet werden. Konkreter gesagt besteht die Motivation darin, die aktuelle Ausprägung von ausgewählten DevOps-assoziierten Werten, Praktiken und Technologiepräferenzen in Softwareentwicklungsprojekten zu erfassen und zu bewerten. Zu diesem Zweck wurde in Kooperation und mit der Unterstützung des Unternehmens die sogenannte DevOps@Novatec Umfrage konzipiert und mit fünf Projekten durchgeführt, sowie ausgewertet.

1.2 Ziele der Arbeit

Die vorliegende Arbeit beschreibt die Konzeption, Durchführung und Auswertung der DevOps@Novatec Umfrage. Dabei handelt es sich um eine Reihe an Experteninterviews, die mit Mitarbeitern der Novatec durchgeführt und anschließend ausgewertet wurden. Im Rahmen des Projektes wurde unter anderem ein Fragenkatalog, ein Online-Fragebogen und eine Vorgehensweise zur Auswertung der gesammelten Daten konzipiert, implementiert und dokumentiert. Die vorliegende Arbeit beschreibt das Vorgehen bei der Durchführung, stellt die entstandenen Artefakte vor und gibt einen Ausblick hinsichtlich der zukünftigen Nutzung.

2 DevOps

2.1 Begriff

Bei dem Begriff DevOps handelt es sich um ein Kofferwort aus Softwareentwicklung (Development) und IT-Betrieb (Operations). Der Begriff symbolisiert damit zugleich den Grundgedanken und die Zielsetzung der gleichnamigen Bewegung: DevOps repräsentiert einen Schulterschluss zwischen den beiden, in der traditionellen Wahrnehmung separierten Abteilungen innerhalb des Software-Lebenszyklus. Der Namensgeber war dabei Patrick Debois, der erstmals 2009 die 'DevOpsDays' veranstaltete, eine Konferenz zu diesem Themenfeld in Gent [4]. Grundlegend für das Verständnis von DevOps ist, dass die Frage danach, wie ein jeweiliger Software-Lebenszyklus konkret aufgebaut und durchgeführt werden muss, um DevOps-konform zu sein, nicht zielführend ist. DevOps ist keine weitere Rolle, Team, Abteilung, Toolchain, Prüfsiegel oder Regulierung, die eingestellt, eingekauft oder erfüllt werden kann [2]. In [1] wird DevOps mittels drei, sich gegenseitig unterstützenden Elementen definiert. Dazu zählen *Capabilities, Cultural Enablers* und *Technological Enablers*. Dabei wird argumentiert, dass Capabilities den Hauptaspekt von DevOps ausmachen. Capabilities umfasst dabei Fähigkeiten wie kontinuierliche Planung, Entwicklung, Integration und

Tests, Freigabe und Bereitstellung, Infrastrukturüberwachung und -optimierung, Überwachung und Optimierung des Nutzerverhaltens sowie Feedback und unverzügliche Wiederherstellung von Dienstausfällen. Um diese Fähigkeiten zu stärken, bedarf es einer Prozessverbesserung in verschiedenen Aspekten. Allerdings bedeutet Prozessverbesserung für verschiedene Unternehmen aufgrund der spezifischen Probleme, mit denen sie konfrontiert sind, verschiedene Dinge. Zu diesem Zweck ist ein relativ gründliches Verständnis über die Leistung des aktuellen Prozesses und eine effektive Anleitung für zukünftige Prozessverbesserungen erforderlich. Um dieser Anforderung gerecht zu werden, wurden verschiedene Referenzmodelle entwickelt und übernommen, darunter das CMM (Capability Maturity Model) und seine neue Version CMMI (Capability Maturity Model Integration), um den aktuellen Prozess zu bewerten und die zukünftige Prozessverbesserung zu steuern. [1]

2.2 Bewegungsstiftender Konflikt

DevOps als Bewegung wird oft über den bewegungsstiftenden Konflikt erklärt. Die erhöhte Frequenz, mit der Software heutzutage ausgeliefert werden kann, verstärkt indirekt einen Konflikt, der branchenweit bereits seit langem besteht, jedoch bis dato größtenteils unter Kontrolle gehalten werden konnte. Gemeint ist der Konflikt zwischen Entwicklung und Betrieb, der ein zentraler Motivationsfaktor zur Entstehung der DevOps-Bewegung war. Um diesen Konflikt zu verstehen, reicht es, die diametral entgegengesetzte Motivation und Zielsetzung der beteiligten Parteien zu betrachten [4]. Auf der einen Seite steht das Development, welche Funktionalitäten immer schneller bauen und liefern können. Auf der anderen Seite steht der IT-Betrieb, dessen übergeordnetes Ziel darin besteht, die Software als Ganzes zu betreiben. Da das Produkt möglichst stabil und sicher laufen soll, stellt jede Änderung eine potenzielle Gefahr und gleichzeitig einen entsprechenden Arbeitsaufwand dar. Der Betrieb wird seitens des Unternehmens für Stabilität der anvertrauten Systeme belohnt, während die Entwicklung für Veränderung belohnt wird. Zusammenfassend lässt sich sagen, dass es einen Balanceakt darstellt, gelieferte Services langfristig auf einem Niveau zu halten, welches Kunden- und Marktanforderungen entspricht. Auf der einen Seite sollen unerwünschte Auswirkungen von Veränderungen vermieden werden, um Stabilität zu gewährleisten. Auf der anderen Seite gilt es, notwendige Innovationen und Veränderungen zu ermöglichen und zu unterstützen. [2]

Die DevOps-Bewegung beschreibt grundlegend Ansätze, um die so genannte Wall of Confusion (dt. Wand aus Verwirrung) zwischen fachlichen Silos einzureißen, sowie den gesamten Software-Lebenszyklus für alle Akteure zu optimieren [4]. Neben der Überwindung des zuvor beschriebenen Konflikts lassen sich diese weiteren, übergeordneten Ziele von DevOps zusammenfassen: Der Weg von der Idee bis zum Kunden soll stromlinienförmiger gemacht werden, d.h. das häufige und sichere Ausrollen von Software soll ermöglicht und erleichtert, die Stabilität und Sicherheit dieser Software soll erhöht und besonders die Arbeitsbedingungen aller Akteure des gesamten Software-Lebenszyklus sollen mittels Werten, Praktiken und Tools kontinuierlich verbessert werden [2].

2.3 24 Key Capabilities

Für die Auswertung wurden die 24 Key Capabilities (KC's) nach [5] genutzt. Dabei handelt es sich um 24 Handlungsempfehlungen, beziehungsweise eine Sammlung an empfohlenen Praktiken mit demselben Ziel aus dem Themenfeld DevOps. Es sind diese Fähigkeiten welche, wenn konsequent umgesetzt, in untersuchten Unternehmen zu signifikanten, messbar besseren Ergebnissen im Bereich der Software Delivery Performance und damit

der organisatorischen Leistung führten. Diese Capabilities sind einfach zu definieren, zu messen und zu verbessern. [5]
Die Tabelle 1 listet alle KC's, welche jeweils einer von 5 Kategorien zugeordnet sind.

Tabelle 1: 24 Key Capabilities nach [5]

Nr.	Key Capability
	Continuous Delivery Capabilities
1	Use version control for all production artifacts
2	Automate your deployment process
3	Implement continuous integration
4	Use trunk-based development methods
	Architecture Capabilities
5	Implement test automation
6	Support test data management
7	Shift left on security
8	Implement continuous delivery (CD)
9	Use a loosely coupled architecture
	Product and Process Capabilities
10	Architect for empowered teams
11	Gather and implement customer feedback
12	Make the flow of work visible through the value stream
13	Work in small batches
14	Foster and enable team experimentation
15	Have a lightweight change approval processes
	Lean Mgnt. and Monitoring Capabilities
16	Monitor across application and infrastructure to inform business decisions
17	Check system health proactively
18	Improve processes and manage work with work-in-process (WIP) limits
19	Visualize work to monitor quality and communicate throughout the team
	Cultural Capabilities
20	Support a generative culture (as outlined by Westrum)
21	Encourage and support learning
22	Support and facilitate collaboration among teams
23	Provide resources and tools that make work meaningful
24	Support or embody transformational leadership

Die in der Tabelle 1 dargestellte Zuordnung einzelner KC's zu Kategorien wurde für die spätere Datenauswertung übernommen, um eine zusätzliche Ebene an Abstraktion zu schaffen.

3 DevOps @ Novatec Umfrage

Im folgenden Kapitel wird die DevOps @ Novatec Umfrage vorgestellt. Es wird auf ausgewählte Details eingegangen, um deren Inhalt, Umfang und die beteiligten Artefakte vorzustellen.

3.1 Projektumfeld

Der Name DevOps@Novatec Umfrage fasst bereits den Zweck des zugehörigen Projektes grob zusammen: Dieser liegt darin, das Erfassen und Bewerten des Ist-Zustandes in Bezug auf DevOps-assoziierte Praktiken in diversen Softwareentwicklungsprojekten der Novatec Consulting GmbH zu ermöglichen.
Es sollen also zwei Unterziele erreicht werden: Erstens soll eine Möglichkeit gefunden werden, den aktuellen Ist-Zustand angemessen zu ermitteln. Zweitens soll der ermittelte Ist-Zustand mit klar definierten, ausgesuchten Referenzwerten beziehungsweise Best Practices aus der DevOps-Literatur verglichen werden. Die Umfrage sollte die eigenen Ergebnisse mit einem angemessenen, belegbaren Ideal vergleichen und somit aus- und bewerten.
Die Durchführung der Umfrage bietet Unterstützung in den folgenden Punkten: Erstens wird ein besserer Überblick über interne Kompetenzen, Stärken und Unzulänglichkeiten im Hinblick auf DevOps geschaffen. Zweitens wird ein Verständnis dafür geschaffen, welche Praktiken aktuell in einem jeweiligen Projekt wie ausgeprägt sind. Auf dessen Grundlage lassen sich beispielsweise auch Kundenanfragen differenzierter entscheiden.

4 Datenerhebung

Im folgenden Abschnitt werden die wichtigsten Maßnahmen für die Planung und Durchführung der Datenerhebung im Rahmen der DevOps@Novatec Umfrage erläutert.

4.1 Experteninterviews

Als Erhebungsmethode wurden Experteninterviews gewählt, welche eine spezielle Form

der Leitfadeninterviews darstellen. Leitfrageninterviews sind dabei eine Methode aus der qualitativen Interviewforschung. Der verwendete Interviewleitfaden dient dazu, den Gesprächsverlauf zu strukturieren [6].
Bei Experteninterviews ist der Interviewpartner durch seine Funktion als Experte in einem bestimmten Themenfeld von Bedeutung für die qualitative Befragung. Das Experteninterview bezieht sich auf einen klar definierten Wirklichkeitsausschnitt. Der Befragte wird nicht als Einzelfall, sondern als Repräsentant einer Gruppe in die Untersuchung einbezogen. [7]
Im Rahmen der erstmaligen Durchführung der DevOps@Novatec Umfrage wurden fünf Experteninterviews abgehalten. In diesen etwa 1 ½ stündigen Interviews wurde der in Abschnitt 4.2. beschriebene Fragenbogen mit dem oder den jeweiligen Interview-Partnern ausgefüllt. Die Auswahl der Interviewpartner basierte dabei auf der Auswahl der teilnehmenden Softwareprojekte. Dabei handelt es sich entweder um den jeweiligen Team-Lead oder einen, von dieser Person benannten, Stellvertreter.
Bei der Interviewform der Experteninterviews spielt der Leitfaden eine besondere Rolle, denn er dient einerseits dazu den Befragten auf das interessierende Expertentum zu begrenzen bzw. festzulegen und andererseits um als Interviewer zu verdeutlichen, mit der Thematik vertraut zu sein. [7]
Auf die Konzeption des Fragebogen, sowie dessen Implementierung in LimeSurvey wird im Folgenden eingegangen.

4.2 Fragebogen

Der verwendete Fragebogen basiert auf der sogenannten DevOps X-Ray individual assessment Online-Umfrage des DevOps Research and Assessment Instituts.
Dieser Fragebogen wurde in Absprache mit allen Projektbeteiligten als inhaltlicher Grundstein für die Konzeption und Implementierung eines eigenen Fragebogens für die DevOps@Novatec Umfrage benutzt. Alle englischen Original-Fragen wurden zunächst analysiert und wenn möglich einer Literaturquelle zugeordnet. Im nächsten Schritt wurden einzelne Fragengruppen ausgeschlossen. Innerhalb der Auswahl relevanter Fragengruppen wurden einzelne Fragen gekürzt, erweitert oder konkretisiert. Zudem wurden die Antwortmöglichkeiten der meisten Fragen abgeändert und erweitert, um eine größere Vielfalt an Daten erheben zu können. Dieses Unterfangen benötigte mehrere Iterationen und Anpassungen bis der richtige Detailgrad gefunden wurde. Zuletzt wurden noch Fragen hinzugefügt, die es ermöglichen sollten, alle relevanten Daten zum teilnehmenden Projekt zu erfassen. Ebenso wurden Fragen hinzugefügt, die explizit DevOps-Vorwissen, sowie eine „DevOps-Arbeitsdefinition“ des Kunden und des befragten Experten abfragten. Das Ergebnis all dieser Iterationen war ein angepasster, ins Deutsche übersetzter Fragenkatalog der im nächsten Schritt implementiert werden konnte. Darauf wird im folgenden Abschnitt eingegangen. Der Fragebogen wurde mittels LimeSurvey implementiert. Die Herausforderung bezüglich der Implementierung bestand dabei in der angemessenen Auswahl von Fragen- und Antwortentypen aus den von LimeSurvey angebotenen Möglichkeiten. Die aktuelle Version des Fragebogens besitzt die folgenden Charakteristiken: Insgesamt sind 75 Fragen enthalten, in diesen werden 111 Datenpunkte erhoben. Davon sind 61 Fragen gewichtet, d.h. Fragen zum Projektnamen, Dauer und ähnlichem sind dabei außerhalb der Wertung. Die Fragen sind thematisch in fünf Fragengruppen (Seiten innerhalb des LimeSurvey-Fragebogens) gegliedert.
Es folgt eine Kurzbeschreibung der verwendeten fünf Fragengruppen, um einen Eindruck für den Inhalt des Fragebogens zu vermitteln. Die Fragengruppe 1 „Projektdaten“ enthält 14 Fragen zum Projekt und Interviewpartner selbst, wie z.B. Dauer, Mitarbeiteranzahl, DevOps-Vorwissen des Teams und Kunden.

Die Fragengruppe 2 „Automation / Tooling“ enthält 20 Fragen zu Themen wie der Handhabung von Version Control, Testautomatisierung und der Autonomie des Teams.
Die Fragengruppe 3 „Development“ fragt mit 20 Fragen die Gestaltung des Entwicklungsprozesses ab, zum Beispiel mit Fragen zu den Themen Arbeiten mit minimum viable products (MVPs), Visual Dashboards und Information Security.
In der Fragengruppe 4 „Monitoring, Notification and Lean Management“ sind 11 Fragen enthalten, die unter anderem die Handhabung des Incident Managements mittels Logging- und Überwachungssystemen abfragen.
Zuletzt besteht die Fragengruppe 5 „Daily Work“ aus 10 Fragen mit einem Schwerpunkt auf Code Deployments, wobei deren Häufigkeit, Art und Aufwand erfasst wird. Zudem gibt es für jede Frage die Möglichkeit einen Kommentar anzugeben, um weitere Kontextinformationen erfassen zu können.
Die erfassten Daten wurden im Anschluss an die jeweiligen Interviews ausgewertet, worauf das nächste Kapitel eingeht.

5 Auswertung

Nachdem im vorherigen Kapitel die Datenerhebung betrachtet wurde, beschäftigt sich dieses Kapitel mit deren Auswertung. Um die erhobenen Daten bewertbar und vergleichbar zu machen, wurden primär zwei Maßnahmen getroffen: Zum einen wurde jeder möglichen Antwortmöglichkeit jeder Frage eine Gewichtung zugeordnet, das heißt mit einem Prozentwert verknüpft. Zum anderen wurde jede gestellte Frage einer Key Capability nach [5] zugeordnet. Beide Maßnahmen werden im Folgenden kurz beschrieben. Das konkrete Ergebnis dieser Arbeit sind ein gewichteter Fragenkatalog und eine Vorgehensweise, um die erfassten Daten anhand klar definierter Kategorien zu vergleichen. Um die Auswertung künftig erhobener Datensätze zu unterstützen, wurde eine ausfüllbaren Excel-Tabelle implementiert. Diese enthält dabei explizit die im Folgenden beschriebene Gewichtung aller Fragen und deren Zuordnung zu übergeordneten Key Capabilities.

5.1 Gewichtung

Dieser Abschnitt beschreibt die Gewichtung einzelner Antworten als Vorbedingung für eine spätere Vergleichbarkeit. Zu diesem Zweck wurde jede mögliche Antwortmöglichkeit jeder Frage mit einem Prozentwert verknüpft und somit gewichtet. Dadurch lassen sich zwei Antworten auf eine Frage vergleichen. Um diese Gewichtung vorzunehmen, war es notwendig Antworten abgestuft in Kategorien von ‘Besser‘ oder ‘Schlechter‘ zu bewerten. Hierzu war es wiederum nötig ein „Ideal“ festzulegen, anhand dessen eine Bewertung erfolgte. Alle Arten von Gewichtung wurden als tabellarische Zuordnung zwischen allen Antwortmöglichkeit einer Frage und Prozentwerten festgehalten. Aufgrund der Vielzahl der verwendeten Fragen- und Antworttypen und da überwiegend Likert-Skalen verwendet wurden, handelte es sich bei der Gewichtung um einen sehr umfangreichen Prozess, der mehrfach überprüft werden musste. Nachdem in diesem Abschnitt erklärt wurde, wie die Antwortmöglichkeiten gewichtet werden, um diese vergleichbar zu machen, wird im nächsten Abschnitt auf eine Zuordnung der Fragen in klar definierte Kategorien eingegangen.

5.2 Zuordnung

Die Zuordnung der Fragen zu einer jeweiligen KC wurde unter anderem deshalb durchgeführt, um eine zusätzliche Ebene an Struktur innerhalb der Datenauswertung zu erhalten. Die Zuordnung zwischen Kategorie und KC ist dabei direkt aus der Literatur entnommen. Im Gegensatz dazu ist die Zuordnung zwischen Fragen und KC’s vom Autor in Absprache mit der Novatec festgelegt worden.

Dies ist bei der Interpretation der Ergebnisse stets zu beachten und kann in zukünftigen Auswertungen geändert werden, falls zusätzliche Fragen oder Fragengruppen hinzukommen. Ebenso ist zu beachten, dass manchen KC's keine Fragen zugeordnet sind. Die nachfolgende Erläuterung soll die Zuordnung zwischen Kategorien, Key Capabilities und verwendeten Fragen verdeutlichen. Bei der vorgenommenen Zuordnung kann jede Kategorie 0-n KC's besitzen, jede KC jedoch nur 0-1 Kategorien zugeordnet sein. Ebenso verhält es sich mit der Beziehung zwischen KC und Frage, denn jede KC kann 0-n Fragen enthalten, jeder Frage ist jedoch nur 0-1 KC zugeordnet.

5.3 Excel-Tabelle

Nach der Durchführung der zuvor beschriebenen Schritte war es notwendig, eine klare, nachvollziehbare Vorgehensweise für die Auswertung festzuhalten. Um dies zu garantieren, wurde eine Excel-Tabelle auf Basis der vorherigen Arbeit implementiert. Um zu verdeutlichen, wie sich diese nutzen lässt, folgt ein Beispiel: Die Frage „Wie viele aktive Branches existieren im Code-Repository der Application oder Service?" hat in der internen Nummerierung die Nummer 1.7. und ist der Key Capability Nr. 4 (Use trunk-based development methods) zugeordnet. Diese KC ist wiederum der Kategorie 1 (Continuous Delivery Capabilities) zugeordnet. Der Wert dieser Frage und der anderen Fragen in dieser KC wurde in die Excel-Tabelle eingetragen, welche den Gesamtwert in Prozent errechnet und geordnet nach Key Capability anzeigt. Ein Ergebnis von 100% sagt dabei aus, dass ein Projekt dem festgelegten Ideal für diese KC völlig entspricht, ein Ergebnis unter 100% kann als Anlass für das Ableiten von Handlungsempfehlungen herangezogen werden.
Alle festgelegten Gewichtungen der Fragen, sowie deren Zuordnung zu KC's wurden in einer ausfüllbaren Excel-Tabelle implementiert. Dieses Artefakt enthält damit Ergebnisse aller zuvor bearbeiteten Teilschritte. Die Excel-Tabelle ermöglicht die folgenden Dinge: 1. Es lässt sich für jede Frage genau ablesen, welche Antwortmöglichkeiten existieren und zulässig sind. 2. Es lässt sich für jede Antwortmöglichkeit die zugehörige Gewichtung ablesen. 3. Sollten weitere Fragen oder Antwortmöglichkeiten hinzugefügt werden, kann leichter auf die vorhandene Arbeit aufgebaut werden. 4. Die Auswertung zukünftiger Datensätze wird automatisch errechnet. Damit erfolgt diese schneller und ist weniger anfällig für Fehler.

