
Erfassung, Erkennung und qualitative Analyse von menschlicher Bewegung

André Ebert

Dissertation
an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität
München

vorgelegt von
André Ebert

München, den 23. Oktober 2018

Erfassung, Erkennung und qualitative Analyse von menschlicher Bewegung

André Ebert

Dissertation
an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität
München

vorgelegt von
André Ebert

1. Berichterstatter/in:	Prof. Dr. Claudia Linnhoff-Popien
2. Berichterstatter/in:	Univ.-Prof. Dr. Peter Reichl
Tag der Einreichung:	23. Oktober 2018
Tag der Disputation:	11. April 2019

Eidesstattliche Versicherung

(siehe Promotionsordnung vom 12.07.11, § 8, Abs. 2 Pkt. 5)

Hiermit erkläre ich an Eides statt, dass die Dissertation von mir selbstständig, ohne unerlaubte Beihilfe angefertigt ist.

München, 15. Mai 2019 André Ebert

Danksagung

Diese Dissertation entstand während meiner Zeit als wissenschaftlicher Mitarbeiter am Lehrstuhl für Mobile und Verteilte Systeme der Ludwig-Maximilians-Universität München. Dabei habe ich von verschiedenen Personen immer wieder wertvolle Unterstützung, Rat und Motivation erfahren und möchte diesen gerne Danken.

Mein ganz besonderer Dank gilt Frau Prof. Dr. Claudia Linnhoff-Popien, welche mich nicht nur durch die formale und inhaltliche Betreuung der vorliegenden Dissertation unterstützte. An ihrem Lehrstuhl konnte ich in einem inspirierenden und inhaltlich wertvollen Arbeitsumfeld wirken und möchte ihr dafür sowie für das stets von großem Respekt und Vertrauen geprägte Verhältnis danken.

Weiterhin gilt mein Dank auch Herrn Prof. Dr. Peter Reichl in seiner Rolle als Zweitberichterstatter, welcher mich auch während der Anfertigung dieser Dissertation durch die kritische Diskussion der Inhalte motivierte sowie Herrn Prof. Dr. Christian Böhm als Vorsitzenden der Prüfungskommission und Herrn Prof. Dr. Andreas Butz in seiner Rolle als Ersatzprüfer.

Meinen Kollegen am Lehrstuhl für Mobile und Verteilte Systeme danke ich für das außerordentlich angenehme, produktive und freundschaftliche Arbeitsklima. Besonderer Dank gilt hierbei Chadly Marouane und Sebastian Feld für die aktive inhaltliche Diskussion meiner Dissertation sowie Kyrill Schmid und Markus Friedrich für die in jeder Hinsicht tatkräftige Unterstützung und Motivation. Mein größter Dank gilt jedoch meiner Schwester Christina, meinen Eltern und Jana, welche mich immer bei der Verfolgung meiner Ziele unterstützen und mir in jeder Situation mit Rat und Tat zur Seite stehen. Vielen Dank euch allen.

Zusammenfassung

Visionen vom *Internet of Things* und der nahtlosen Einbettung der virtuellen Welt in den physischen Alltag des Menschen sind durch ubiquitäre Vernetzung, stationäre und mobile Computer sowie miniaturisierte Sensorik längst Realität geworden. Zusammen mit Algorithmen des Data Minings und der künstlichen Intelligenz werden so kontextsensitive Dienste und vernetzte Alltagsgegenstände geschaffen, welche einen immensen Mehrwert im privaten, kommerziellen und industriellen Raum schaffen.

Im Rahmen dieses Szenarios vielbeachtete Forschungsgebiete sind die Erschließung von menschlichem Kontext und die menschliche Aktivitätserkennung mithilfe von mobiler Sensorik. Während es auf dem Gebiet der rein quantitativen Erkennung von menschlicher Aktivität bereits viele Verfahren zur Vorhersage und Erkennung von Bewegungsereignissen auf Basis von Bewegungs- oder Tiefeninformationen sowie visueller Sensorik gibt, sind Konzepte zur feingranularen, automatisierten Analyse mit qualitativem Schwerpunkt bislang kaum existent. Typische Anwendungsgebiete für diese sind zum Beispiel die Identifikation von Notfallsituationen im medizinischen Bereich oder die Erkennung von Fehlstellungen und Anomalien bei physischer, menschlicher Aktivität.