6 Zukünftige Nutzung

Das folgende Kapitel gibt einen Ausblick bezüglich der zukünftigen Nutzung der DevOps@Novatec Umfrage.

6.1 Reifegradmodelle

Eine Möglichkeit, die in dieser Arbeit beschriebenen Artefakte zu nutzen, ist als Vorarbeit für ein DevOps-Reifegradmodell. Der nachfolgende Abschnitt enthält eine kurze Einführung in das Wesen und die Nutzung von Reifegradmodellen (engl. Maturity Modells) nach [8].
Basierend auf der Annahme vorhersehbarer Entwicklungs- und Veränderungsmuster beinhalten Reifegradmodelle in der Regel eine Abfolge von Ebenen (oder Stadien), die zusammen einen erwarteten, gewünschten oder logischen Weg von einem Ausgangszustand zur Reife bilden. In diesem Zusammenhang geben die einzelnen Reifegrade die aktuellen (oder wünschenswerten) Fähigkeiten eines Unternehmens in Bezug auf eine bestimmte Klasse von Unternehmen an. Reifegradmodelle werden üblicherweise angewendet, um die Ist-Situation zu beurteilen, Verbesserungsmaßnahmen abzuleiten und zu priorisieren und den Fortschritt zu kontrollieren.[8]
Die im Rahmen der DevOps@Novatec Umfrage entstandenen Artefakte dienen als Vorarbeit für die Konzeption eines deskriptiven

DevOps-Reifegradmodells. Ein Reifegradmodell dient einem deskriptiven, d.h. beschreibenden Verwendungszweck, wenn es für Ist-Analysen eingesetzt wird, bei denen die aktuellen Fähigkeiten des untersuchten Unternehmens anhand vorgegebener Kriterien bewertet werden. Das Reifegradmodell wird als diagnostisches Instrument eingesetzt und die zugewiesenen Reifegrade können an interne und externe Stakeholder berichtet werden.[8]
Seit das Software Engineering Institute vor fast zwanzig Jahren das Capability Maturity Model (CMM) auf den Markt gebracht hat, wurden Hunderte von Reifegradmodellen von Forschern und Praktikern aus verschiedenen Bereichen vorgeschlagen und betrachtet, so unter anderem in [9].
Reifegradmodelle zielen beispielsweise darauf ab, Unternehmen mit digitaler Verwaltung, IT-Management oder Wissensmanagement zu unterstützen. Auch im Bereich des Business Process Management (BPM) wurden diverse Reifegradmodelle vorgeschlagen. Dies liegt wahrscheinlich in der hohen Bedeutung von Prozessorientierung und kontinuierlicher Prozessverbesserung für die Organisationsgestaltung begründet. In der Praxis wird erwartet, dass die allgemeine Akzeptanz von Reifegradmodellen zunehmen wird, was durch die zahlreichen proprietären Modelle von Softwareunternehmen und Beratungsunternehmen bestätigt wird. Auch die jüngste Literatur berichtet von einem zunehmenden wissenschaftlichen Interesse an Reifegradmodellen.[8]

6.2 Interpretation der Ergebnisse

Die Ergebnisse der Auswertung für die teilnehmenden fünf Softwareentwicklungsprojekte wurde firmenintern kommuniziert, können in dieser Ausarbeitung jedoch nicht näher beschrieben werden. Naheliegend ist natürlich die Nutzung des entstandenen Setups, um weitere Daten in Softwareentwicklungsprojekten zu erheben und auszuwerten. Für die Interpretation dieser Ergebnisse sind die folgenden Punkte erwähnenswert: Die Ergebnisse der Auswertung der teilnehmenden Projekte lassen sich nutzen als Visualisierung einer Momentaufnahme in Bezug auf DevOps-assoziierte Praktiken. Die Rohdaten der Interviews, d.h. die Antworten, besonders in Verbindung mit den gegebenen Kommentaren können genutzt werden um Kompetenzen, Stärken, Herausforderungen und Mankos im Projekt aufzuzeigen. Eine Stufe darüber können die Ergebnisse der einzelnen KC's genutzt werden um zu filtern, welche DevOps-assoziierten Inhalte oder Best Practices besonders explizit praktiziert werden (und warum), sowie welche bislang vernachlässigt praktiziert werden. Ein mögliches, wenn auch stark kondensiertes Ergebnis ist die Summe der Prozentsätze aus allen gewerteten KC's. Diese Summe lässt sich mit dem durch die Literatur gegebenen Ideal von 100% vergleichen. Ein genaueres Bild ergibt sich dabei natürlich durch den Vergleich einzelner Capabilities. Die Interpretation der Ergebnisse kann also mittels drei verschiedener Sichten in zunehmender Granularität vorgenommen werden: Nach Kategorie, nach KC und auf Ebene der einzelnen Fragen.
Die DevOps@Novatec Umfrage, deren Artefakte und die zugehörige Dokumentation kann abschließend als eine generelle Vorauswahl an, beziehungsweise eine Fokussierung auf Best Practices dienen. Diese kann ohne weitere Recherche für eine Evaluation in Projekten genutzt werden, zum Beispiel von Personen, welche sich erst seit kurzem mit DevOps-Inhalten beschäftigen.

7 Fazit

In diesem Abschnitt werden die wichtigsten Inhalte der Arbeit zusammengefasst. Die DevOps@Novatec Umfrage entstand in Kooperation mit und unter der Betreuung des Unternehmens Novatec Consulting GmbH. Das allgemeine Ziel des Projektes lag darin, einen Überblick über die Ausprägung von

DevOps-assoziierten Inhalten und Praktiken in diversen Novatec-Projekten zu ermöglichen. Als Methode der Erhebung wurden fragebogengestützte Experteninterviews ausgewählt und durchgeführt. Ein Vergleich mit Best Practices aus der Literatur sollte es ermöglichen, die einzelnen Projekte zu bewerten, um Mängel und Stärken erkennen zu können. Zu den abgefragten Praktiken gehörten unter anderem Version Control, Continuous Deployment, generative Lernkultur, agile Entwicklungsmethoden und Testautomatisierung. Ein Fragenkatalog des DevOps Research and Assessment Instituts wurde als Grundlage für die Konzeption und Implementierung eines eigenen Fragebogens in LimeSurvey ausgewählt. Dieser wurde in einer Reihe von Experteninterviews mit Novatec-Mitarbeitern in diversen Softwareentwicklungsprojekten ausgefüllt und anschließend ausgewertet. Für die Auswertung wurden die 24 Key Capabilities nach [5] genutzt. Die geleistete Arbeit lässt sich generell als Hilfsmittel nutzen, um in einem gegebenen Projekt DevOps-verwandte Praktiken zu identifizieren und anhand von Messwerten zu bewerten. Damit wird gleichzeitig eine Art Vorauswahl zur Verfügung gestellt mit der ohne weitere Recherche gearbeitet werden kann. Das Resultat der vorliegenden Arbeit ist ein nutzungsbereiter Setup, der unter anderem die folgenden Artefakte umfasst: Eine Vorauswahl DevOps-assoziierter Praktiken, ein gewichteter Fragenkatalog, ein implementierter Fragebogen in LimeSurvey, eine Vorgehensweise für die Auswertung in Form einer Excel-Tabelle, sowie umfassende Dokumentation. Zudem wurde mit der bisherigen, hier dokumentierten Arbeit, der Grundstein für eine Nutzung als deskriptives Reifegradmodell gelegt.

Literatur

[1] G. Rong, H. Zhang und D. Shao, *CMMI Guided Process Improvement for DevOps Projects: An Exploratory Case Study*, https : / / ieeexplore . ieee . org / document/7831578, Juni 2016.

[2] G. Kim, J. Humble, P. Debois, J. Willis und J. Allspaw, *The DevOps Handbook: How to Create World-Class Agility, Reliability, and Security in Technology Organizations*, , Juni 2017.

[3] N. Forsgren, *2017 State of DevOps Report - Technical Report*, https://puppet.com/resources/whitepaper/2017-state-of-devops-report, Juli 2017.

[4] P. Peschlow, *Die DevOps Bewegung*, https://www.codecentric.de/wissen/publikation/die-devops-bewegung/ Zuletzt aufgerufen am 20.10.2019, 2016.

[5] N. Forsgren, J. Humble und G. Kim, *Accelerate: The Science of Lean Software and DevOps: Building and Scaling High Performing Technology Organizations*, , März 2018.

[6] U. Kuckartz, T. Dresing, S. Rädiker und C. Stefer, *Qualitative Evaluation – Der Einstieg in die Praxis*, , Apr. 2007.

[7] H. Mayer, *Interview und schriftliche Befragung – Grundlagen und Methoden empirischer Sozialforschung*, , Nov. 2012.

[8] J. Pöppelbuß und M. Röglinger, *What makes a useful maturity model? A framework of general design principles for maturity models and its demonstration in business process management*, https://aisel.aisnet.org/ecis2011/28, Juni 2011.

[9] M. Khoshgoftar und O. Osman, *Comparison of maturity models*, https : / / ieeexplore.ieee.org/document/5234402, Sep. 2009.

Application of Artificial Intelligence to optimize forecasting capability in procurement

Daniel Kiefer
ESB Business School, Reutlingen University
Daniel.Kiefer@Reutlingen-University.de

Annette Ulmer
ESB Business School, Reutlingen University
Annette.Ulmer@Reutlingen-University.de

Abstract

The aim of this paper is to show to what extent Artificial Intelligence can be used to optimize forecasting capability in procurement as well as to compare AI with traditional statistic methods. At the same time this article presents the status quo of the research project ANIMATE. The project applies Artificial Intelligence to forecast customer orders in medium-sized companies.

Precise forecasts are essential for companies. For planning, decision making and controlling. Forecasts are applied, e.g. in the areas of supply chain, production or purchasing. Medium-sized companies have major challenges in using suitable methods to improve their forecasting ability.

Companies often use proven methods such as classical statistics as the ARIMA algorithm. However, simple statistics often fail while applied for complex non-linear predictions.

Initial results show that even a simple MLP ANN produces better results than traditional statistic methods. Furthermore, a baseline (Implicit Sales Expectation) of the company was used to compare the performance. This comparison also shows that the proposed AI method is superior.

Until the developed method becomes part of corporate practice, it must be further optimized. The model has difficulties with strong declines, for example due to holidays. The authors are certain that the model can be further improved. For example, through more advanced methods, such as a FilterNet, but also through more data, such as external data on holiday periods.

Key words

Artificial Intelligence, Machine Learning, Deep Learning, Artificial Neural Network, Forecasting, Time series prediction

CR-Category

- **Computing methodologies** → *Artificial Intelligence*
- **Computing methodologies** → *Planning under uncertainty*
- **Applied computing** → *Forecasting*

University supervisor: Prof. Dr. Clemens van Dinther
ESB Business School
Reutlingen University
Clemens.Van_Dinther@
Reutlingen-University.de

Wissenschaftliche Vertiefungskonferenz
21. November 2019, Reutlingen University
Copyright 2019 Daniel Kiefer, Annette Ulmer

1 Introduction

Artificial Intelligence is a branch of computer science that deals with the investigation of mechanisms of intelligent human behavior (intelligence). This is done by simulation using artificial artifacts, usually programs on a computer. The term "artificial intelligence" was invented by the American computer scientist John McCarthy (*1927). He used it in the title of a project proposal for a conference lasting several weeks, which took place in 1956 at Dartmouth College in the USA [1].

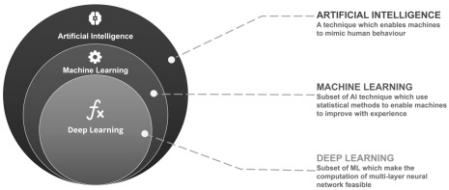


Figure 1: Artificial Intelligence, Machine Learning and Deep Learning [2]

Figure 1 shows the relationship between AI, ML and DL. They are subsections with specific methods and algorithms. AI algorithms are able to act and optimize independently, e.g. the subsection Machine Learning learns through training with huge data sets [3]. Multiple real-life industry and business applications have already shown the potential of this technology.

An important sector is the prediction of future demand not only according to past demand, but also considering factors like economic growth, weather forecasts or social status of the buyers etc. [4, 5]. AI hence can be of positive economic impact for the enterprises that use it. In industry processes for example, predictive maintenance algorithms using Machine Learning (ML) or Deep Learning (DL) methods that analyze data provided by IOT-sensors to predict upcoming defects reliably [6].

A concrete example is the implementation of a forecast model for a spare parts supplier in the agricultural industry. Here, a model was developed which predicts which components will fail or break on a tractor.

For this purpose, internal customer data (tractor model, daily running time, GPS data, etc.) as well as external data (humidity, temperature, season, etc.) are used.

This ensures that the right components are ordered and delivered just in time. As a result, the customer has faster repairs and downtime is reduced. At the same time, the supplier's tied-up capital is reduced and customer satisfaction increases.

The following chapters in this paper are structured as follows: The research project with its goals is explained in the following chapter. Chapter 3 will explain the state of the art technics for demand forecasting in greater detail. Afterwards the Solution Approach with an MLP ANN is presented in chapter 4. Finally, we will explain the next steps in the project: How we plan to increase the accuracy of mid-term forecasting and implement a short-term forecast as well. The last chapter leads to a conclusion and outlook.

2 Description of the research project ANIMATE

The acronym ANIMATE stands for: *„**An**wendung der künstlichen **I**ntelligenz im **M**ittelstand: Data **A**nalytics – **T**echnologie, Kompetenz und **E**rfolgsfaktoren in der betrieblichen Anwendung“.*

The project started in March 2019 and will continue until February of 2021. The research project is intended to answer questions about the practical use of artificial neural networks, the associated problem areas, their requirements and the suitability of their prognosis:

- Which requirements exist for the used data (data availability, dataset size, quality, etc.)?
- How can the structuring and preparation of the data be supported?
- How can ANN tools be designed so that they can be used relatively easily by companies in their own context with as few adjustments as possible?

- Can problem clusters be identified for this?
- Are the forecasts of the ANN tools better suited than the methods used in companies to date and how can these be further optimized by practical experience?

This article intends to give insight into the ongoing research, display the already found results and map the future research focus.

2.1 Digitalization in the "Mittelstand"

The German term *Mittelstand* is not only defined by quantitative measures such as the size of the companies, but also by qualitative aspects such as the unity of ownership and entrepreneurial responsibility [7]. Therefore, enterprises that exceed the usual thresholds (250 employees, a turnover of 50 Mio and/or 43 Mio in balance sheet total) that are set to define SME's by the European Union [8] are still considered part of the *Mittelstand*.

One of the largest studies on the topic of digitization in medium-sized businesses was carried out by Becker et al [9]. They conclude that digitization has an important to very important meaning for almost all of the participating companies. In relation to the industry, the importance of digitization is also important. In contrast to the industries of the participating companies, the companies in the areas of financing, leasing and corporate service providers in particular assign great importance to digitization, followed by companies from the manufacturing sector. A minor importance was put on record by companies in the construction industry [9].

Yet only 5% of businesses in Germany used Artificial Intelligence in 2018, and especially in medium-sized businesses it is rarely used [10]. Based on the conclusion that artificial intelligence is rarely used in SMEs, the project aims at the transfer of AI methods to the medium-sized German companies.

To transfer methods that are already proved in larger enterprises and research facilities, so that they are usable for smaller enterprises, poses distinctive challenges. This is caused by the business structure and culture of SMEs, as they often have only limited resources to invest in digitalization and are more risk averse [11].

2.2 Project-Team

The ANIMATE project is a cooperation between the REIFF Group, the management consultancy 4flow AG and the research group "Digitalization & Management" of the ESB Business School at Reutlingen University.

Its aim is to research possible business application for Artificial Intelligence (AI) in the *Mittelstand* by implementing applications to forecast demand of the goods traded by REIFF. The research focuses not only on the technical feasibility but also on the organizational demands posed by the business structure and culture of the *Mittelstand*.

2.3 REIFF Group

The REIFF group is based in Reutlingen, Germany and is one of the leading retailers for technical products in the Business-to-Business (B2B) sector. It has a product portfolio of more than 140.000 various products purchased from more than 10.000 suppliers and about 10.000 customers. With a work force of about 875 it is not a SME, but as a family owned and lead business it qualifies as *Mittelstand* [12].

With a highly diversified and specialized portfolio, managing demand and supply of products is challenging for the buyers, as they must estimate future demand of a large quantity of products and customers by high uncertainty. This demand is influenced furthermore by outside factors like for example economic growth. Currently the estimated numbers of future demand are determined by the buyers' individual experience.

The Project ANIMATE aims to support those decisions by using state of the art methods to forecast demand more accurately and also integrate external factors like economic climate.

3 State of the art

Forecasting of future demand as exact as possible is of high economic value. Erroneous forecast can have multiple negative effects. If demand is underestimated, outstanding orders cannot be fulfilled and therefor revenue is lost. Overestimating demand can lead to excess stock, fixed capital and storage costs [4]. Both scenarios are hence to be avoided.

The state of the art in science still differs significantly from business practice regarding forecasting methods [4, 13]. While companies mainly use judgmental and statistical methods for forecasting, e.g. Autoregressive integrated moving average (ARIMA). Artificial Neural Networks (ANN's) have been used for these purposes in research around the turn of the millennium[14]. For some time now, AI forecasting methods have been gradually finding their way into the business practices of large companies [15].

The main reasons for entering practice are primarily:

- Big Data – The amount of data is increasing rapidly, at the same time the price of storing enormous amounts of data is decreasing [16].
- AI Democratization – Through AI packages from e.g. TensorFlow, AI methods are distributed more widely, which reduces the complex hurdle of using them [17].
- Low-cost computing power – Due to the continuous improvement of computing power, simple applications can be calculated on almost any computer. Flexible cloud computing services, such as AWS (Amazon Web Services), reduce the high investment requirements of companies in their own data centers [18].

Some studies show that medium-sized companies have recognized digitization as an important trend, but are far away from large corporations in terms of implementation [9].

The question arises why especially medium-sized companies do not yet use state of the art forecasting methods such as Artificial Intelligence. The reasons for this are multi-layered, but amount to fundamental technical deficits in the IT infrastructure, the lack of a digital strategy as well as psychological hurdles of the changes and the shortage of skilled IT workers [19].

For this reason, there is the research project ANIMATE, which aims to support small and medium-sized enterprises in the application of Artificial Intelligence. Therefore, the focus in this second chapter is on the description of Artificial Intelligence as state of the art method for forecasting purposes.

3.1 Description of the currently used methods in the field of forecasting

In this chapter, the current forecasting methods used by companies in practice are briefly examined.

Surveys conducted with companies show that companies today still use judgmental forecasting methods based on experience rather than quantitative methods [20].

Research has shown that quantitative methods to be superior to judgmental methods, both in terms of accuracy and timeliness. Moreover, it has been shown that human biases influence the assessment of forecasts, including lack of consistency, tendency to over forecast and wishful thinking [20].

The following table 1 gives an overview of widely used statistical methods for time series prediction based on linear algebra.

Table 1. Overview of classic statistical methods for forecasting purposes [21]

Method	Example
Autoregressive Integrated Moving Average (ARIMA)	Gold price
Seasonal Autoregressive Integrated Moving-Average (SARIMA)	Electricity Consumption
Vector Autoregression Moving-Average (VARMA)	Temperature
Holt Winter's Exponential Smoothing (HWES)	Visitors in Australia
Generalized Auto Regressive Conditional Heteroscedasticity (GARCH)	Stock Market

The advantage of these methods for companies is based on the fact that these methods are widely used since decades and are better understood by companies than artificial intelligence methods [21]. The classical statistical methods can achieve reliable results with a multitude of possible applications. Since these approaches are based on linear relationships, they are clearly limited and fail in complex applications. Artificial Intelligence methods, such as neural networks, achieve significantly better results because they can recognize complex non-linear cause-effect relationships [4].

The presented methods all had their justification in the past, but due to the permanent advancement of technology, software and research, there are new and better opportunities for companies to conduct forecasting today [17]. Overall it can be stated that the current methods of companies can be improved in practice. In the context of relevance, influence and dependencies of good forecasts, companies should constantly adapt and optimize their methods [4].

Due to the developments explained at the beginning of chapter 3, a shift to AI-based forecasts is possible, even for SMEs. For this reason, the possibility of Artificial Intelligence in forecasting will be discussed in more detail in the following chapter.

3.2 Artificial Intelligence for forecasting purposes

As shown in Figure 1 in Chapter 1 Introduction, the generic term artificial intelligence consists of the further subareas Machine Learning and Deep Learning.

Artificial Intelligence enables machines to learn from their experiences. The machines adapt their response to new inputs and thus perform human-like tasks by processing large amounts of data and recognizing patterns in them. Machine Learning is a subset of artificial intelligence. It enables the machines to learn and make predictions based on their experiences (data). ML uses specific algorithms such as Naive Bayes or Multiple Linear Regression. Deep Learning is a subset of machine learning inspired by the functionality of our brain cells. It simply takes data connections between all artificial neurons and adapts them according to the data pattern. More neurons and layers are needed if the data set is complex and not linear. It automatically provides learning at multiple levels of abstraction, allowing a system to learn complex functional mappings without being dependent on a particular algorithm. Different neural network architectures are used, such as Convolutional Neural Networks for image processing [2, 3, 18, 22].

More than 5,000 academic publications have been indexed in the AI forecasting field by ISI (International Science Indexing), which demonstrates the increased interest of AI in the context of forecasting [5].

According to Zimmermann, ANN's are particularly suitable for solving nonlinear problems with a high number of variables [23]. They extend the instruments of conventional methods of linear algebra, which are suitable for the solution of linear systems of equations

with a multitude of linear variables, and the methods of integral and differential calculus, which are suitable for the solution of nonlinear systems with few variables.

Therefore ANN's are particularly suitable for forecasts of complex, non-linear but also linear circumstances as well as time series [24].

Figure 2 illustrates the above-mentioned facts.

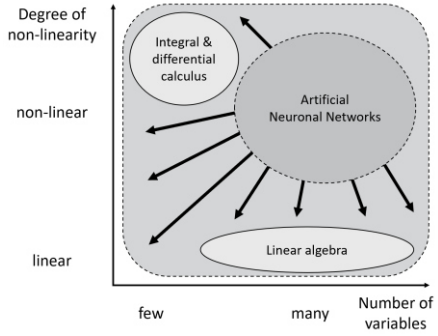


Figure 2: Areas of application of Artificial Neural Networks [4]

Neuronal network-based algorithms have become very popular in recent years and are in the focus of Artificial Intelligence [25]. Artificial Neural Networks (ANN) are systems that consist of many simple units and can process information. They can be programmed individually or an already existing program of an ANN can be used for training and prognosis [22]. Due to the many, often free of charge available programs, which can also be used by people without far-reaching computer science knowledge, the use of the ANN is also practically applicable in enterprises.

Areas of application for Artificial Intelligence include customer loyalty management for forecasting customer departures, credit management for creditworthiness checks, exchange rate forecasts, marketing for sales and demand forecasting, controlling for early warning or in the energy supply sector for load forecasts or electricity price forecasts for day-ahead electricity prices as well as the prediction of the price of CNC turned parts based on the technical drawing [4, 6, 13, 16, 19, 23, 26, 27].

The following figure 3 illustrates an Artificial Neural Network.

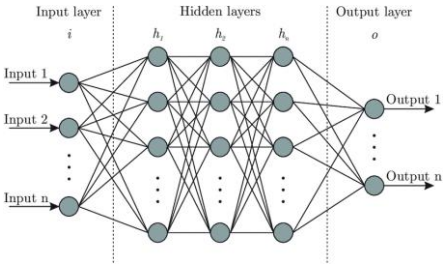


Figure 3: Artificial Neural Network architecture [28]

The network shown consists of the Input Layer, the Hidden Layers and the Output Layer. Since the displayed network has more than one hidden layer, it is a so-called Deep Learning network [29].

The next chapter describes the concrete possibility of Artificial Intelligence to create prognoses in the context of the ANIMATE research project.

4 Solution Approach

The REIFF Group faces numerous challenges. The company serves many customers in different industries with a variety of products. Service is perceived as a decisive competitive feature. For this reason, it is of great interest to REIFF to optimize inventory levels and reduce delivery times for customers.

In the following chapter, the procedure for using an Artificial Neural Network in the ANIMATE research project is explained in more detail in the following structure.

First, the research process is presented. This is followed by the Explanation of the target forecast and the data structure as well as the approach for training the model. This is followed by the model architecture and its parameters. Finally, the forecasting capability is evaluated.

4.1 Research Process

The REIFF group provided the research team with inventory numbers on monthly basis and all incoming orders from 2008 until 2019. After consolidating the data, which was delivered as multiple .csv data files, into an SQL-data base, this data was used to train and test machine learning algorithms, that forecast the demand of product groups and industry sectors.

A baseline was set up to compare the accuracy of the applications forecasting with the status quo. The scope of the forecasting is separated in mid-term and short-term forecasts. Mid-term forecasts study a time range from 3 up to 6 months for different product categories. Short-term forecasts focus on estimating the demand of the individual products on a basis of 1-2 weeks. As a first step, the middle-term forecasts were implemented by training a Multilayer perceptron (MLP) with the data provided by REIFF and several outside factors.

Alongside, we will also research the current business processes of REIFF via a number of qualitative expert interviews to then subsequently conceptualize how an AI application can be integrated in the default business processes.

4.2 Explanation of the target forecast

The present data structure consists of eight divisions of the branches, such as drive technology, hydraulics, sealing technology, plastics, etc.. Sales mainly consists of the number of orders and the articles ordered per month. In the first step, a mid-term forecast of the sales of the internal branches of REIFF is made on a monthly basis. The medium-term forecast was requested by the project leader from the partner company REIFF. This poses the following challenges. Medium to long-term forecasts at product level (n>140.000) are currently difficult to implement, due to the long forecast period and the large number of completely different products and customers.

For this reason, mid to long-term forecasts are realized at the branches level. As soon as the mid-term forecasts on branch level are implemented with a sufficient error rate, we will continue to develop short-term forecasts on product level. Figure 4 illustrates this relationship between short-term, mid-term and long-term forecasts.

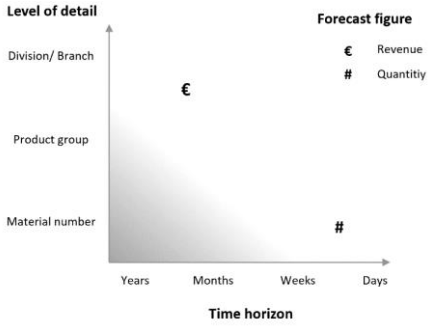


Figure 4: Comparison of the forecast time horizon

The purpose of the graph is to illustrate that a forecast at the lowest detail level with the highest information content can only be made on a very short time horizon. In turn, a forecast at industry level on a significantly longer forecast period is possible

4.3 Structure of the data

For a better understanding, the data structure is now discussed, as this serves as input for the forecast (output). ML and DL algorithms work best if they can use a large set of data that is labeled congruently. Under real-life conditions though, data availability and consistency pose a challenge though, as data sets are often incomplete or repetitive [30].

As already mentioned, the data structure is basically divided into branches such as drive technology, hydraulics, sealing technology, plastics, etc.. The customer order contains information about the order, like the customer, the branches, the ordered article, the quantity, etc. We have access on the stock numbers in a monthly interval and all the customer orders beginning in 2008.

During the project period we will receive regular updates, which means newer data of the future quarters as well.

In addition to the data of REIFF, the following external data has been included: ifo Business Climate Index, DeStatis economic data, Google Trends (using Keywords like “PVC”, “REIFF” or customer names). This data is then combined to a single .csv by a python script. The data is then preprocessed by using the StandardScaler method included machine learning library Sklearn to standardize the input, and then rescaled to values between 0 and 1 by the MinMax-Scaler by Sklearn.

4.4 Training structure

For the training of the neural network the data explained above are used. It should be emphasized that instead of a classical training method, a rolling cross validation training method (Time Series Split) has been used. With a classical training method, the data set is usually divided randomly into a training set (80%) and a test set (20%). The model has access to the full training set from the beginning.

Using the time series split method, the training set (80%) is split once more into training data and validation data. The test set (20%) is used for a subsequent evaluation. The model is trained iteratively, with an increasing amount of data. This has the advantage that the chance of a randomly "simple" data set or division has been reduced. Furthermore, the model becomes less susceptible to the effect of overfitting.

This is a training method which is advantageous for time series prediction because time series modeling and forecasting are tricky and challenging. Cross validation methods must be used to optimize model quality and reduce possible overfitting. Time Series Split is a variation of k-fold which returns first k folds as train set and the (k+1) fold as test set. The idea about the Time Series Split is to divide the training data into 2 parts.

In each iteration the training set is used to train the model and at the same time a part of the training set is used directly for validation, with data that the model has not yet used [31].

As an example, it can be assumed that the model is trained with the data for the months January to March. April, which is predicted, serves as validation. In the next iteration the model is trained with the months January to April and predicts the month May for validation. The following figure 5 illustrates the training method schematically.

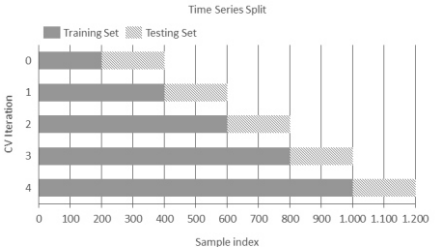


Figure 5: Exemplary representation of the rolling Time Series Split [31]

4.5 Model architecture and parameters

In a first step, a prototype network architecture of the Multi-Layered Percepton (MLP) is tested. The open source library scikit-learn, which is based on the Python programming language, is used to build the neural network. The concrete name of the used function/class is sklearn.neural_network.MLPRegressor.

The model has 3 hidden layers consisting of 500, 100 and 10 neurons each as well as an input and output layer. The activation function rectifier (ReLu) is used. Adam is used as a gradient-based optimization of stochastic objective functions.

The other Parameters are illustrated in the following table 2.

Table 2. Parameters of the MLP ANN

Parameter	Value
Alpha	0.001
Batch_size	Auto
Learning_rate	Invscaling
Power_t	0,5
Max_iter	1000
Shuffle	True
Random_state	1
Verbose	False
Warm_start	False
Momentum	0.9
Nesterovs_momentum	True
Early_stopping	False
Alidation_fraction	0.1
Beta_1	0.9
Beta_2	0.999
Epsilon	$1 * 10^{-8}$

Various hyperparameters were tested on a trial-and-error basis to find the overall best result.

4.6 Outcome and Evaluation

To evaluate the model quality, the neural network is compared with standard forecasting methods as well as a baseline, which represent the performance of REIFF. Since there is currently no explicit sales planning, the warehouse stocks are used to estimate sales expectations. The REIFF baseline is calculated as follows.

$$I = \frac{s\ [€]}{p\ [months]} \qquad (1)$$

with:

I = Implicit sales expectations

s = stock

p = planned delivery time

We use linear regression and SARIMA as classic forecast methods. For the comparison between the methods Root Mean Square Error (RSME) and Coefficient of Determination (R^2) are used.

The following table 3 displays the results of the methods (The row highlighted in gray is the MLP ANN based on Artificial Intelligence).

Table 3: Comparison of different RMSE and R^2 for different models

Model	RSME	R^2
Baseline	99 k€	78%
Lin. Regression	88 k€	87%
SARIMA	137 k€	70%
MLP ANN	46 k€	95%

For a better understanding regarding the prognosis quality, the actual sales and the forecast from the MLP ANN for 2018 are presented graphically in Figure 6.

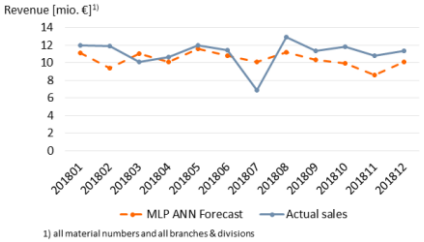


Figure 6: Comparison of the MLP ANN forecast with the actual sales in 2018

The chart shows that the MLP ANN has very similar trends to the current sales trend. However, it can also be noted that the difference of forecast and actual data needs to be improved. Additionally, the July's outlier was not correctly identified by the ANN The outlier itself might be explained by a holiday period.

As a conclusion it can be stated that artificial intelligence has its justification in the field of forecasts.

On the one hand, the MLP ANN was able to achieve better results, i.e. more accurate forecasts, than classical static methods (see Table 3). At the same time, the result is not yet satisfactory. Further research is needed to optimize the result. One is to improve the data basis. Part of this is to increase the amount of data through internal data from REIFF. At the same time, external data such as holiday periods are added. Other deep learning architectures are also being tested, such as FilterNet, a convolutional neural network.

5 Conclusion and outlook

The first implementation of the ANN shows that using an Artificial Intelligence is superior. The ANN performs better than the baseline of REIFF. It also performed better than both the linear regression and the SARIMA algorithm (see Table 3). The current results could already be used by REIFF to estimate which branches will probably increase their order value and can hence be specially targeted by the sales department. Yet, none of the algorithms could exceed the threshold of 40 k€ with the current clustering and the used external input.

The next research step therefor will aim to enhance the accuracy of the mid-term forecast.

5.1 Enhancing the mid-term forecast accuracy

As a first step the quality of the model in respect of the external data will be verified by Principal Component Analysis (PCA), Correlation and Collinearity Analysis. Also, the clustering of the data will be reviewed. The current used clustering methods combined customers by the branches they were assigned to by REIFFs internal guidelines. These guidelines are partially due to historic developments though. Instead of using REIFFs clustering, in a new approach is used to cluster customer groups, as they are clustered by correlation in their time series. This way businesses with similar demand variation are combined in one cluster.

5.2 Short-term forecasts

Reliably forecasting the demand of product groups and/or single products would enable the buyers of REIFF to minimize the costs of wrong order quantities. Short-term forecasting of single products imposes a set of challenges that we aim to investigate. The product portfolio of REIFF with 140.000 various products is not only large but also highly diverse. Computing such numbers of various products with ANNs requires substantial computing power. And as many of the 140.000 various products are sometimes only ordered every few years, the validity and meaningfulness of a demand forecast of such a product has to be called in question. Hence, before implementing any technical solution, there needs to be an evaluation of the products (ABC- and XYZ-Analysis) themselves and an evaluation what kind of output is usable for REIFF. Also, a new baseline must be set up to compare the future results too. After this preliminary work, several types of ML algorithms as well as Neural Networks will be evaluated on the data.

5.3 Integration in business processes

The ANIMATE project not only aims to show the validity of AI algorithms for SME processes but also to integrate them in daily business, as our research goal is to enable REIFF and other SME's to profit from AI technology. As a first step, expert interviews with REIFF buyers will be conducted to understand the status quo: The tools they use, the people involved, the different kind of orders and the corresponding ordering process. The next task is then to conceptualize how to integrate the technology into the daily business. If necessary, current business processes will be modelled with BPMN (Business Process Model Notation). This visualization can then be used to develop process models that integrate AI forecast technology. The needs and worries of both employees and management also has to be taken in consideration and evaluated.

6 List of references

[1] A. Wichert, “Künstliche Intelligenz,” *Spektrum,* https://www.spektrum.de/lexikon/neurowissenschaft/kuenstliche-intelligenz/6810, 2000.

[2] A. Nandan, “AI vs Machine Learning vs Deep Learning,” *edureka!,* https://www.edureka.co/blog/ai-vs-machine-learning-vs-deep-learning/, 2019.

[3] M. Hänlein and A. Kaplan, “A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence,” *California Management Review*, vol. 61, no. 4, pp. 5–14, 2019.

[4] S. F. Crone, *Neuronale Netze zur Prognose und Disposition im Handel*. Wiesbaden: Gabler, 2010.

[5] G. Zhang, B. Eddy Patuwo, and M. Y. Hu, “Forecasting with Artificial Neural Networks,” *International Journal of Forecasting*, vol. 14, no. 1, pp. 35–62, 1998.

[6] W. Zhang, D. Yang, and H. Wang, “Data-Driven Methods for Predictive Maintenance of Industrial Equipment: A Survey,” *IEEE Systems Journal*, vol. 13, no. 3, pp. 2213–2227, 2019.

[7] IfM Bonn, *Mittelstandsdefinition des IfM Bonn.* [Online] Available: https://www.ifm-bonn.org/definitionen/mittelstandsdefinition-des-ifm-bonn/. Accessed on: Oct. 13 2019.

[8] European Union, *User guide to the SME Defintion.* [Online] Available: https://ec.europa.eu/regional_policy/sources/conferences/state-aid/sme/smedefinitionguide_en.pd. Accessed on: Oct. 13 2019.

[9] W. Becker, P. Ulrich, and M. Vogt, *Digitalisierung im Mittelstand*. Bamberg: Otto-Friedrich-Universität, 2013.

[10] Tobias Weber, Irene Bertschek, Jörg Ohnemus, Martin Ebert, *Monitoring Report - Wirtschaft DIGITAL 2018.*

[11] T. Botzkowski, *Digitale Transformation von Geschäftsmodellen im Mittelstand: Theorie, Empirie und Handlungsempfehlungen*. Wiesbaden: Springer, 2018.

[12] Reiff Gruppe, *Zahlen 2019.* [Online] Available: https://www.reiff-gruppe.de/.

[13] H. Min, “Artificial Intelligence in supply chain management: Theory and applications,” *International Journal of Logistics Research and Applications*, vol. 13, no. 1, pp. 13–39, 2010.

[14] S. W. Foo, Ed., *Global engineering management: emerging trends in the Asia Pacific: Proceedings of 1995 IEEE Annual International Engineering Management Conference, 28-30 June, 1995, Singapore*. Piscataway, NJ: IEEE Service Center, 1995.

[15] H. C. Leung, “Neural networks in supply chain management,” in *Global engineering management: emerging trends in the Asia Pacific: Proceedings of 1995 IEEE Annual International Engineering Management Conference, 28-30 June, 1995, Singapore*, Singapore, 1995, pp. 347–352.

[16] E. Brynjolfsson and A. McAfee, *Die Welt von morgen: Wie die nächste industrielle Revolution unserer aller Leben verändern wird*. Kulmbach: Plassen, 2014.

[17] A. D. Fiore, “Analytics: Why AI Will Shift Decision Making from the C-Suite to the Front Line,” *Harvard Business Review*, 2018.

[18] U. Lämmel and H. Cleve, *Künstliche Intelligenz,* 4th ed. München: Carl Hanser, 2012.

[19] M. Saam, S. Viete, and S. Schiel, “Digitalisierung im Mittelstand: Status Quo, aktuelle Entwicklungen und Herausforderungen: Forschungsprojekt im Auftrag der KfW Bankengruppe,” *Zentrum für Europäische Wirtschaftsforschung*, 2016.

[20] N. R. Sanders, “Corporate Forecasting Practices in the Manufacturing

Industry," *Production and Inventory Management Journal*, vol. 33, no. 3, 1992.
[21] R. Majid and S. A. Mir, "Advances in Statistical Forecasting Methods: An Overview," *Economic Affairs*, vol. 63, no. 4, pp. 815–831, 2018.
[22] A. Zell, "Einführung in Künstliche Neuronale Netze," in *Computational Intelligence, Betriebswirtschaftliche Anwendungen des Soft Computing: Neuronale Netze, Fuzzy-Systeme und Evolutionäre Algorithmen*, H. A. Biethahn J., Kuhl J., Leisewitz MC., Nissen V., and Tietze M., Eds., Wiesbaden: Vieweg+Teubner, 1998.
[23] H.-G. Zimmermann, C. Tietz, R. Grothmann, and T. Runkler, "Recurrent Neural Networks for Industrial Procurement Decisions," *KI - Künstliche Intelligenz*, vol. 26, no. 4, pp. 403–406, 2012.
[24] Z. Tang, C. de Almeida, and P. A. Fishwick, "Time series forecasting using neural networks vs. Box- Jenkins methodology," *SIMULATION*, vol. 57, no. 5, pp. 303–310, 1991.
[25] K. Ramasubramanian and A. Singh, *Machine learning using R: With time series and industry-based uses in R,* 2nd ed. New York: Apress, 2019.
[26] R. J. Kuo, Y. S. Tseng, and Z.-Y. Chen, "Integration of fuzzy neural network and artificial immune system-based back-propagation neural network for sales forecasting using qualitative and quantitative data," *Journal of Intelligent Manufacturing*, vol. 27, no. 6, pp. 1191–1207, 2016.
[27] A. Eberl *et al.,* "Methoden zur Sofortpreis-kalkulation von CNC-Drehteilen," *ZWF*, vol. 113, no. 12, pp. 835–839, 2018.
[28] F. Bre, J. M. Gimenez, and V. D. Fachinotti, "Prediction of wind pressure coefficients on building surfaces using artificial neural networks," *Energy and Buildings*, vol. 158, pp. 1429–1441, 2018.
[29] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural networks*, vol. 61, pp. 85–117, 2015.
[30] X. Li, S. Deng, S. Wang, Z. Lv, and L. Wu, "Review of Small Data Learning Methods," in *2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC)*, Tokyo, Japan, Jul. 2018 - Jul. 2018, pp. 106–109.
[31] D. Cournapeau, "Time Series Split," *scikit-learn.org,* https://scikit-learn.org/stable/auto_examples/model_selection/plot_cv_indices.html#sphx-glr-auto-examples-model-selection-plot-cv-indices-py, 2019.