Um solche Fragestellungen aus dem Bereich der Erfassung, Erkennung und qualitativen Analyse von menschlicher Bewegungsaktivität zu adressieren, wird in dieser Arbeit zunächst ein ganzheitliches, verteiltes Sensorsystem, welches auf Basis von Bewegungsinformationen menschliche Bewegungsaktivität untersucht, spezifiziert. Anschließend wird ein Vorgehen zur automatisierten und qualitativen Analyse individueller, wiederkehrender, menschlicher Bewegungsereignisse, mithilfe eines adaptiven Segmentierungsverfahrens und eines Konzepts zur Formalisierung und Diskretisierung von subjektiven Qualitätsmerkmalen in menschlichen Bewegungsabläufen, vorgestellt. Im Anschluss steht die qualitative Untersuchung von nicht vorhersehbarer, menschlicher Bewegungsaktivität im Fokus. Hierzu werden neue Konzepte zur Segmentierung und zur generischen Projektion der physischen, menschlichen Leistung des Menschen auf diskrete Merkmalsvektoren vorgestellt.

Zusammengefasst stellt die vorliegende Arbeit ein umfassendes Paket zur generischen Untersuchung von menschlicher Bewegungsaktivität vor. Mit diesem lassen sich quantitative und qualitative Analysen von Bewegungsaktivität für sowohl wiederkehrende als auch für nicht vorhersehbare, menschliche Bewegungsereignisse effizient umsetzen.

Abstract

The visions created by the *Internet of Things*, which encompass the seamless embedding of the virtual world into daily human life have become reality by now. One important reason for that is the ubiquitous availability of fast communication connections, of stationary as well as of mobile computers, and of miniaturized sensors and wearables. Thereby, the combination of this infrastructure along with data mining algorithms and concepts of artificial intelligence allow the creation of context-aware services and interconnected items of daily life, which provide vast added value for potential users.

Related to that, the recognition of human activity by using mobile sensors as a basis and the extraction of human context are currently much-noticed aspects within this field of research. By now, concepts focusing on quantitative recognition of human activity are a well studied area while examinations targeting qualitative analysis are currently handled only sparsely. Typical scenarios of usage for such concepts, which target qualitative analysis, are, e.g., the detection of emergency cases within medical environments or of anomalies within human movement during the conduction of physical activities.

In order to target research questions related to these qualitative topics, a holistic, distributed sensor system, which is capable of capturing and analyzing human motion is developed within this work at first. Subsequently, a concept for automated, qualitative assessment of individual, recurrent human motion events is presented. Therefore, it makes usage of a new adaptive segmentation algorithm and proposes an advance for formalization and discretization of subjective characteristics of quality within physical human activities. Afterwards, this work deals with the qualitative analysis of non-recurrent human motion. For this purpose, new concepts of segmentation are necessary and a new approach for generic projection of physical human performance onto discrete feature vectors is introduced.