Advanced Analytics: Analyse und Ablauf der Pre-Implementation

Marek Mauser
Hochschule Reutlingen
Marek.Mauser@Student.Reutlingen-University.de

Abstract

[Context/motivation] Daten sind im heutigen Zeitalter so wichtig wie nie zuvor. Daher ist eine innovative Verarbeitung und Verwendung von Daten innerhalb eines Unternehmens nicht mehr wegzudenken. Bei Advanced Analytics handelt es sich um zukunftsorientierte Analyseverfahren für Daten, welche Unternehmen neue Möglichkeiten im Bezug auf Datenanalyse bieten.
[Question/problem] Advanced Analytics ist heutzutage vielen Unternehmen ein Begriff. Eine Studie belegt, dass Advanced Analytics für die Mehrheit von Unternehmen erwünscht ist. Trotzdem hat Advanced Analytics einen Verbreitungsgrad von nur 15 %. Ursachen hierfür sind unter anderem unzureichende Ressourcen, kein qualifiziertes Personal und zu großen Respekt davor, etwas verändern zu wollen - wenngleich es sich für Unternehmen - richtig angewendet - durchaus lohnen würde. Wie können sich Unternehmen also sicher sein, ob sie bereit für Advanced Analytics sind?
[Principal ideas/results] Es soll ein Einblick in die organisatorischen und technischen Faktoren zur Einführung von Advanced Analytics in ein Unternehmen gegeben werden. Dazu gehören Attribute wie z. B. die Architektur, die Umstellung des Unternehmens und die gewählten Strategie. Zusätzlich werden zielführende Methoden von Advanced Analytics analysiert und erläutert, um einen Überblick über sämtliche existierende Methoden zu bekommen.
[Contribution]Diese Ausarbeitung beschäftigt sich mit der strategischen Einführung von Advanced Analytics in ein Unternehmen. Ziel der Arbeit ist es, ein Konzept für Unternehmen anzubieten, mit dem sie die Einführung von Advanced Analytics erfolgreich organisieren können.

Keywords

Business Analytics, Advanced Analytics, Predictive Analytics, Prescriptive Analytics, Methoden, Einführung, Unternehmen

1 Einleitung

Daten sind das neue Öl; und das nicht nur für die weltweit größten Konzerne wie Google, Amazon oder Apple. Auch kleinere Unternehmen können sich dank intensiver Datenanalyse Vorteile gegenüber Konkurrenten verschaffen. Business Intelligence (BI) wird heutzutage in vielen Unternehmen erfolgreich eingesetzt. Hierbei geht es in erster Linie um die Verarbeitung von gesammelten Daten. Der Bereich bezieht sich dabei allerdings nur auf vergangene oder aktuelle Geschehnisse mit Fragen wie: Was ist geschehen? Warum ist es geschehen? Was geschieht gerade? Advanced Analytics priorisiert, im Gegensatz zu BI, die Analyse zukünftiger Ereignisse. Man spricht auch von Predictive Analytics und Prescriptive Analytics – also der Prognose und Optimierung von Ereignissen. Dieses Datenanalyseverfahren

Betreuer Hochschule: Prof. Dr. Christian Kücherer
Hochschule Reutlingen
Christian.Kuecherer@Reutlingen-University.de

Wissenschaftliche Vertiefungskonferenz
27. November 2019, Hochschule Reutlingen
Copyright 2019 Marek Mauser

ist sehr mächtig und erbringt Unternehmen Chancen, beispielsweise durch die Analyse von zukunftsorientierten Fragen wie: Was wird passieren? Wie können wir etwas passieren lassen?. In dieser Ausarbeitung werden folgende Forschungsfragen beantwortet:

Research Question 1:
Welche organisatorischen und technischen Faktoren müssen vor der Implementierung von Advanced Analytics bedacht werden?

Research Question 2:
Welche zielführenden Methoden stehen Unternehmen mit Advanced Analytics zur Verfügung?

Die Klärung der Fragen ist notwendig, um Probleme vorzeitig zu lösen und damit eine reibungslose Einführung gewährleistet zu können. **RQ1** fragt nach den organisatorischen und technischen Aspekten, die vor einer Einführung von Advanced Analytics berücksichtigt werden müssen. Mithilfe von **RQ2** werden Methoden vorgestellt, mit denen Unternehmen ihre Ziele mit Advanced Analytics realisieren können. Zu jeder existierenden Technik wird daher eine Methode als Beispiel beschrieben.

2 Literaturanalyse

Die Methodik der vorliegenden Ausarbeitung folgt dem Ansatz einer systematischen Literaturrecherche. Im Verlauf dieses Abschnitts wird dabei näher auf den Ablauf dieser eingegangen. Für die Recherche wurden im ersten Schritt die Suchterme erstellt. Anhand dieser Suchterme wurde in den Fachdatenbanken IEEE, ACM, ScienceDirect, DBLP, Springer, Tylor & Francis und dem IBS BW nach Ausarbeitungen gesucht. Es wurden ebenfalls Artikel durch Snowballing (Durchsuchen der Referenzlisten) hinzugezogen. Folgende Suchterme und deren Kombination untereinander wurden für die Literaturrecherche verwendet: organizational, usage, advanced analytics, analytics, introduction, enterprises, case study, requirements, data analytics, risks, launch, technical, development, advantages, big data, methods, predictive, prescriptive und preparation. Nach dem Abschluss der erster Recherchephase, wurden in Phase 2 die Arbeiten und Bücher durch Lesen der Abstracts oder Inhaltsverzeichnisse grob gefiltert. Die dritte und letzte Phase der Literaturanalyse bezog sich auf das Lesen wichtiger Abschnitte und Kapitel der Ausarbeitungen. So entstand letztendlich eine Auswahl von 17 wissenschaftlichen Artikeln und 5 Fachbüchern, die in dieser Ausarbeitung verwendet und referenziert werden. Zusätzlich wird in der Ausarbeitung graue Literatur verwendet, welche als Fußnoten gekennzeichnet sind.

3 Grundlagen

Im folgenden Kapitel werden die Gemeinsamkeiten und Unterschiede der verschiedenen Analyse-Kategorien erläutert. Es geht in erster Linie darum, zwischen Business Analytics, BI und Advanced Analytics zu unterscheiden.

3.1 Advanced Analytics, Business Intelligence und Business Analytics

In der heutigen Zeit finden sich viele verschiedene Begriffe in der Datenanalyse wieder. Um ein klares Verständnis davon zu bekommen, werden die wichtigsten Begriffe vorgestellt. Business Analytics stellt den Oberbegriff für alle Unterkategorien zur Datenanalyse dar. Während man mit dem Begriff Business Intelligence und derer Herangehensweisen mit Descriptive, Diagnostic und Real-time Analytics primär eine Erhöhung der Effizienz von Geschäftsprozessen anstrebt, zielt der Einsatz von Advanced Analytics darauf ab, zukünftige Trends zu erkennen und das Unternehmen

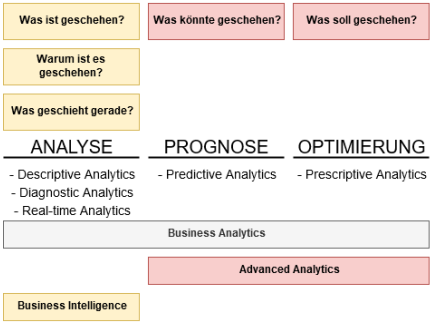


Abbildung 1: Analytics-Begriffe und deren Verfahren (eigene Abb. in Anlehnung an [1])

auf bevorstehende Herausforderungen vorzubereiten [1]. Advanced Analytics lässt sich in Abbildung 1 zu den Rubriken Prognose und Optimierung zuordnen. Zu den Teilgebieten von Advanced Analytics gehören unter anderem Predictive und Prescriptive Analytics. Advanced Analytics hilft Unternehmen dabei, sich nicht nur mit den klassischen Business Intelligence Methoden auseinanderzusetzen, sondern auch präzise Zukunfts- und Optimierungsherausforderungen mithilfe von Predictive und Prescriptive Analytics lösen zu können.

3.2 Predictive Analytics

Predictive Analytics berücksichtigt kürzlich vergangene Daten, um zukünftige Trends, Vorurteile, Tendenzen, Verhaltensweisen usw. durch Ursachen und Korrelationen vorherzusagen [2]. Das Hauptziel ist dabei immer dasselbe: Es geht um den Prozess zur Entwicklung eines Werkzeuges oder Modells, durch das genaue Vorhersagen generiert werden können [3]. Die Methoden sollen neues Wissen generieren. Unter Wissen versteht man interessante Muster in den Daten die allgemein gültig sind, nicht trivial, neu, nützlich und verständlich [4]. Abbildung 2 zeigt die Techniken mit ihren untergeordneten Methoden für Predictive Analytics, auf welche in Kapitel 5 weiter eingegangen wird.

3.3 Prescriptive Analytics

Nachdem sich das vorhergehende Kapitel mit der Klärung von Zukunftsfragen beschäftigt hat, befasst sich folgendes Kapitel mit Prescriptive Analytics und damit einhergehend mit der Optimierung von Ergebnissen. Der Begriff stammt aus [2], und wird als eine Reihe von mathematischen Techniken beschrieben, mit denen sich viele alternative Maßnahmen und Entscheidungen anhand komplexer Ziele und Anforderungen rechnerisch bestimmen lassen. Im Vergleich zu Predictive Analytics ist Prescriptive Analytics weitaus weniger ausgereift [3]. Aktuell zieht Prescriptive Analytics jedoch zunehmend Forschungsinteresse auf sich und wird inzwischen als nächster Schritt zur Steigerung der Datenanalyse und zur Optimierung der Entscheidungsfindung für die Verbesserung der Geschäftsleistung angesehen [5] [4]. Diese Art von Analyse ist die fortschrittlichste Art der Geschäftsanalyse und kann Unternehmen den größten Nutzen bringen [6]. In Abbildung 3 werden die Methoden für Prescriptive Analytics dargestellt.

3.4 Gründe für den Einsatz von Advanced Analytics im Unternehmen

Ein großer Vorteil von Advanced Analytics ist die enorm hohe Flexibilität. Die Methoden von Advanced Analytics können in vielen verschiedenen Bereichen eines Unternehmens

[1] R. Gleich, K. Grönke, M. Kirchmannund, J. Leyk, Strategische Unternehmensführung mit Advanced Analytics, Ser. Haufe Fachbuch. München, 2017, ISBN: 364810422

[2] I. Lustig, B. Dietrich, C. Johnson und C. Dziekan, "The analytics journey", 2010, http://www.analytics-magazine.org, zuletzt zugegriffen am 20.10.2019

[3] Gartner und J. Hagerty, "Planning Guide for Data and Analytics", 2016, https://www.gartner.com, zuletzt zugegriffen am 20.10.2019

[4] D. den Hertog und K. Postek, "Bridging the gap between predictive and prescriptive analytic", 2016, http://www.optimization-online.org, zuletzt zugegriffen am 20.10.2019

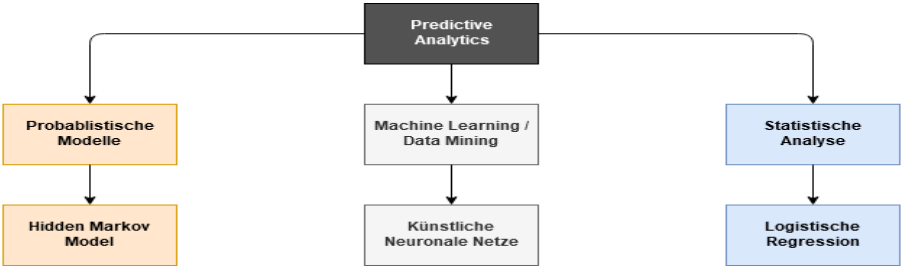


Abbildung 2: Einteilung der Methoden für Predictive Analytics, angelehnt an [1]

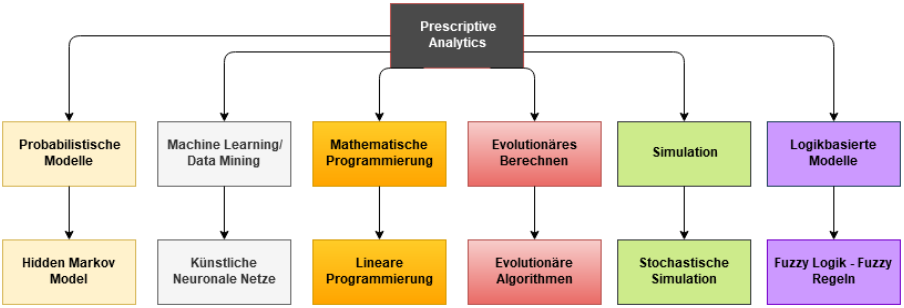


Abbildung 3: Einteilung der Methoden für Prescriptive Analytics, angelehnt an [1]

effektiv und effizient eingesetzt werden. Abteilungen wie Human Resources, Marketing oder das Controlling können die verschiedenen Methoden erfolgsbringend anwenden. Beispiele dafür werden in Kapitel 5 dargestellt. Controllern wird bei der Einführung von Advanced Analytics eine Schlüsselrolle zukommen, deren Kernaufgabe in der Koordination der Informationsversorgung mit dem Informationsbedarf einerseits und dem Planungs- und Kontrollsystem andererseits liegt [7]. Betrachtet man nun beispielsweise den Controlling-Prozess zur Berechnung periodischer Vorausrechnungen, so können mithilfe von Advanced Analytics Schwachstellen im zeitlichen und ressourcenbezogenen Aufwand sowie der Präzision minimiert werden [8].

4 Organisatorische und technische Faktoren

Vor der Einführung von Advanced Analytics in ein Unternehmen, müssen im Vorfeld organisatorische und technische Aspekte geklärt sein. Denn nicht jedes Unternehmen sollte Advanced Analytics anwenden. Das liegt vor allem an den benötigten Ressourcen zur Umsetzung. Deshalb werden in diesem Kapitel die wichtigsten Faktoren erläutert, die bereits vor der Einführung von Advanced Analytics klar definiert sein müssen.

4.1 Organisatorische Faktoren

4.1.1 Strategie und Use Cases Die erfolgreiche Einführung von Advanced Analytics in ein Unternehmen setzt eine gut durchdachte Strategie voraus. Daher sollten sich die Verantwortlichen in erster Linie mit der Unternehmensstrategie auseinandersetzen und

analysieren, in welchen Bereichen Advanced Analytics eingesetzt werden soll. Konzipierte Use Cases helfen bei der Entscheidungsfindung. Es gilt zu beachten, dass es nicht damit getan ist, einen Data Scientist einzustellen und die aktuellste Advanced Analytics Software zu erwerben. Ein wichtiger Aspekt bei der Einführung von Advanced Analytics in das Unternehmen ist die Auswahl der Durchführung. Eine Version stellt das sogenannte High Impact Modell dar. Dabei geht es darum, dass die Einführung für das gesamte Unternehmen substanziell, einflussreich und entscheidend sein sollte [9]. Die Herausforderung sollte groß, das Modell aufschlussreich und die Geschäftsvision vollständig sein [5]. Advanced Analytics wird mit diesem Modell also bereits bei der Planung so konzipiert, dass es dem Unternehmen voll und ganz bei der effektiveren und genaueren Analyse von Daten zur Verfügung steht.

Das Low Cost Modell fokussiert sich stattdessen auf einer dem High Impact Modell entgegengesetzten Strategie. Hier geht es in erster Linie darum, sich mit Advanced Analytics vertraut zu machen und vorerst keine Ziele oder hohe Anforderungen an Advanced Analytics zu stellen. Unternehmen sollen einige relativ einfache Analysen definieren, für deren Ausführung nicht viel Zeit oder Daten notwendig sind [6]. Es geht darum, Advanced Analytics kennenzulernen und durch Beispiele oder Tests geeignete Anforderungen und Use Cases erstellen zu können. Anhand einiger intuitiver Beispiele kann das Unternehmen sehen, was die Daten leisten können und gewährleisten, dass dieser Ansatz Ergebnisse liefert, die leicht zu testen sind, um festzustellen, welche Art von Auftrieb die Analyse bietet [6]. Dieses Modell benötigt zwar längere Zeit um für das Unternehmen einen signifikanten Mehrwert zu erbringen, jedoch birgt es auch weniger Risiken.

Die Untersuchung und daraus resultierende Arbeit von [9] stellt ein geeignetes Prozess Modell vor, welches durch eine qualitative Untersuchung von fünf Unternehmen, die eine Einführung in Advancend Analytics vorgenommen haben, erarbeitet wurde. Das in Abbildung 4 dargestellte Prozess-Modell dient dem Unternehmen dazu, alle relevanten Phasen und deren dazugehörigen Leistungen, die für die Einführung von Advanced Analytics notwendig sind, zu erbringen. Das Modell startet mit der Kickoff Phase und soll Antworten bezüglich des zuständigen Teams und der Verantwortlichkeiten innerhalb des Teams festlegen. In Phase 2 werden anschließend die Verhandlung von Geschäftsproblemen geregelt. Dazu gehören die Definition der Geschäftsprobleme, wie diese Probleme momentan geregelt werden und die Anforderungen an die Daten. Die Datenvorbereitung in Phase 3 soll Fragen bezüglich Datenquellen, Definition der Daten, Datenverfügbarkeit und

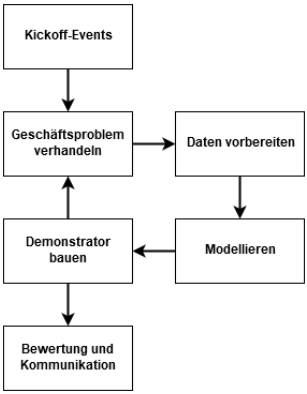


Abbildung 4: Vorgeschlagenes Prozess-Modell nach [9]

[5] S. LaValla, E. Lesser, R. Shockley, M. S. Hopkins und N. Kruschwitz, "Big Data, Analytics and the Path From Insights to Value", http://foresight.ifmo.ru, zuletzt zugegriffen am 20.10.2019

[6] B. Franks, "To Succeed with Big Data, Start Small", 2012, https://hbr.org/2012/10/to-succeed-with-big-data-start, zuletzt zugegriffen am 20.10.2019

Datenbereinigung klären. Die Modellieren-Phase beschreibt die vorgehende algorithmische Strategie mithilfe von Methoden. In Phase 5 wird die gebaute Umgebung in das System des Unternehmens eingepflegt und liefert erste Ergebnisse. Die letzte Phase dient der Evaluation des gesamten Prozesses. Hier werden die Ergebnisse visualisiert und mit den Ausgangswerten vergleichen. Zudem werden die Auswirkungen der Ergebnisse analysiert und stetig weitere Verbesserungen kommuniziert und ausgetauscht.

4.1.2 Veränderte Organisationsformen In der heutigen Zeit ist das wirtschaftliche Umfeld geprägt durch Personalabbau, Fusionen, Übernahmen und erhöhte Unsicherheit, wodurch sich der Druck auf die Unternehmen enorm erhöht hat [10]. Eine gute Möglichkeit diesen Risiken bzw. Herausforderungen zu trotzen, sind Shared Services. Bei einer Shared Service Organisation (SSO) besteht das Hauptziel darin, gleichartige Prozesse aus unterschiedlichen Abteilungen eines Unternehmens zusammenzufassen und von einer dafür bereitgestellten Abteilung erbringen zu lassen. Besonders in Verbindung mit Advanced Analytics ist eine veränderte Organisationsform empfehlenswert. Grund hierfür ist in erster Linie, dass mehrere Fachabteilungen für Advanced Analytics zusammenkommen müssen.

4.1.3 Neuordnung der Rollen Neue Arbeitsbereiche erfordern auch neue Rollen innerhalb der Organisation. Damit Advanced Analytics funktioniert, benötigt man vor allem geeignetes Personal. Ein **Chief Digital Officer (CDO)**, der alle digitalen Themen in einem Unternehmen vorantreibt und für die Zusammenarbeit und Koordination zwischen unterschiedlichen Bereichen im Unternehmen verantwortlich ist, ist aufgrund seiner Qualifikationen besonders wichtig. Ebenfalls benötigt werden sogenannte **Data Scientists**, die im Stellenmarkt sehr selten zu finden sind. Ein **Data Scientist** kennt sich hervorragend mit statistischen Verfahren und Algorithmen aus. Der moderne **Data Scientist** muss sich im Gegenteil zu anderen **Data Scientists**, welche bereits Erfahrungen in der Verwendung von BI haben, auch sehr gut mit mathematischen Kenntnissen und Machine Learning Prozessen auskennen. Ein **Data Engineer** sorgt für die Zusammenführung der Daten. Er kümmert sich z. B. darum, dass die Daten problemlos von der Maschine bis hin zur Analyse-Software gelangen. Ebenfalls wird ein **Data Steward** benötigt, welcher stets die Qualität und Integrität der Daten überwacht. Ebenfalls abgelöst bzw. erweitert wird das **Business Intelligence Competency Center (BICC)**. Ob am Ende der Planungsphase für die Einführung von Advanced Analytics ein erweitertes **BICC** oder aber vollkommen neue Konzepte wie etwa ein **Analytics Competence Center (ACC)** stehen wird, kann aus heutiger Sicht nicht eindeutig geklärt werden [11]. Für Unternehmen ist es dennoch schwer, an qualifizierte Personen zu kommen, die diese Stellen besetzen können. Aus diesem Grund greifen immer mehr Unternehmen auf externe Mitarbeiter zurück. Unternehmen, welche Advanced Analytics einführen möchten, sollten sich aber auch auf ihre BI-Erfahrungen besinnen, um auf diese Weise die Einführung und Nutzung effizient zu gestalten [11].

4.1.4 Veränderung der Anforderungen Aufgrund der Digitalisierung steht auch das Controlling vor neuen Herausforderungen. Die Digitalisierung des Controllings entwickelt sich nur sehr langsam und allmählich. Big Data, In-Memory-Technologien und Advanced Analytics verändern das Controlling nicht immens; vielmehr ist es ein Prozess, der mit einzelnen Use Cases beginnt, schrittweise ausgebaut wird und damit zu einer nachhaltigen und umfassenden Steigerung von Effizienz und Effektivität im Controlling führt [7].

[7] R. Gleich, K. Grönke, M. Kirchmannund, J. Leyk, Strategische Unternehmensführung mit Advanced Analytics, Ser. Haufe Fachbuch. München, 2017, ISBN: 364810422

Durch die Einführung von Advanced Analytics können innerhalb eines Unternehmens völlig neue Anforderungen gestellt werden, welche bislang nicht möglich waren. Zudem werden die geschäftlichen und technischen Anforderungen in [8] als anders beschrieben. Damit ist gemeint, dass Anforderungen durch Advanced Analytics weitreichender getroffen werden können, was so noch nicht möglich war.

4.2 Technische Faktoren

Um die Vorteile von Analytics für die Erzielung optimaler und automatisierter Entscheidungen in nahezu Echtzeit zu nutzen, wird von [12] eine Verbundarchitektur vorgeschlagen, die Descriptive, Predictive und Prescriptive Analytics sowie verschiedene Datengeneratormodelle und ein Data Warehouse umfasst. Die Abbildung 5 zeigt diese Verbundarchitektur, dargestellt als Verhalten des Ablaufs einer Analyse. Es wird eine Umgebung durch ein Prescriptive Modul, welches von einem Predictive Modul mit zusammentragenden Funktionen betrachtet wird, beschrieben [12]. Zudem wird jedes Predictive Modul als Descriptive Modul mit zugeordneten Fähigkeiten betrachtet. Im folgenden werden die Abbildungen beschrieben.

Data and Business Generator Models: sind für die Bereitstellung verschiedener Arten von Daten und domänenspezifischer Informationen verantwortlich.

Descriptive Analytics Module: ist die Datenverdichtungseinheit, welche gesammelte Daten durch Datenbereinigungseinheiten vorbereitet und anschließend zu bestimmten Formen zusammenfasst, die für andere Analysemodule verständlich sind.

Predictive Analytics Module: dabei handelt es sich um die Prognoseeinheit. Hier erstellt das Modul ein Vorhersagemodell, welches vom beschriebenen Modul bereitgestellt

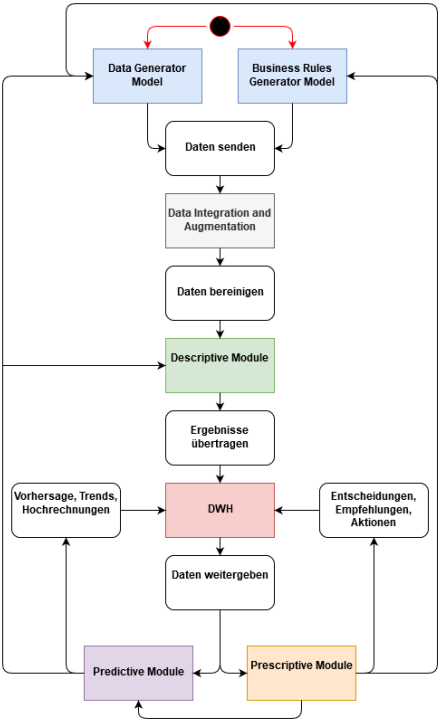


Abbildung 5: Aktivitätsdiagramm einer Architektur für Advanced Analytics, angelehnt an [12]

wird. Um die Genauigkeit des Vorhersagemodells zu verbessern, sendet dieses Modul eine ordnungsgemäße Rückmeldung zu den Daten und dem Descriptive Analytics Modul und fordert Änderungen der generierten Daten an.

Prescriptive Analytics Module: bei diesem Modul handelt es sich um die Empfehlungseinheit und beinhaltet Einheiten für Simulation, Optimierung, Bewertung und Rückmeldung. Um die Übereinstimmung der Vorschriften mit den Systemzielen zu überprüfen, sendet dieses Modul eine Rückmeldung an die Data and Business Generator Models und ebenfalls an das Predictive Analytics Module. Die Kernelemente eines

[8] P. Russom, "Big Data Analytics", 2011, https://tdwi.org, zuletzt zugegriffen am 20.10.2019

jeden Prescriptive Analytics Modul sind Simulation, Optimierung sowie Auswertung.
Data Warehouse: dient als zentrale Datenspeichereinheit.
Die Architektur aus [12] sieht vor, zunächst die Daten aus den Data and Business Generator Models zu sammeln und in das Descriptive Module zu übertragen. Das Data Integration and Augmentation Module befasst sich mit der Vereinheitlichung der Datenströme und bereinigt die gesammelten Daten. Das Descriptive Module speichert dann die zusammengefassten Daten. Als nächstes fragt das Predictive Module die aktualisierten zusammengefassten Daten aus dem Speicher als Eingabe ab und erstellt ein genaues Vorhersagemodell auf der Grundlage der historischen Daten. Ausgehend von den historischen Daten mit ihren Wahrscheinlichkeitswerten können Unternehmen anhand der Vorhersagen und zukünftigen Trends ihre Zukunftschancen und wahrscheinlichen Risiken unterscheiden. Abschließend übernimmt das Prescriptive Module die aktualisierten Vorhersagen sowie die Geschäftsregeln und generiert mehrere Szenarien durch die Simulationseinheit. Durch eine Liste möglicher umsetzbarer Empfehlungen, lässt sich für das Unternehmen die Frage, was zu tun ist, beantworten. Die Optimierungseinheit validiert anschließend simulierte Szenarien zusammen. Das am besten geeignete und optimierteste Szenario wird ausgewählt und zur endgültigen Validierung an die Evaluierungseinheit weitergegeben. Die Optimierungseinheit beantworten abschließend die Frage, warum das Unternehmen das machen sollte. Ergebnisformen können Ja/Nein-Vorschläge sein, aber auch Werte oder Vektormatrizen mit empfohlener Preise und Mengen.

5 Zielführende Methoden von Advanced Analytics für Unternehmen

Das folgende Kapitel beschäftigt sich mit den Techniken von Advanced Analytics und den daraus folgenden anwendbaren Methoden. In Kapitel 3.2 und 3.3 wurde bereits kurz auf die vorhandenen Methoden für Predictive und Prescriptive Analytics eingegangen, ohne sie jedoch näher zu betrachten. Für das vorgeschlagene Prozess-Modell in Abbildung 4 muss eine algorithmische Strategie entschieden werden. Die wichtigsten und relevantesten Methoden zeigen die Abbildung 2 und Abbildung 3. Sie zeigen die Techniken Probabilistische Modelle, Machine Learning / Data Mining, Statistische Analyse, Mathematische Programmierung, Evolutionäres Berechnen, Simulation und Logikbasierte Modelle mit jeweils einer Beispielmethode.
Die verfügbaren Techniken und die daraus resultierenden Methoden unterscheiden sich für Predictive und Prescriptive Analytics. Das liegt vor allem daran, dass man mit diesen zwei Analyseverfahren unterschiedliche Probleme angeht. Welche Methode für ein Unternehmen geeignet ist, lässt sich so nicht sagen. Jeder Anwendungsfall benötigt andere Methoden und da Advanced Analytics in so vielen Bereichen einer Organisation Anwendung finden kann, lässt sich diese Frage nach der Einen passenden Methode nicht erläutern. Unternehmen müssen sich die Frage stellen, was genau sie eigentlich analysieren möchten. Folgende Beispiele von einsetzbaren Methoden geben einen Einblick und erläutern, warum es so schwer ist diese Frage allgemein zu beantworten.

5.1 Probabilistische Modelle

In der Predictive und Prescriptive Analytics werden probabilistische Modelle verwendet, um die Wahrscheinlichkeit des Auftretens bestimmter Ereignisse zu berechnen [1]. Dafür kommt beispielsweise die Stochastik, was die Wahrscheinlichkeitstheorie und die Statistik

zusammenführt, zum Einsatz. Das Hidden Markov Model ist ein stochastisches Modell, welches unter anderem im Finanzbereich eines Unternehmens eingesetzt werden kann. In einem Artikel von [13] wird eine Hybridmethode vorgestellt, die Text Mining und das Hidden Markov Model zur Vorhersage der finanziellen Notlage für börsennotierte Unternehmen verwendet.

5.2 Machine learning/Data Mining

Die Verwendung von künstlichen neuronalen Netzen z. B. für das Einschätzen des Kreditrisikos ist aufgrund der stetig steigenden Datenmengen für Unternehmen problemlos anwendbar [14]. Mithilfe solcher Netze lassen sich Zukunftsprognosen und Trends besser bestimmen. Auch im Bereich Optimierung und Simulation finden künstliche neuronale Netze Anwendung. In Artikel [15] wird beispielsweise die kollaborative Supply Chain mithilfe eines künstlichen neuronalen Netzwerk Ansatzes geplant.

5.3 Statistische Analyse

Die statistische Analyse ist ein Teilgebiet der Mathematik, das sich mit Datensammlung, Organisation, Analyse, Interpretation und Präsentation befasst [16]. Zu dieser Technik gehören beispielsweise viele Methoden der Regression an. In der Ausarbeitung [17] wird ein logistisches Regressionsmodell entwickelt, um Unternehmen auf der Grundlage von Pünktlichkeits- und Budgetmaßnahmen zu klassifizieren.

5.4 Mathematische Programmierung

Die mathematische Programmierung im Kontext der Prescriptive Analytics wird sehr oft verwendet. Insbesondere befasst sich die überwiegende Mehrheit der von [1] überprüften Arbeiten mit Optimierungsmethoden und -algorithmen. Die lineare Programmierung und ihre Erweiterungen sind wohl die am häufigsten verwendeten Optimierungsmethoden in der Prescriptive Analytics [1]. Die lineare Optimierung ist eine Technik zur Optimierung einer linearen Zielfunktion unter Berücksichtigung linearer Gleichheits- und Ungleichheitsbedingungen [18]. Einsatzgebiete sind zum Beispiel die Produktion und die Logistik eines Unternehmens.

5.5 Evolutionäres Berechnen

Für Probleme die zu kompliziert sind, um mit mathematischer Programmierung eine optimale Lösung zu finden, können evolutionäre Berechnungsmethoden verwendet werden, um Lösungen effizient bereitzustellen [19]. Evolutionäre Berechnungsmethoden haben sich als Optimierungswerkzeuge etabliert [1]. Evolutionäre Algorithmen sind ein Bestandteil der evolutionären Berechnungsmethoden. Eine Ausführlich gelistete Tabelle in [20] zeigt verschiedene Einsatzmöglichkeiten von evolutionären Algorithmen innerhalb eines Unternehmens. Einsatzgebiete sind beispielsweise Finanzdienstleistungen, die Energiewirtschaft und das Personalwesen.

5.6 Simulation

Das zukunftsorientierte Simulieren von gewünschten Szenarios kann für Unternehmen eine wichtige Technik darstellen. Grund hierfür sind die vielen Anwendungsgebiete wie die Produktion, die Logistik, Investitionen oder die Supply Chain des Unternehmens. Exemplarisch wurde eine stochastische Simulation genutzt, um den Nutzen des Produktportfolios eines Unternehmens im Einklang mit den Unternehmenszielen zu maximieren [21].

5.7 Logikbasierte Modelle

Logikbasierte Modelle sind hypothetische Beschreibungen der Kette von Ursachen und Wirkungen, die zu einem Ergebnis von Interesse führen und im Kontext der Informationssysteme regelbasierte Systeme, die Darstellung von Expertenwissen und die Ermittlung von Domänenwissen umfassen

können [1]. Ein Beispiel sind Fuzzy Regeln innerhalb eines Fuzzylogik Systems. In einem Artikel von [22] werden die Vorteile der Verwendung von Fuzzy-Logik im Handel aufgezeigt, denn die menschliche Beurteilung und Bewertung folgt weder der Booleschen Logik noch einer anderen herkömmlichen mathematischen Disziplin [22].

Die exemplarisch beschriebenen Methoden zeigen, wie unterschiedlich Predictive und Prescriptive Analytics im Unternehmen angewendet werden können. Es gibt noch weitaus mehr zur Verfügung stehende Methoden, welche in dieser Ausarbeitung jedoch nicht erläutert werden. Die Ausarbeitung soll nur einen Blick in die verfügbaren Methoden geben um die Breite von Advanced Analytics widerzuspiegeln.

6 Zusammenfassung und Fazit

Die Ausarbeitung zeigt, mit was sich Unternehmen vor der Einführung von Advanced Analytics befassen müssen. Dabei handelt es sich um organisatorische Aspekte wie einer geeigneten Strategie, eine veränderte Organisationsform, der Neuordnung von Rollen innerhalb des Unternehmens oder der Veränderungen der Anforderungen. Auch die technischen Aspekte werden sich mit der Einführung von Advanced Analytics verändern, weshalb ein Blick in eine vorgeschlagene Architektur empfehlenswert ist. Diese Erkenntnisse beantworten die in der Einleitung gestellte **RQ1**. Der konkrete Einsatz von Advanced Analytics durch verschiedene Techniken und derer dazugehörigen Methoden beschreibt, wo sich das Analyseverfahren überall einsetzen lässt. **RQ2** lässt sich durch die abschließende Abbildung 6 beantworten. Die Grafik zeigt die Ergebnisse von wichtigen organisatorischen und technischen Faktoren, die vor der Implementierung geklärt werden müssen. Die zielführenden Methoden in der Grafik wurden gewählt, da sie sich im Unternehmen sehr gut einbringen lassen und von hoher Bedeutung sein können. Sämtliche Ergebnisse wurden aufgrund der Literaturrecherche gewählt.

RQ1: Organisatorische und technische Faktoren	RQ2: Zielführende Methoden
- Strategie und Use Cases - Veränderte Organisationsformen - Neuordnung der Rollen - Veränderung der Anforderungen - vorgeschlagene Architektur	- Hidden Markov Model - Künstliche Neuronale Netze - Lineare Programmierung - Evolutionäre Algorithmen - Stochastische Simulation - Fuzzy Logik / Fuzzy Regeln

Abbildung 6: Erkenntnisse der Ausarbeitung

7 Ausblick

In dieser Ausarbeitung wurde nur eine Auswahl an Aspekten betrachtet, die zur Einführung von Advanced Analytics notwendig sind. Weitere Details zu den technischen Faktoren werden in der folgenden Masterarbeit des Autors herausgearbeitet. Zusätzlich werden in der Masterarbeit weitere wichtige Fragen wie z. B. die Risiken der Einführung, verfügbare Tools und die Datenaufbereitung geklärt. Sämtliche Ergebnisse werden auf ein Unternehmen angewendet und anschließend durch ein Experteninterview validiert.

Literatur

[1] K. Lepenioti, A. Bousdekis, D. Apostolou und G. Mentzas, "Prescriptive analytics: Literature review and research challenges", *International Journal of Information Management*, 2020. DOI: 10.1016/j.ijinfomgt.2019.04.003.

[2] B. C. Menezes, J. D. Kelly, A. G. Leal und G. C. Le Roux, "Predictive, Prescriptive and Detective Analytics for Smart Manufacturing in the Information Age", *IFAC-PapersOnLine*, S. 568–573, 2019. DOI: 10.1016/j.ifacol.2019.06.123.

[3] M. Kuhn und K. Johnson, *Applied predictive modeling*. New York: Springer, 2016, ISBN: 978-1-4614-6849-3.

[4] T. A. Runkler, *"Data Mining: Methoden und Algorithmen intelligenter Datenanalyse"*, Ser. Computational intelligence. Wiesbaden: Vieweg + Teubner, 2010. DOI: 10.1007/978-3-8348-9353-6.

[5] D. Larson und V. Chang, "A review and future direction of agile, business intelligence, analytics and data science", *International Journal of Information Management*, S. 700–710, 2016.

[6] L. Šikšnys und T. B. Pedersen, "Prescriptive analytics", *Encyclopedia of database systems*, 2016.

[7] P. Horváth, *Controlling*, 12., vollst. überarb. Aufl., Ser. Vahlens Handbücher der Wirtschafts- und Sozialwissenschaften. München: Vahlen, 2012. DOI: 10.15358/9783800644551.

[8] G. Satzger, C. Holtmann und S. Peter, "Advanced Analytics im Controlling - Potenzial und Anwendung für Umsatz- und Kostenprognosen", *Controlling*, S. 229–235, 2015. DOI: 10.15358/0935-0381-2015-4-5-229.

[9] J. Rose, M. Berndtsson, G. Mathiason und P. Larsson, "The advanced analytics Jumpstart: definition, process model, best practices", *Journal of Information Systems and Technology Management*, S. 339–360, 2017. DOI: 10.4301/S1807-17752017000300003.

[10] B. P. Bergeron, *Essentials of shared services*. New York: John Wiley & Sons, 2003, ISBN: 0471250791.

[11] U. Haneke, S. Trahasch, M. Zimmer und C. Felden, *Data Science: Grundlagen, Architekturen und Anwendungen*, Ser. Edition TDWI. Heidelberg: dpunkt, 2018, ISBN: 3864906105.

[12] R. Soltanpoor und T. Sellis, *"Prescirptive Analytics for Big Data"*. Springer International Publishing, 2016. DOI: 10.1007/978-3-319-46922-5.

[13] S. Lai, J. Seng und W. Lin, "Prediction of financial distress with text mining and hidden Markov model", *PACIS 2018 Proceedings*, 2018.

[14] X. Huang, X. Liu und Y. Ren, "Enterprise credit risk evaluation based on neural network algorithm", *Cognitive Systems Research*, S. 317–324, 2018. DOI: 10.1016/j.cogsys.2018.07.023.

[15] M. Chiu und G. Lin, "Collaborative supply chain planning using the artificial neural network approach", *Journal of Manufacturing Technology Management*, S. 787–796, 2004. DOI: 10.1108/17410380410565375.

[16] Y. Dodge, Hrsg., *The Oxford dictionary of statistical terms*, First published in paperback 2006. Oxford: Oxford Univ. Press, 2006, ISBN: 9780199206131.

[17] E. Steyerberg, "Stepwise Selection in Small Data Sets A Simulation Study of Bias in Logistic Regression Analysis", *Journal of Clinical Epidemiology*, S. 935–942, 1999. DOI: 10.1016/S0895-4356(99)00103-1.

[18] G. Sierksma und Y. Zwols, *"Linear and Integer Optimization: Theory and Practice, Third Edition"*, Ser. Advances in Applied Mathematics. CRC Press, 2015, ISBN: 978-1-4987-4312-9.

[19] L. Duan und Y. Xiong, "Big data analytics and business analytics", *Journal of Management Analytics*, S. 1–21, 2015. DOI: 10.1080/23270012.2015.1020891.

[20] V. Nissen, *Evolutionäre Algorithmen: Darstellung, Beispiele, betriebswirtschaftliche Anwendungsmöglichkeiten*. Wiesbaden: Deutscher Universitätsverlag, 1994. DOI: 10.1007/978-3-322-83430-0.

[21] M.-H. Jank, C. Dölle und G. Schuh, "Product Portfolio Design Using Prescriptive Analytics", in *Advances in Production Research*, Springer International Publishing, 2019, S. 584–593. DOI: 10.1007/978-3-030-03451-1_57.

[22] V. Bezděk, "Using Fuzzy Logic in Business", *Procedia - Social and Behavioral Sciences*, S. 371–380, 2014. DOI: 10.1016/j.sbspro.2014.02.498.