To sum up, this thesis presents a comprehensive package for generic analysis of human motion activity. Therefore, it enables the detailed and efficient examination of recurrent as well as of non-recurrent human motion and moreover, allows a qualitative as well as a quantitative focus of analysis.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation und Zielsetzung dieser Arbeit	2
1.2	Zugrundeliegende Vorarbeiten	4
1.3	Aufbau dieser Arbeit	7
2	Definitionen und Grundlagen	9
2.1	Definitionen	9
2.1.1	Wiederkehrende und nicht vorhersehbare Bewegung	9
2.1.2	Qualitative und quantitative Untersuchung	10
2.1.3	Einfache und komplexe Bewegungsereignisse	10
2.2	Grundlagen der Signalverarbeitung	11
2.2.1	Filtertechniken	11
2.2.2	Signalkorrelation	12
2.2.3	Sliding-Windows	14
2.2.4	Signalstärke und spektrale Energie	16
2.2.5	Interpolation und Resampling	16
2.2.6	Vergleich von Zeitreihen	17
2.3	Maschinelles Lernen	19
2.3.1	Bereiche des maschinellen Lernens	20
2.3.2	Datensätze zum maschinellen Lernen	27
2.3.3	Feature-Engineering	29
2.3.4	Hyper-Parameter-Optimimierung	32
2.4	Aktivitätserkennung	33
2.4.1	Untersuchung von Bewegungsaktivität	34
2.4.2	Prozesskette zur Untersuchung von Bewegungsaktivität	34
2.5	Zusammenfassung	42
3	Erfassen und Erkennen von menschlicher Bewegung	43
3.1	Vorveröffentlichungen	43
3.2	Grundidee und Motivation	44
3.3	Verwandte Arbeiten	45
3.3.1	Arbeiten auf Basis von visueller Information	45
3.3.2	Arbeiten auf Basis von Tiefeninformationen	47
3.3.3	Arbeiten auf Basis von Bewegungsinformationen	48
3.3.4	Weitere Ansätze	50
3.3.5	Bestandsaufnahme	50
3.4	Konzeption einer verteilten Sensorplattform	51

3.4.1	Anforderungsanalyse	51
3.4.2	Spezifikation einer Grundarchitektur	52
3.5	Umsetzung von SensX	54
3.5.1	Hardware Ebene	54
3.5.2	Logische Ebene	58
3.6	Evaluation	64
3.6.1	Studiendesign	64
3.6.2	Untersuchung unterschiedlicher Sensorkonfigurationen	66
3.6.3	Erkennungsperformanz nach Ereignisklassen	68
3.6.4	Laufzeituntersuchung	69
3.6.5	Einordnung der Ergebnisse	71
3.7	Zusammenfassung	73
4	Qualitative Analyse von wiederkehrender Bewegungsaktivität	75
4.1	Vorveröffentlichungen	76
4.2	Motivation und Herausforderungen	76
4.3	Verwandte Arbeiten	78
4.3.1	Segmentierung von Bewegungsaktivität	78
4.3.2	Qualitative Untersuchung und Bewertung von Bewegung	80
4.4	Konzept zur qualitativen Untersuchung und Bewertung	82
4.4.1	Adaptive Segmentierung von Bewegungsereignissen	82
4.4.2	Extrema-Fingerprints zur Ereignisdetektion	86
4.4.3	Qualitative Bewertung und Formalisierung von Bewegung	87
4.4.4	Aufbau komprimierter Merkmalsvektoren	88
4.5	Evaluation	91
4.5.1	Automatisierte, qualitative Bewertung von Bewegung	92
4.5.2	Adaptive versus statische Segmentierungsverfahren	94
4.5.3	Aktivitätserkennung mit komprimierten Merkmalsvektoren	97
4.5.4	Laufzeituntersuchung	98
4.6	Zusammenfassung	100
5	Qualitative und Quantitative Analyse von nicht vorhersehbarer Bewegungsaktivität	103
5.1	Vorveröffentlichungen	104
5.2	Motivation und Herausforderungen	104
5.3	Grundlagen und verwandte Arbeiten	106
5.3.1	Untersuchung von Kletteraktivität	106
5.3.2	Beschreibung von physischer Leistung im Klettersport	109
5.4	Konzept zur Untersuchung nicht vorhersehbarer Bewegungsaktivität	110
5.4.1	Definition einer Prozesskette	110
5.4.2	Erfassung von Kletterbewegung	112
5.4.3	Zweistufiger Segmentierungsprozess	113
5.4.4	Differenziertes Merkmalsmodell zur Beschreibung menschlicher Bewegung	117
5.4.5	Bildung von diskreten Merkmalsvektoren	124