Frequenzmodulation-Parametrisierung für Radarwellen zur Erkennung von Gesten

Markus Gromer
Hochschule Reutlingen
Markus_Elia.Gromer@Student.Reutlingen-University.de

Abstract

In dieser Ausarbeitung wird ein System beschrieben, das zur Entwicklung von radarbasierten Human-Machine-Interfaces (HMI) eingesetzt wird. Das System basiert auf der von Texas Instruments zur Verfügung gestellten Hardware. Die Ausarbeitung beschäftigt sich hauptsächlich mit dem Hardwareaufbau, den Radareinstellungen und der Vorverarbeitung der Daten, so dass diese anschließend von geeigneter Algorithmik zur Erkennung von Gesten eingesetzt werden können. Ein wesentlicher Beitrag dieser Ausarbeitung ist die Beschreibung der Parametrisierung des Modulationsverfahrens der Radarhardware. Dabei werden die mathematischen Grundlagen beschrieben und es wird auf die Parameter eingegangen, die eine Erkennung von Gesten durch Radarwellen ermöglichen sollen.

CCS Concepts

• **Human-centered computing** → *Gestural input*;

Keywords

mmWave Radar, Gestensteuerung, IWR6843, 60 GHz, FMCW, Chirp

1 Einleitung

Das Forschen und Entwickeln an modernen User Interfaces ist ein breit gefächertes Gebiet. Diverse Einsatzgebiete erfordern nicht selten eigens hierfür angepasste Bedienmöglichkeiten. Für viele Anwendungen sind dabei kontaktlose Bedienmöglichkeiten von Maschinen oder Geräten von Vorteil und können Arbeitsabläufe vereinfachen. Im Virtual- oder Mixed-Reality-Bereich werden solche Human-Machine-Interfaces (HMI) bereits eingesetzt. Als Beispiel lässt sich hier die Mixed Reality Brille HoloLens[1] von Microsoft nennen. Auch im Bereich der Consumer Elektronik wird ein kontaktloses HMI schon eingesetzt. Als Beispiel kann hier Spark[2], eine Drohne des chinesischen Herstellers DJI, aufgeführt werden. Selbst in der Automobilbranche wird die Gestensteuerung mittlerweile als Ergänzung zu herkömmlichen Knöpfen und Schaltern eingesetzt. Moorstedt schreibt, dass in BMW Modellen der Bordcomputer schon jetzt über Gesten gesteuert werden kann.[1] Kontaktlose HMI's, die zur Gestensteuerung eingesetzt werden, werden hauptsächlich auf Basis von Vision, also kamerabasierten Systemen durchgeführt. Neue HMI-Entwicklungen, abseits von kamerabasierten Systemen, werden beispielsweise von Google mit ihrem Projekt Soli vorangetrieben. Google entwickelt in seinem Projekt Radarwellen-Sensorik zur Gestensteuerung und stellt verschiedene Möglichkeiten zum Einsatz dieser Technologie vor.[2] Der erste Einsatz des

[1] https://www.microsoft.com/de-de/hololens (besucht am 19.10.2019)
[2] https://www.dji.com/spark (besucht am 19.10.2019)

Betreuer	Cristobal Curio
Hochschule:	Hochschule Reutlingen Cristobal.Curio@Reutlingen-University.de
Betreuer	Dirk Staneker
Firma	Robert Bosch GmbH Dirk.Staneker@de.bosch.com

Wissenschaftliche Vertiefungskonferenz
27. November 2019, Hochschule Reutlingen
Copyright 2019 Markus Gromer

Chips mit dieser Technologie wurde durch Googles Smartphone Pixel 4 auf den Markt gebracht.[3] Radarwellen sind für bestimmte Anwendungen als HMI besonders interessant. Ob in einem Fahrzeug als Bedienung des Infotainmentsystems, in einer Fertigungshalle, in einem Labor oder im Krankenhausumfeld, wo Handschuhe getragen werden müssen und die Bedienung eines Touchscreens nicht möglich ist, oder zu Hause zur Steuerung der Musikanlage oder des Fernsehers - die Einsatzgebiete sind sehr vielseitig. Radarsensoren können dabei, anders als bei kamerabasierten Systemen, hinter einer Oberfläche verbaut werden und sind somit unsichtbar. Des Weiteren ist die Detektion von Objekten durch Radarwellen unabhängig von den gegebenen Lichtverhältnissen und kann im Freien, bei Tag und Nacht sowie in geschlossenen Räumen mit guter oder schlechter Ausleuchtung eingesetzt werden. Im Vergleich zu Kamerasystemen schreiben Smith *et al*, dass beispielsweise Tiefenkameras für das Erkennen von Positionen gut geeignet wären, diese jedoch Nachteile in der Erkennung von Richtungen oder verschiedener Handformen aufweisen.[3] Zusätzlich bietet Radar Vorteile bezüglich des Datenschutzes. Es ist vorstellbar, dass die Akzeptanz eines Radarsystems bei den Verbrauchern größer ist als bei visionbasierten Systemen, da dabei nicht gefilmt werden muss. Schon der Einsatz von Kameraüberwachung an öffentlichen Plätzen wird in den Medien kritisch diskutiert.[4] [5]
In dieser Ausarbeitung werden die Grundlagen der Entwicklung eines gestengesteuerten HMI beschrieben. Dabei geht es um den Hardware-Grundaufbau der Radar-Sensorik und der Anbindung an einen Computer zur Weiterverarbeitung der Radardaten. Es wird auf das `Chirp-Design` eingegangen, das ein wichtiger Aspekt bei der Einstellung des Modulationsverfahrens ist und einen entscheidenden Beitrag zum Erfolg der Gestensteuerung liefert. Des Weiteren wird auf die notwendigen Schritte eingegangen, um die Rohdaten der eingesetzten Sensorik zu übertragen und eine erste Aufbereitung der Daten durchzuführen.

2 Grundlagen

In diesem Abschnitt wird auf die Grundlagen des Themas eingegangen. Dabei werden neben dem Grundaufbau des Systems die notwendigen Grundlagen des Radars beschrieben. Hierzu zählt beispielsweise das Frequenzmodulationsverfahren - FMCW[6]. Außerdem wird auf die Grundlagen der Gesten eingegangen, die in einer aufbauenden Arbeit erkannt werden sollen.

2.1 Grundaufbau

In diesem Abschnitt wird auf den Grundaufbau des Gesamtprojekts eingegangen. Die grundsätzliche Signalverarbeitungskette wird in Abbildung 1 dargestellt. Zu Beginn wird ein Radarsensor und eine zugehörige Peripherie benötigt, mit der ein Radarsignal generiert, versendet und empfangen werden kann. Das zurückempfangene Signal bedarf anschließend einer umfangreichen Signalverarbeitung. Danach sollen die vorverarbeiteten Radardaten von geeigneter Algorithmik weiter verarbeitet werden. Dies ist der Teil, der die Gesten klassifizieren soll. Nach der

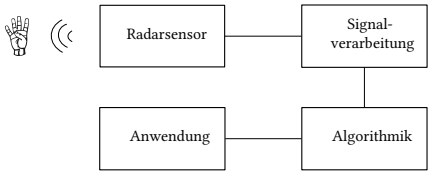


Abbildung 1: Grundaufbau des Projekts

[3] https://heise.de/-4555806 (besucht am 19.10.2019)
[4] https://heise.de/-3952482 (besucht am 19.10.2019)
[5] https://www.tagesspiegel.de/berlin/debatte-um-videoueberwachung-gesichtserkennung%2Dsoll-trefferquote-von-99-prozent-aufweisen/21220060.html (besucht am 19.10.2019)

[6] Frequency Modulation Continous Wave

Klassifizierung kann anschließend eine Anwendung bedient werden. In dieser Ausarbeitung wird auf die Radarsensorik sowie die Signalverarbeitung eingegangen. Die Klassifizierung durch Algorithmik und mögliche Anwendungen sind nicht Teil dieser Arbeit und werden in einem Folgeprojekt bearbeitet.

2.2 Radarsensor

In diesem Abschnitt werden die Grundlagen des Radarsignals beschrieben, die notwendig sind, um das eingesetzte Radarsystem und die Einstellungsmöglichkeiten zu verstehen. Eine Radarwelle ist grundsätzlich ein hochfrequentes Signal, das von einem Sender als Welle ausgestrahlt, von Objekten reflektiert und anschließend von einem Empfänger zurückempfangen werden kann. Abbildung 2 zeigt ein Blockdiagramm, das die hierfür benötigten Komponenten darstellt. Über den Signalgenerator wird das Signal generiert, anschließend über die TX-Antenne versendet, von einem Objekt reflektiert und von der RX-Antenne zurückempfangen. Der Mischer kombiniert anschließend die beiden Signale zu einem Zwischenfrequenzsignal (ZF-Signal), das anschließend ausgewertet werden kann. Aus dem Vergleich der gesendeten Welle mit der wieder zurückempfangenen Welle können Informationen über das reflektierte Objekt entnommen werden. Hierzu zählen beispielsweise die Position, die Entfernung, die Geschwindigkeit und die Bewegungsrichtung des Objekts.[7] [8][4] Im Folgenden wird auf das Frequenzmodulationsverfahren eingegangen. Es werden die einzelnen Schritte aufgezeigt, wie die Informationen aus dem Signal entnommen werden können.

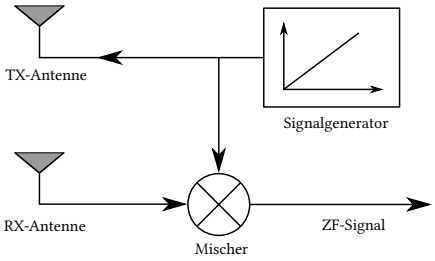


Abbildung 2: Blockdiagramm des FMCW Radars

2.2.1 Frequenzmodulation - FMCW Die Frequenzmodulation ist eines von mehreren Modulationsverfahren, das im Radarbereich eingesetzt wird. Bei FMCW-Radaren wird eine Welle ausgesendet, bei der die Frequenz über die Zeit linear steigt. Ein solches Signal wird auch als `Chirp` bezeichnet. Abbildung 3 zeigt oben einen beispielhaften Sinus-Chirp, aufgetragen über die Zeit. In Abbildung 3 unten wird das gleiche Signal in Abhängigkeit der Frequenz dargestellt. In beiden Abbildungen

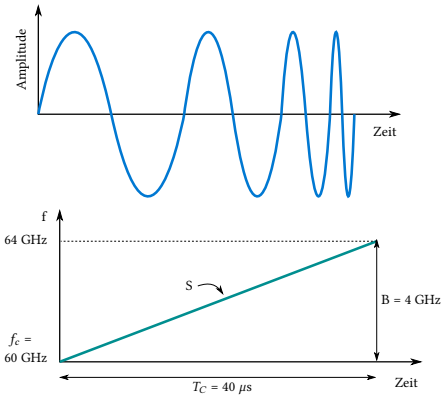


Abbildung 3: Oben: Zeitlicher Verlauf eines FMCW-Chirps. Unten: Frequenz aufgetragen über die Zeit eines FMCW-Chirps

[7] https://training.ti.com/intro%2Dmmwave%2Dsensing%2Dfmcw%2Dradars%2Dmodule%2D1%2Drange%2Destimation?cu=1128486 (besucht am 19.10.2019)
[8] https://training.ti.com/sites/default/files/docs/mmwaveSensing%2DFMCW%2Dofflineviewing_0.pdf (besucht am 19.10.2019)

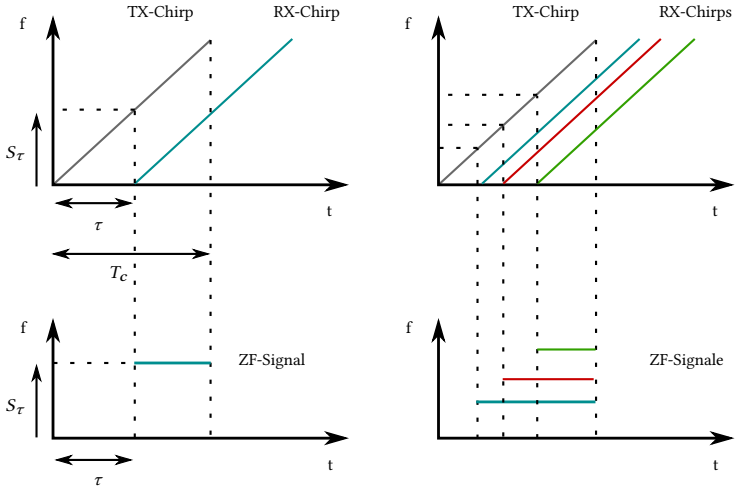


Abbildung 4: (1) ZF-Signal des Mischers aus dem TX- und RX-Chirp (2) ZF-Signale aus der Reflektion mehrerer verschiedener Objekte

ist der Anstieg der Frequenz über die Zeit zu erkennen. Die Eigenschaften des Signals sind eine Startfrequenz f_c von 60 GHz, einer Bandbreite B von 4 GHz und einer Dauer T_c von 40μs. Aus diesem Beispiel ergibt sich eine Steigung des Signals von S = $\frac{B}{T_c}$ = 100 $\frac{MHz}{\mu s}$. Die Steigung entspricht also der Änderungsrate des Frequenzsignals. Nachdem ein Chirp ausgesendet und zurückempfangen wurde, wird das Signal von einem Mischer verarbeitet. Der Mischer ist eine Komponente, die aus dem gesendeten und dem empfangen Signal ein neues Signal, das Zwischenfrequenzsignal bildet. Wird angenommen, dass die beiden Eingangssignale des Mischers aus einem Sinus $x_1 = sin(\omega_1 t + \phi_1)$ und einem Sinus $x_2 = sin(\omega_2 t + \phi_2)$ bestehen, so hat das ausgehende ZF-Signal folgende Eigenschaften. Die aktuelle Frequenz des ZF-Signals ist die Differenz aus den Frequenzen der Eingangssignale des Mischers. Ebenso ist die Phase des ZF-Signals die Differenz der Phasen der Eingangssignale. Durch den Mischer entsteht so ein neues Signal $x_{out} = sin((\omega_1 - \omega_2)t) + (\phi_1 - \phi_2)$ mit einer neuen Frequenz und einer neuen Phase. Abbildung 4 (1) veranschaulicht das neue ZF-Signal. Aus den beiden Signalen Tx und Rx wird ein neues Signal mit einer konstanten Frequenz gebildet, das durch Formel 1 beschrieben werden kann.[9][10][4]

$$f_{ZF} = \frac{S * 2d}{c} \tag{1}$$

Da das RX-Signal dem zeitversetzten TX-Signal entspricht, kann der Zeitversatz τ aus $\tau = \frac{2d}{c}$ berechnet werden[11] . Umgekehrt kann der Abstand des reflektierten Objektes durch die

[9] https://training.ti.com/intro%2Dmmwave%2Dsensing%2Dfmcw%2Dradars%2Dmodule%2D1%2Drange%2Destimation?cu=1128486 (besucht am 19.10.2019)
[10] https://training.ti.com/sites/default/files/docs/mmwaveSensing%2DFMCW%2Dofflineviewing_0.pdf (besucht am 19.10.2019)
[11] Lichtgeschwindigeit $c \approx 300.000.000 \frac{m}{s}$

Umstellung der Formel mit $d = \frac{\tau c}{2}$ berechnet werden. Wichtig dabei ist, dass das ZF-Signal nur in dem Zeitraum von τ bis zum Ende der gesendeten Frequenzrampe gültig ist. Wird das TX-Signal an mehreren Objekten reflektiert, werden mehrere Signale zurückempfangen. Wie in Abbildung 4 (2) zu erkennen, werden vom Mischer dann verschiedene ZF-Signale gebildet.

2.3 Gesten

In diesem Abschnitt wird auf die Gesten eingegangen, die durch die Radarwellen erkannt werden sollen. Grundsätzlich sollen Gesten erkannt werden, die von einer einzelnen Hand ausgeführt werden. Dabei spielt die Bewegung der einzelnen Finger eine wesentliche Rolle. Die Geste soll in einem Nahbereich von wenigen Dezimeter erkannt werden. Lien *et al* beschreiben beispielhafte Gesten, die mit verschiedenen bekannten Bewegungen von herkömmlichen Eingabegeräten assoziiert werden können. Hierzu zählen beispielsweise das Drücken eines Knopfes oder das Verschieben eines Sliders.[2] Abbildung 5 zeigt einige solcher beispielhafter Bewegungen, die von der Radarapplikation erkannt werden sollen. Für das Erkennen solcher Gesten und vorallem das Erkennen der Bewegungen ist es wichtig, dass die Radareinstellungen so gewählt werden, dass eine möglichst hohe Auslösung der Geschwindigkeit erziehlt werden kann.

3 Umsetzung

In diesem Abschnitt wird die Umsetzung auf der Texas Instruments Hardware beschrieben. Dabei wird zuerst allgemein auf die Hardware und anschließend die Einstellungen des Sensors eingegangen.

3.1 Texas Instruments Entwicklermodul

In diesem Abschnitt wird auf die Hardware von Texas Instrument eingegangen, auf der die theoretisch behandelten Verfahren umgesetzt werden. Texas Instruments bietet umfangreiche Hardware im Bereich der Radarsensorik im Bezug auf Anwendungen im automobilen sowie industriellen Bereich. Für dieses Projekt wird eine Hardwarekombination aus einem Radarsensor `IWR6843`, der in einem Plug-in-Modul integriert ist, einer Sensorplattform `Sensor Carrier Card Plattform - MMWAVEICBOOST` und einem Realtime Datenübertragungsadapter `Data Capture Card - DCA1000` verwendet. Der Radarsensor hat eine Grundfrequenz von 60 GHz und eine maximale Frequenzbandbreite von 4 GHz. Da die Hardware komplett von Texas Instruments zur Verfügung gestellt wird, ist ein verhältnismäßig unkomplizierter Hardwareaufbau möglich. Die Komponenten haben aufeinander abgestimmte Schnittstellen, die in diesem Projekt verwendet werden. Abbildung 6 stellt den schematischen Aufbau und die verwendeten Schnittstellen dar. In Abbildung 7 wird der reale Aufbau gezeigt. Zwischen der DCA und einem Computer wird eine Ethernetverbindung aufgebaut, um die Rohdaten so schnell wie möglich an einen Computer zu übertragen. Die USB-Verbindung zur DCA wird zur Übermittlung von Einstellungen an

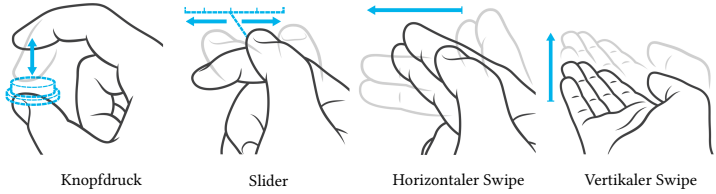


Abbildung 5: Beispielhafte Gesten von Lien *et al* (nach [2])

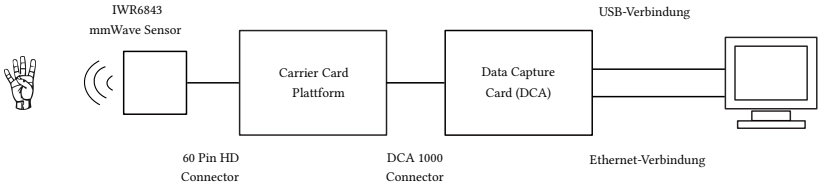


Abbildung 6: Schematischer Aufbau der Hardware

den Sensor verwendet. Dabei wird die Verbindung von der DCA an die Carrier Card Plattform durchgeschleift. Auf dem Computer läuft eine von Texas Instruments zur Verfügung gestellte Sofware `mmWave Studio`, die dabei unterstützt, eine Verbindung zu den einzelnen Komponenten herzustellen und Daten und Einstellungen zu übertragen. Die Software hilft dabei, einen Überblick über die möglichen Einstellungen zu bekommen und Radardaten per Knopfdruck zu übertragen. Für einen kompletten Verarbeitungsprozess ist die Software jedoch nicht geeignet. In Abschnitt 3.3 wird im weiteren Verlauf auf die angepasste Datenübertragung eingegangen.

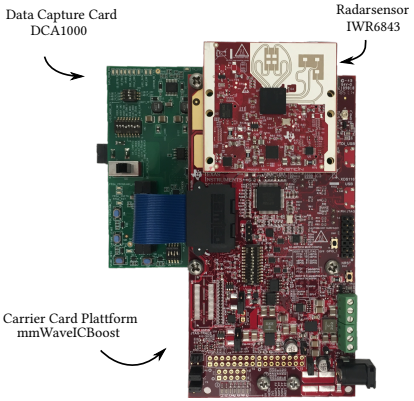


Abbildung 7: Aufbau der Hardware bestehend aus DCA1000, Carrier Card Plattform und dem Radarsensor IMR6843

Im folgenden Abschnitt werden die Einstellungen des Sensors näher betrachtet.

3.2 Einstellungen des Sensors

In diesem Abschnitt wird detailliert auf die Einstellung des Radarsensors eingegangen. Zur Ermittlung der passenden Einstellungen gibt es Tools und Hilfestellungen von Texas Instruments. Diese helfen, einen Überblick über die vielen Einstellungsmöglichkeiten zu bekommen. Neben einigen Grundeinstellungen geht es hauptsächlich darum, den Chirp der Modulation zu dimensionieren und eine geeignete Verkettung mehrerer Chirps in sogenannten Frames zu entwerfen. Im Folgenden wird auf die Parametrisierung des Chirps und der Frames eingegangen.

3.2.1 Die Chirp-Parametrisierung Die Parametrisierung des Chirps beinhaltet die Einstellung der Frequenzrampe, Wartezeiten und die Einstellung der Analog-Digital-Wandlung. Im Abschnitt 2.2.1 wurden die Grundlagen eines Chirps beschrieben. In diesem Abschnitt werden nun die konkreten Parameter beschrieben, die für die Anwendung der Gestensteuerung sinnvoll sind. Die Charakteristik eines Chirps hängt von den Eigenschaften des zu messenden Objektes ab. In Abschnitt 2.3 wurde schon festgestellt, welche Eigenschaften der Radarsensor bzw. der Chirp haben soll. Die wichtigste Eigenschaft ist eine gute Auflösung der Geschwindigkeit, da vor allem präzise Bewegungen gut erkannt werden sollen. Weniger wichtig ist beispielsweise das Messen von Objekten in großer Entfernung. Abbildung 8 zeigt einen einzelnen Chirp und die

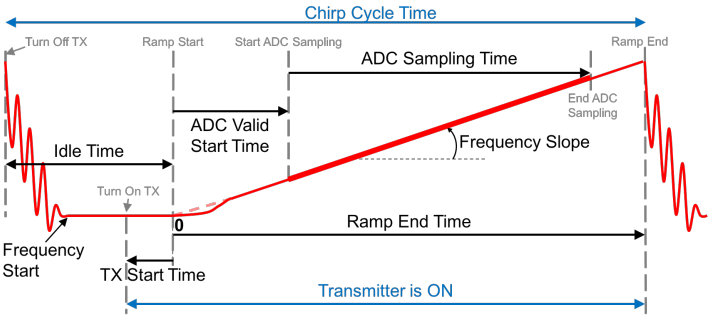


Abbildung 8: Einzelner Chirp und zugehörige Parameter. Frequenz aufgetragen über die Zeit (nach [5])

dazugehörigen Parameter. Ein Chirp-Zyklus besteht aus einer Leerlaufzeit (Idle Time) und der Zeit in der die Frequenzrampe ausgesendet wird ($T_c = Ramp_{End} - Ramp_{Start}$). Die Leerlaufzeit wird vom Frequenzgenerator benötigt, um die Frequenz aus dem letzten Zyklus wieder auf die Startfrequenz zu setzen. Die Frequenzrampe wird durch die Frequenzsteigung S (Frequency Slope) bestimmt und definiert über die Laufzeit der Rampe T_c die Bandbreite des Chirps. Die Bandbreite ist die Multiplikation der Rampenlaufzeit und der Frequenzsteigung und ist durch Formel 2 gegeben.[12][13]

$$B = T_c S \tag{2}$$

Die Zeitpunkte ADC-Start und ADC-End definieren den Zeitrahmen, in dem die Analog-Digital-Wandlung der Signale stattfindet. Abbildung 8 gibt einen Überblick, zu welchem Zeitpunkt welche Funktionen gestartet und beendet werden können.

3.2.1.1 Entfernung der Objekte Die Entfernung der Objekte kann durch die Umstellung der Frequenz des ZF-Signals aus Formel 1 berechnet werden und wird so durch Formel 3 beschrieben.

$$d_{max} = \frac{IF_{max} c}{2S} \tag{3}$$

Die maximale Entfernung ist also von der Steigung der Frequenzrampe abhängig. Je flacher die Rampe, desto höher ist die messbare Entfernung. Bei gleichbleibender Laufzeit T_c wird dadurch die Bandbreite der Frequenz auch geringer. Dies hat wiederum einen Einfluss auf die Auflösung der Entfernung.

3.2.1.2 Entfernungsauflösung Die Auflösung der Entfernung ist entscheidend, um mehrere Objekte voneinander zu unterscheiden. Je besser die Auflösung, desto näher können zwei Objekte beieinander sein und trotzdem als zwei verschiedene Objekte erkannt werden. Die Auflösung ist dabei von der Bandbreite eines Chirps abhängig und wird durch Formel 4 beschrieben.

$$d_{res} = \frac{c}{2B} \tag{4}$$

Je höher die Bandbreite eines einzelnen Chirps, desto höher wird auch die Auflösung der Entfernung und dadurch die Möglichkeit zwei

[12] https://training.ti.com/sites/default/files/docs/mmWave_Radar_Programming_Model.pdf (besucht am 19.10.2019)
[13] https://training.ti.com/mmwave-radar-programming-model?cu=1128486 (besucht am 19.10.2019)

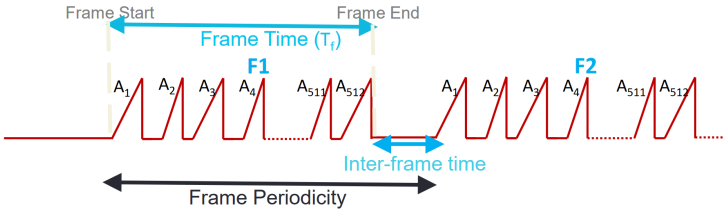


Abbildung 9: Mehrere Chirps zusammengestellt zu Frames (nach [5])

naheliegende Objekte voneinander zu unterscheiden.[14][15] Bei einer maximal möglichen Bandbreite von 4 GHz des Radarsensors kann also eine maximale Auflösung von 3,75 cm erziehlt werden.
Eine hohe Auflösung kann für das Erkennen von verschiedenen Gesten von Vorteil sein. Beispielsweise können einzelne Finger dadurch besser voneinander unterschieden werden. Da für das Erkennen von Gesten keine große Entfernung notwendig ist, kann eine steile Rampe mit maximaler Bandbreite und damit hoher Auflösung verwendet werden.

3.2.2 Die Frame-Parametrisierung Nach der Parametrisierung eines einzelnen Chirps kommt es nun darauf an, wie mehrere Chirps aneinander gehängt werden, um einen Frame zu bilden. Ein Frame ist also eine Sequenz aus vielen einzelnen Chirps und bildet damit das FMCW-Signal. Die Hardware von Texas Instruments ermöglicht das Aneinanderhängen von bis zu 512 Chirps in einem Frame. Abbildung 9 zeigt einen Frame mit mehreren Chirps und die verschiedenen Frame-Parameter. Ein Frame kann aus einem Chirp bestehen, der sich über die Framezeit T_f wiederholt oder aber auch aus mehreren verschiedenen Chirps, die sich wiederholen. Es können folglich 512 verschiedene Chirps definiert werden und zu einem Frame zusammengestellt werden. Zu den Parametern des Frames gehören die Periodizität und die Anzahl der Frames, die insgesamt gesendet werden sollen. Diese Frame-Paramter haben Einfluss auf die Auflösung der Geschwindigkeit der zu messenden Objekte. Die Geschwindigkeitsauflösung kann durch Formel 5 beschrieben werden.[16][17]

$$v_{res} = \frac{\lambda}{2T_f} \tag{5}$$

Wie in Formel 5 zu sehen, hängt die Auflösung der Geschwindigkeit von der Wellenlänge λ und der Framezeit T_f ab. Die Auflösung der Geschwindigkeit kann durch das Erhöhen der Framezeit T_f verbessert werden. Da dies für das Erkennen von Bewegungen bei Gesten essentiell ist, muss eine hohe Geschwindigkeitsauflösung erziehlt werden und damit die Framezeit T_f möglichst hoch gewählt werden.

3.2.3 Setzen der Einstellungen In diesem Abschnitt wird gezeigt, wie die in den letzten Abschnitten behandelten Einstellungen an die Hardware übertragen werden können. Zuerst muss ein Profil definiert werden, das die Haupteinstellungen eines Chirps beinhaltet.

[14] https://training.ti.com/sites/default/files/docs/mmWave_Radar_Programming_Model.pdf (besucht am 19.10.2019)
[15] https://training.ti.com/mmwave-radar-programming-model?cu=1128486 (besucht am 19.10.2019)
[16] https://training.ti.com/sites/default/files/docs/mmWave_Radar_Programming_Model.pdf (besucht am 19.10.2019)
[17] https://training.ti.com/mmwave-radar-programming-model?cu=1128486 (besucht am 19.10.2019)

Folgende Parameter werden in einem Profil eingestellt:

- Startfrequenz
- Leerlaufzeit
- ADC Startzeit
- TX Startzeit
- Samplerate
- Anzahl der Samples
- Rampenlaufzeit
- TX Energiereserve
- Cutoff-Frequenz 1
- Cutoff-Frequenz 2
- RX Verstärkung

Es können bis zu vier verschiedene Profile erstellt werden. Anschließend wird ein Chirp über die Chirp-Konfiguration eingestellt. Jeder Chirp wird zuerst einem Profil zugeordnet und bekommt somit die groben Einstellungen des Profils übertragen. Anschließend kann der Chirp mit weiteren Parametern verfeinert werden. Die Chirp Konfiguration hat folgende Parameter:

- Profil Index
- Startfrequenz
- Frequenzsteigung
- Leerlaufzeit
- ADC Startzeit
- TX-Kanal
- Phasenmodulation

Zum Definieren eines Chirps muss mindestens eine Zuordnung zu einem Profil erfolgen. Alle weiteren Parameter können von der Chirpeinstellung überschrieben werden, sofern dies erwünscht ist. Im nächsten Abschnitt wird auf die Datenübertragung eingegangen. Dies beinhaltet auch die Übertragung der behandelten Einstellungen an die Hardware.

3.3 Datenübertragung

In diesem Abschnitt wird auf die Datenübertragung vom Computer zur Hardware eingegangen. Da es mit der von Texas Instruments zur Verfügung gestellten Software nicht möglich ist einen guten Verarbeitungsprozess durchzuführen, der die Sensoreinstellungen übertragen, die Daten zurückempfangen und darüber hinaus direkt weiterverarbeiten kann,

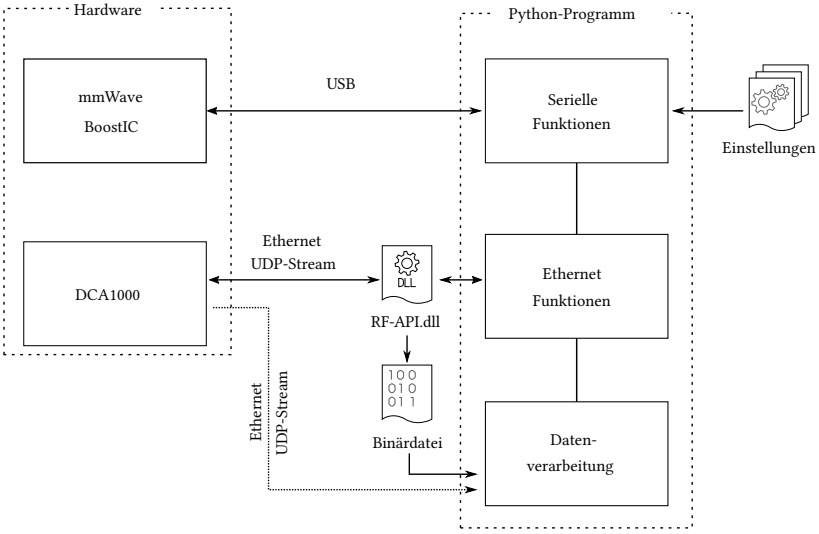


Abbildung 10: Aufbau der Software und Zusammenspiel mit der Hardware

ist es notwendig, eine Lösung für dieses Problem zu entwickeln. Für die Anbindung an die Hardware mit Datenübertragung wurde ein Python-Programm geschrieben, das diese Aufgaben übernehmen kann. Dies ermöglicht eine flexible Erweiterbarkeit und eine Schnittstelle zu einer aufbauenden Algorithmik. Die Sensoreinstellungen, die in Abschnitt 3.2 beschrieben werden, werden über eine serielle Schnittstelle an die Hardware übertragen. Die Einstellungen werden dazu aus einer Datei gelesen, die die ermittelten Chirp- und Frame-Parameter enthält. Alle Einstellungen für die High-Speed-Übertragung über die DCA1000 werden durch eine Bibliothek `RD-API.dll` über die Ethernetschnittstelle und das Netzwerkprotokoll UDP[18] übertragen. Die Kommunikation des Python-Programms mit der Hardware wird in Abbildung 10 dargestellt. Nachdem die Einstellungen an die Hardware übertragen wurden, kann der Sensor und damit die Datenübertragung durch einen Startbefehl über die serielle Schnittstelle gestartet werden. Anschließend werden die Rohdaten von der DCA1000 über einen UDP-Datenstrom an die Software gesendet. Die Daten werden dann von der RF-Api-Bibliothek in eine Binärdatei gespeichert. Parallel kann der Datenstrom auch direkt abgegriffen und weiter verarbeitet werden. Die UDP-Pakete müssen dazu zuerst entschlüsselt werden. Die Struktur der gespeicherten Daten kann in einem Manual von Texas Instruments nachgelesen werden.[6] Nachdem die Daten entschlüsselt sind, folgen Datenverarbeitungsschritte wie die Fourriertransformation. Das Python-Programm kann nun dafür verwendet werden, um einzelne Gesten aufzunehmen und um eine Algorithmik aufzusetzen.

4 Fazit und Ausblick

In dieser Ausarbeitung wurden Grundlagen über FMCW-Radare beschrieben und erläutert, wie die Einstellungen dieser Radare zu einer guten Möglichkeit führen, Gesten mittels Radarwellen zu erkennen. Es wurde detailiert darauf eingegangen, welche Parameter für das Design der Radarwellen notwendig sind. Die Zusammenhänge zwischen verschiedenen Parametern wurden erklärt. Außerdem wurde auf die Zusammenstellung des Signals mittels Chirps und Frames eingegangen.
Des Weiteren wurde eine sinnvolle Möglichkeit aufgezeigt, um die Einstellungen an einen TI-mmWave-Radar zu übertragen und Rohdaten von der Hardware zurück zu empfangen. Hierfür wurde ein Python-Programm implementiert, das diese Aufgaben übernimmt. Es können nun Radardaten zu einzelnen Gesten aufgenommen werden. Darauf aufbauend können die Rohdaten verwendet werden, um von einer geeigneten Algorithmik weiter verarbeitet zu werden und Gesten aus den Radardaten zu klassifizieren.

Literatur

[1] M. Moorstedt, “Technik: Aus die Maus”, *Focus*, 2015, https://www.focus.de/digital/multimedia/technik-aus-die-maus_id_5084053.html. (besucht am 19. 10. 2019).

[2] J. Lien u. a., “Soli: Ubiquitous Gesture Sensing with Millimeter Wave Radar”, 2016, http://doi.acm.org/10.1145/2897824.2925953.

[3] K. A. Smith u. a., “Gesture Recognition Using mm-Wave Sensor for Human-Car Interface”, 2018, https://doi.org/10.1109/LSENS.2018.2810093.

[4] C. Iovescu und S. Rao, “The fundamentals of millimeter wave sensors”, 2017, http://www.ti.com/lit/wp/spyy005/spyy005.pdf. (besucht am 19. 10. 2019).

[5] V. Dham, “Programming Chirp Parameters in TI Radar Devices”, 2017, http://www.ti.com/lit/an/swra553/swra553.pdf. (besucht am 19. 10. 2019).

[6] “Mmwave Radar Device ADC Raw Data Capture”, 2018, http://www.ti.com/lit/an/swra581b/swra581b.pdf. (besucht am 19. 10. 2019).

[18] User Datagramm Protocol

Model-View Design Patterns

Clemens Weissenberg
Hochschule Reutlingen
Clemens_Oliver.Weissenberg@Student.Reutlingen-University.de

Abstract

Industry applications with Graphical User Interfaces (GUIs) are often sold to customers as an intuitive experience to their internal, existing Application Programming Interface (API). While design patterns already cover basic guidelines on how to develop such GUIs, they lack information for details, for example with an API. The focus of this paper is to dive deeper into the design pattern Model-View-ViewModel (MVVM) and demonstrate the best practices based on an application implemented with the latest .Net Framework Windows Presentation Foundation (WPF). It is shown that separating the three modules View, ViewModel and Model into isolated C# projects supports the basic idea of MVVM of separation of concerns. It is also demonstrated that wrapping a complex API in the Model keeps the ViewModel simple. Furthermore, asynchronous API calls are accumulated in separate classes, called Manager. Also best practices for applications without an API are mentioned. These guidelines can be used for any GUI which is designed following the MVVM and uses an API.

CCS Concepts

• **Human-centered computing** → **Graphical user interfaces**; • **Software and its engineering** → **Software design engineering**; **Design patterns**; • **Applied computing** → Industry and manufacturing;

Keywords

GUI, Design Pattern, MVVM, WPF, Best Practises

1 Introduction

Applications in the industry often require a GUI application for an API. Design patterns like MVVM are not precise enough to restrict the developer from creating unmanageable source code. This is not only the fault of the developer. APIs are often not solely designed for GUIs and they consequently lack important features to support an easy integration for the GUI. This paper extends the design pattern MVVM to provide additional guidelines to lead application developers to a maintainable solution even with a non optimal API. This paper is not only designated to applications with APIs, but also applications without. All best practises are developed based on a Transducer Electronic Data Sheet (TEDS) Editor of the company Computer-Systeme-Messtechnik GmbH (CSM), called CSM TEDS Editor (CTE). The CTE is implemented with C# and WPF.

In section 2 insights in the state of the art of design patterns is provided, especially MVVM and WPF, and also shows other applications. The CTE application, which is the basis of this work, is explained in section 3 in detail. Section 4 lists the lessons learned and best practices which are founded by the CTE. Those same lessons learned and best practises are

Betreuer Hochschule:	Prof. Dr.-Ing. Marcus Schöller Hochschule Reutlingen Marcus.Schoeller@Reutlingen-University.de
Betreuer Firma	Paul VanOphem CSM GmbH PVanophem@CSMProductsInc.com

Wissenschaftliche Vertiefungskonferenz
27. November 2019, Hochschule Reutlingen
Copyright 2019 Hanah Büchner

then discussed in section 5. Lastly, the whole discussion is concluded in section 6.

2 State of the Art

In order to categorize the MVVM, a brief survey in Model-Design Patterns is given. Next, a more detailed overview of the MVVM is given.

2.1 Model-View Design Patterns

In the history of software with user interfaces, many design patterns were introduced. For example the Model-View-* (MV*) family, which includes Model-View-Controller (MVC), Model-View-Presenter (MVP) and MVVM to mention the most popular ones. All of these design patterns serve different purposes in different use cases. Syromiatnikov and Weyns demonstrate the essentials and differences of each pattern, as well as an illustration of the selection of the pattern. [1]

2.2 Model-View-ViewModel

MVVM is a three layer design. Starting from the lowest to the highest, we have the Model, the ViewModel and the View. The Model is the data model, which consists of the properties to be shown. It also describes the business logic of the data. The ViewModel is an abstraction layer of the Model and serves to prepare the data for the View by providing conversions for the properties of the Model. It also grants the View access to functions of the Model by defining commands. The View is the layer which defines how the data is visualized, by implementing graphical frameworks. The concept of MVVM enhances separation of concerns, which improves testability and readability. [2]

3 Transducer Electronic Data Sheet Editor

A common understanding of the TEDS editor first requires a brief survey of the products of a measurement equipment manufacturer and TEDS. Figure 2 shows the CTE prototype from CSM.

Figure 1: CSM CAN Minimodules product family (Computer-Systeme-Messtechnik GmbH: www.csm.de)

3.1 Measurement Manufacturer

There are many measurement equipment manufacturers whose modules support TEDS. In this paper the focus is on the company CSM. "CSM is a leading manufacturer of mobile measurement equipment and data acquisition systems. The company implements systems for fast, precise and flexible measurements in the vehicle development sector."[1]
Its main selling product family is measurement modules (in the following they are referred as modules), like the Controller Area Network (CAN) Minimodules, shown in figure 1. A common use case in the automobile industry for customers of measurement manufacturer is the following testing setup: A new car is tested by positioning voltage, temperature and strain gauge sensors. These sensors are then connected to the modules. The modules are daisy chained via the bus system. The bus system either is directly connected to a computer or to a data logger, which sends its data to a computer remotely or saves its data to storage, which can be read by a computer. The test developer finally interprets the data on the computer with the measurement software of choice.

[1] Computer-Systeme- Messtechnik GmbH: www.csm.de

3.2 Transducer Electronic Data Sheet

Each module needs to be configured via software in order to interpret the raw values of the sensor output in the form the user prefers. The configuration is necessary because the module must be made aware of which sensor is connected. The measurement setup may change over time, which requires a new configuration of the modules. To make life easier for the users, TEDS was invented in 2004.

The IEEE standard 1451 implements a protocol and interface for analog transducers to communicate digital [3]. In detail the IEEE 1451.4 standard defines TEDS. The TEDS, like the name suggests, is a definition of electronic data sheets for a transducer. These data sheets contain "[...] the critical information to identify, characterize, interface, and properly use the signal from an analog sensor." [4]

The IEEE 1451.4 standard supports both storing the TEDS on an embedded memory (in the following referred as TEDS chip) or as a file. Therefore it is possible to store the measurement configuration for an sensor on a sensor cable with TEDS support. [4]

The IEEE 1451.4 standard also defines a list of templates, excluding the Basic TEDS template. These templates define the how to characterize, interface and properly use the analog sensor, while the Basic TEDS template identifies the TEDS chip. TEDS chips always require a Basic TEDS template and can only contain one of each of the templates depending what kind of sensor is defined.

CSM and their competitors sell cables, which include TEDS chips with or without sensors attached. Specific CSM modules support TEDS and can read and write to the TEDS chip. CSM also implemented a proprietary TEDS file format, to store and load TEDS on a computer.

3.3 Purpose

The TEDS API was implemented to fulfill both purposes: Reading and writing TEDS from and to TEDS chips via CSM modules, and to load and store TEDS as files from and to a computer. A console application and the CTE are both built upon the TEDS API.

The main purpose of the CTE is to provide an intuitive GUI to the TEDS API. While the console application basically provides the same functionality, it lacks a good overview and support for user interactions. The CTE gives an overview over the parameters read from a TEDS chip and copies them locally. A user with proper privileges can change parameters. For this purpose, the GUI provides parameter specific graphical elements which supports inserting values and also validates them. The local copy can then be written on the TEDS chip. Additionally, the GUI provides the function to create new items for devices and TEDS templates. The CTE is designed for all of CSM's measurement modules with TEDS support.

3.4 Design

The design of the CTE is based on the MVVM design pattern. Therefore it consists of the Model, the ViewModel and the View. It will be shown which practises were used to fulfill the purpose in this subsection. First, the view is described to show features for implementation. Then, the ViewModel is explained focusing on the provision of those features. Lastly, the Model describes the implementation of the features.

3.5 View

The basic purpose of the View is linking graphical elements to the properties and commands of the ViewModel. In the case of the CTE, it visualizes the TEDS on the connected hardware as well as in files. The two main modes are the standard and the expert mode. While the standard mode only supports visualizing and editing the TEDS on the connected hardware, the expert mode also provides editing TEDS files. The main layout of the GUI is a ribbon window, which, as the name suggests, has a ribbon. Depending on whether the standard or expert mode is being used, it has two

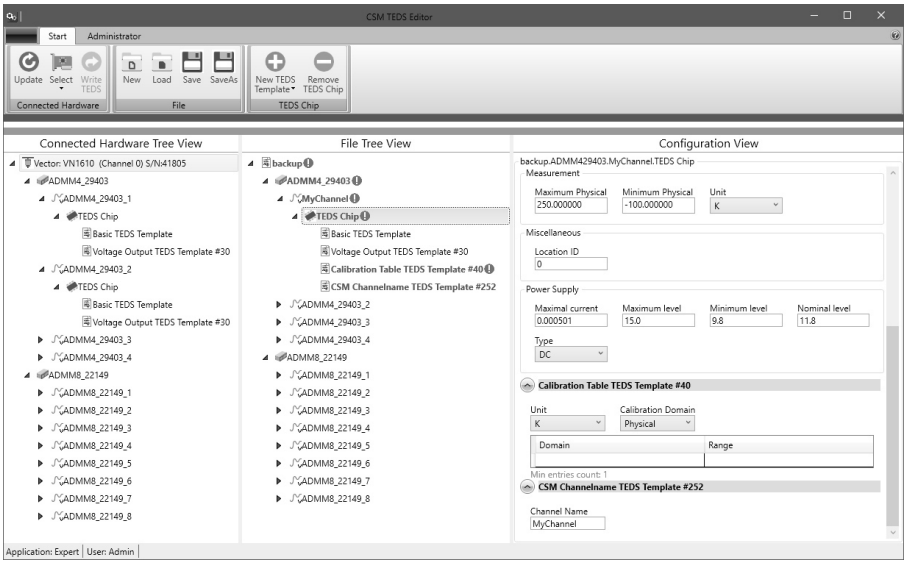


Figure 2: CSM TEDS Editor Prototype (Expert Mode)

or three columns. The following description based on the expert mode, which can be seen in figure 2. The first column on the left hand side is the "Connected Hardware Tree View". It shows the selected Hardware Interface, its modules, channels and connected TEDS chips in a tree view. The column in the middle is the "File Tree View" which shows the loaded file contents, such as modules or TEDS chips. The "Configuration View" on the right hand side, displays the parameter of the TEDS templates of the selected item of either Tree View. They are visualized as text boxes, drop down boxes and tables, depending on the parameter type. Privileged users can change these parameters. The status bar at the bottom of the window shows the current state of the application. The ribbon provides the basic functionality to interact with the columns. It has buttons to update and select the connected Hardware Interface and write changed TEDS templates to the connected hardware. In addition there are buttons for creating new TEDS templates and removing old ones. In the case of the expert mode there also buttons for creating and removing new modules and TEDS chips. These functions can also be found in the context menu of each tree item. Note, in the first attempt the native Microsoft *RibbonWindow* class was used, which got replaced by the telerik *RibbonWindow* class. This change was made due to insufficient documentation and deviation from conventions of WPF of the Microsoft *RibbonWindow* class. To make copying of tree view items easier, drag and drop is implemented for the tree views.

3.6 ViewModel

The main purpose of the ViewModel in the CTE is to wrap the Model, by introducing plain old Common Language Runtime (CLR) object (POCO)s. Another task which the ViewModel processes is validating properties by implementing validation attributes. By adding these attributes, WPF natively handles validation errors in the GUI and adds warnings. Also a generic tree view and a tree view item are implemented in the ViewModel. For each

POCO element a container is defined. This overhead is tolerated in order to preserve a the modularity of the POCO elements. To implement the drag and drop functionality, each tree view item container implements an interface which is responsible for the drag and drop.

3.7 Model

The Model wraps the TEDS API and also provides additional features, like Converters, Managers and Factories (classes which have only the purpose of creating objects). The TEDS API introduces one main class object as the handle to all functionalities of the TEDS API. Also, it defines interfaces for the hardware interface (in the case of the CAN, the CAN interface), the CSM TEDS file, the modules, the channels of the module, a TEDS chip, and the supported TEDS templates. By defining new POCO classes for all mentioned interfaces, the Model completely hides the TEDS API, so the ViewModel is completely independent from the TEDS API. In the case of the module and the channel interfaces, the TEDS API introduces two object definitions: One in context of the hardware and one in the context of the CSM TEDS files. The Model handles the conversions from the two objects. All service functionality regarding the reading and writing of TEDS, and saving and loading files is also handled by the Model. Each service functionality collection is implemented in a so called Manager. This Manager implements asynchronous function calls by utilizing C#'s *async* and *await* keywords, as well as the *Task* class. In order to prevent any race conditions, the one lock for each Manager is implemented. This guards every service call. In order to prevent race conditions, the Managers are implemented as Singletons. The GUI provides commands to create modules and TEDS templates based on types. To populate these items with the correct parameters a function is needed. Rather than implementing the logic in the respective constructor, factories are introduced for each class. These factories create these module and TEDS template objects and populate them with default parameters.

4 Results

In this section, best practices and lessons learned for MVVM and WPF are listed. They were developed while implementing the CTE. In section 5, their advantages and disadvantages are discussed.

4.1 Separation of Model, ViewModel and View in isolated projects

The MVVM design pattern defines the Model as unaware of the ViewModel, and the View and the ViewModel as unaware of the View. Additionally, the View is unaware of the Model. The CTE developed in this work strictly implements this concept of MVVM by dividing the modules Model, ViewModel and View in separate C# projects. The View references the ViewModel, and the ViewModel references the Model. Thus a violation of the dependencies defined by MVVM, is easily verifiable in the references.

4.2 Management of APIs

The GUI described in this work is built upon an API. While in theory an API can directly be used as the Model, in practice that is not always the case.

The GUI implemented in this work requires some object conversions of classes for different contexts. While the classes are well defined in the API, the object conversion definitions are missing. Thus, the Model handles these object conversions by introducing a new simple object, like a POCO, which inherits these conversions. It wraps the API to allow the ViewModel to have a uniform access to the different class objects. Alternatively, the ViewModel class references the objects itself, and implements the object conversion. This significantly increases complexity of the ViewModel compared with wrapping

the API. This expands the complexity of the ViewModel. Instead, all missing conversions which the API lacks, are handled by converters in the Model.
The ViewModel objects in the MVVM are responsible for wrapping the properties of the Model objects for the View. That includes type conversion of parameters, and offering function calls as commands and handling selection. The MVVM is very imprecise in defining which other functionalities belong to the ViewModel. It conforms to MVVM to implement service functionality in the ViewModel [5]. In contrast, the CTE implements all service functionality in separated and bundled service classes called managers. These managers can therefore keep track of each service state and provide them to the View, for example for the status bar.
In MVVM the View conventionally accesses the Model through the ViewModel. Still for some class type definitions, this strict separation is not always required. Simple types, which only consist of properties, without functions, can be passed through to the View, for example enumeration. The only information which can be added by the ViewModel is the localization information, the detailed language specific translation of the display text. So, instead of wrapping these types by the ViewModel, they are directly accessed by the View.
Localization is important in GUI implementations for international target groups. The MVVM design does not explicitly define where the localization should be located. Possible locations are the View and ViewModel. Since the ViewModel is responsible for conversions of parameters in the CTE, it gets implemented in the ViewModel.

4.3 *Async, await & Task*

A lot of APIs are not designed to run asynchronously, as the one mentioned in the CTE. A GUI, on the other hand, should always be designed to run asynchronously, to give the user a good responsiveness. In detail this means when the user executes a function by pressing a button, the GUI will still react to further user interactions while the current function is processed. For the purpose of this work, the *async* & *await* functionality with *Tasks* of C# was utilized in combination with WPF and MVVM. Processing asynchronous tasks presents some hurdles to the application developer. The first obvious one is the race condition. Because all of the tasks are service tasks, which are already collected in the manager classes, the race condition is solved by locks. One lock for each manager utilizes the tasks of the manager in serialized manner and consequently, automatically queues user actions with the GUI. Note, only the current GUI process is allowed to change properties that are visualized in the GUI. The View visualizes the properties of the ViewModel; the ViewModel references the properties of the Model; the Model properties are changed by the service tasks; these service tasks are processed asynchronously. Therefore, it leads to an error in the program if asynchronous service tasks change a Model property. To avoid this error the change of property needs to done in the GUI process by explicitly switching processes in the program code. Because it is required to change visualized properties in the current GUI process, the question arises where to handle the process management. While it would be possible to wrap every setter of a property of a ViewModel with a switch to the GUI process, it would add a lot of unnecessary boilerplate code. Instead, switches to the current GUI process are implemented in the manager classes in the CTE.

5 Discussion

The best practices and lessons learned in section 4 are discussed, in detail in this section.

5.1 Separation of Model, ViewModel and View in isolated projects

Having isolated projects for each module, namely Model, ViewModel and View, comes with a stricter separation than having different directories in one project. The greatest advantage of having this modularity is having separate folders even in the highest hierarchy. A new developer, or one who has not worked for a long time on the GUI, can instantly see and find the module he wants to work on. Additionally, errors which pop up when debugging the software are visualized with the information of the root path which consists of the related project path. Having separate projects comes with the advantage of having obvious dependencies. All dependencies are reflected in the project dependencies. A pure and simple MVVM dependency design, in which the View only depends on the ViewModel and the ViewModel only depends on the Model, can then be verified by the dependencies. Using MVVM with WPF often requires the implementation of some framework interfaces such as ICommand and INotifyPropertyChanged. These implementations can then be implemented in a separated library which can be reused by other MVVM with WPF projects. Also having additional graphical elements is sometimes necessary. Thanks to these strict separation and the obvious dependencies which come with it the structure of the code gets cleaner. Only the modules which really need these additional projects can add them as dependencies. Having separate modules improves working in parallel on one GUI. Thus a designer focuses on implementing the view, while the software developer can implement the ViewModel and the Model. It is even possible to implement multiple views in parallel for the same ViewModel and Model, by creating new projects for each View. Since there are no dependencies on the views even different frameworks for each View can be used without touching the other modules.

5.2 Management of APIs

Having a Model which wraps all unnecessary complexity of an API has several advantages. When the complexity is handled in the Model, the design of the ViewModel is simpler. So the ViewModel can be designed following the Keep it stupid simple (KISS) principle. Then GUI designers can directly see the properties of these classes and build their View on top of it. Having a simple ViewModel also implies that only the minimum of information necessary is available and accessible in the ViewModel. All redundant class types are filtered.

Since there are converters in place for converting classes of the same kind to different objects, this overhead is moved away from the ViewModel classes.

This also applies for managers. All service functionality can be implemented in the ViewModel as simple implementation of the ICommand interface, while all logic is located in the managers.

The localization is done in the ViewModel, which is already simplified. Therefore translation information can easily be found in the POCO elements of the ViewModel. Updating translations then only requires a change to the POCO elements of the ViewModel.

5.3 *Async, await & Task*

The services of the API are channeled in manager classes. These classes implement asynchronous function calls, to utilize a parallel processing of the GUI and the API. Having all services of the same nature in one class improves the debugging of this services. If issues appear, the developer can narrow it down to the one class. Since all managers group service tasks of one resource in one class, one shared lock in the class is enough to handle all race conditions which result using asynchronous function calls. This not only obligates

the developer to have a better structure in his code, but also enhances the transparency.
As mentioned in section 4 it is necessary to switch to the current GUI process to change a property which is visualized. One could argue that it is a potential problem for all ViewModel and therefore all these properties should handle switching to the current GUI process. However the cause of the problem is subject of asynchronous functions and thus should be handled in them. Because all asynchronous tasks are collected in the manager classes, they should be responsible for the switch of processes.

6 Conclusion

In this paper best practices of MVVM with WPF for applications with API are discussed, since applications with GUI are implemented often in the industry and need maintenance for many years. The discussion in section 5 demonstrates how these best practices improve the structure of applications which follow these requirements. Since MVVM is just a design pattern, following additional rules, such as wrapping the API in the Model and implementing asynchronous function calls in a Manager, improves quality of the application. Consequently maintaining such applications gets easier. In contrast following just MVVM without a consequent structure might lead to unmanageable source code. Despite the best practices mentioned in this work, not all guidelines are covered in this paper. Depending on the application, requirements may vary and might be covered in following research. An application that sticks to a well-defined structure will always achieve the best quality.

References

[1] A. Syromiatnikov and D. Weyns, "A journey through the land of model-view-design patterns", in *2014 IEEE/IFIP Conference on Software Architecture*, Apr. 2014, pp. 21–30. DOI: 10.1109/WICSA.2014.13.

[2] I. Manolov. (Mar. 2012). Model-view-ViewModel (MVVM) applications: General introduction – ivo manolov's blog, [Online]. Available: https://blogs.msdn.microsoft.com/ivo_manolov/2012/03/17/model-view-viewmodel-mvvm-applications-general-introduction/ (visited on 08/22/2019).

[3] Technical Committee on Sensor Technology, "IEEE standard for a smart transducer interface for sensors and actuators mixed-mode communication protocols and transducer electronic data sheet (TEDS) formats", *IEEE Std 1451.4-2004*, pp. 1–454, Dec. 2004. DOI: 10.1109/IEEESTD.2004.95745.

[4] National Instruments, Ed. (Aug. 2016). IEEE 1451.4 sensor templates overview - national instruments, [Online]. Available: http://www.ni.com/product-documentation/3468/en/ (visited on 09/19/2019).

[5] J. Smith. (Feb. 2009). Patterns - WPF apps with the model-view-ViewModel design pattern, [Online]. Available: https://msdn.microsoft.com/en-us/magazine/dd419663.aspx (visited on 08/22/2019).

Autorenverzeichnis

B
Büchner, H. 22

G
Gromer, M. 92

K
Kiefer, D. 69
König, G. 60

L
Lampprecht, T. 39

M
Malik, M. 49
Mauser, M. 81

S
Salb, D. 32

T
Taxidis, P. 1

U
Ulmer, A. 69

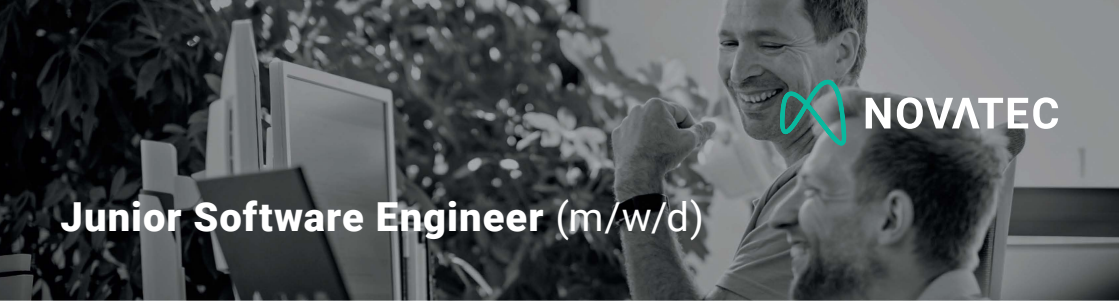
W
Weißenberg, C. 102
Wiedemer, T. 12

Reutlingen University
Fakultät Informatik
Human-Centered Computing
Alteburgstraße 150
D-72762 Reutlingen

http://www.huc.reutlingen-university.de
Telefon: +49 7121 / 271-4002
Telefax: +49 7121 / 271-4042

E-Mail: wvk@reutlingen-university.de
Website: wvk.reutlingen-university.de

Hochschule Reutlingen
Reutlingen University



Junior Software Engineer (m/w/d)

Standorte
Berlin, Frankfurt, Hamburg, Hannover, Karlsruhe, Leinfelden-Echterdingen, München, Zweibrücken

Das sind wir!
Zielgerichtet, zugewandt, direkt begeistert! Mit unseren rund 230 Mitarbeitern, an 9 Standorten in Europa, helfen wir unseren Kunden die Potenziale der Digitalisierung auszuschöpfen und ihr Geschäft voranzubringen. Wie das geht? Wir stehen für maßgeschneiderte IT-Beratung, Soft- wareentwicklung und Begleitung hin zu agilen Unternehmensorganisationen. Unsere Projekte werden zu 100 % agil umgesetzt, denn bei uns ist ‚Agilität' nicht nur ein Schlagwort. Sie ist Teil unserer DNA! Unseren Erfolg verdanken wir dabei jedem einzelnen Mitarbeiter. Ihre Kompetenz ist das Fundament, auf dem wir aufbauen. Für den Erfolg unserer Kunden entwickeln wir uns weiter – Tag für Tag!

Ihre Herausforderungen

In interdisziplinären agilen Teams entwickeln Sie gemeinsam mit Ihren Kollegen maßgeschneiderte Software-Lösungen für unsere Kunden. Dabei arbeiten Sie eng mit dem Kunden zusammen, diskutieren mit ihm seine Anforderungen und verstehen es, diese in ein sauberes Design und letztlich Clean Code zu überführen.

Neben der Entwicklung von Features unterstützen Sie die kontinuierliche Weiterentwicklung der Build Pipeline und weiterer Infrastrukturen.

Ihre Kollegen unterstützen Sie bei der Entwicklung qualitativ hochwertigen Codes durch Coaching, Pair Programming und Feedback-based Development.

Ihr Profil

Java-Ökosystem # Continuous Integration, Delivery und Deployment #Container und Cloud # Agilität # Neugierde # Eigeninitiative # Freude an Kundeninteraktion # Problemlösungskompetenz

Novatec Consulting GmbH
Human Resources
Dieselstraße 18/1
70771 Leinfelden-Echterdingen
Tel.: +49 711 22040-700
www.novatec-gmbh.de/karriere

Interessiert?
Wir freuen uns über Ihre Bewerbung (PDF) mit Angabe des Gehaltswunsches und möglichem Eintrittstermin an **karriere@novatec-gmbh.de**.

9 783000 642364