5.5	Evaluation	125
5.5.1	Studie zur Erfassung von Kletterbewegung	126
5.5.2	Feature-Review und Subset-Bildung	127
5.5.3	Automatisierte Bestimmung von Routenschwierigkeit	131
5.5.4	Automatisierte Unterscheidung von Studienteilnehmern	134
5.6	Diskussion und Zusammenfassung	136
6	Zusammenfassung und Ausblick	139
	Abkürzungsverzeichnis	145
	Abbildungsverzeichnis	150
	Literaturverzeichnis	151

1 Einleitung

Die im vorhergehenden Jahrzehnt aufgekommene Vision vom *Internet of Things* (IoT), welche durch die rasante Entwicklung und Verbreitung von Smartphones und Wearables sowie die Omnipräsenz von künstlicher Intelligenz und mobilen Kommunikationskanälen noch zusätzlich befeuert wird, ist in vielen alltäglichen Anwendungsszenarien mittlerweile Realität geworden. Zahllose smarte Entitäten und Objekte können drahtlos oder im mobilen Internet miteinander kommunizieren, die virtuelle Welt wird nahtlos in das physikalische Umfeld des Menschen eingebunden. Mithilfe der Anwendung von Algorithmen des Data Minings und der künstlichen Intelligenz werden personalisierte, kontextsensitive Services von zuvor nicht gekannter Innovationskraft und immensem Mehrwert für den Nutzer geschaffen [135, 192]. Somit sind der intelligente Kühlschrank, die persönliche Assistenz via Sprachsteuerung oder autonom navigierende Fahrzeuge im Straßenverkehr längst keine Utopie mehr, während selbstlernende Systeme gleichzeitig damit beginnen, auch industrielle Produktions- und Fertigungsprozesse unter dem Sammelbegriff *Industrie 4.0* zu revolutionieren [192, 94, 59].

Als ein großer Erfolgsfaktor für die Umsetzung dieser Vision kann die rapide Weiterentwicklung von leistungsfähiger Hardware, beispielsweise bezüglich Rechenkraft und Speicherkapazität, gesehen werden. Aber erst zusammen mit der Miniaturisierung und Mobilmachung von einstmalig nur stationär einsetzbaren Sensoren, Displays und Rechensystemen besteht auch die Möglichkeit, komplexe Vorhersagemodelle und Anwendungen künstlicher Intelligenz in mobilen Szenarien für verschiedenste Anwendungszwecke einzusetzen. Während zum Beispiel tragbare Sensorsysteme vor einigen Jahren noch aus unpraktischen, den Menschen behindernden Microcontrollern und Einplatinen-Computern oder teurer Spezialhardware bestanden, welche genau eine Art von Information messen konnten, gibt es nun Multisensorplattformen, welche eine Vielzahl von Umweltinformationen (z.B. Lichtstärke, Temperatur, Beschleunigung, Rotation, Höhe, u.a.) zugleich messen können und dabei immer noch ein physisches Format besitzen, welches kleiner als eine 1 Euro Münze ist [173, 111]. Darüber hinaus bieten auch Hersteller von Smartphones und den dazugehörigen Betriebssystemen mit ihren Produkten omnipräsente Multisensorplattformen an, welche schon aufgrund ihres geringen Preises für nahezu jedermann verfügbar sind und offen von Entwicklern mit immer neuen Anwendungen und innovativen Diensten bestückt werden können [170, 171].

Diese umfangreichen Möglichkeiten zur Ausstattung von Menschen mit minimal invasiver Sensorik und erheblicher mobiler Rechenkraft haben in den

letzten Jahren zu einem starken Anstieg der Entwicklung und Nutzung von Anwendungen zur Erkennung von menschlicher Aktivität, von kontext- und ortssensitiven Diensten sowie interaktiven Dienstleistungen und Applikationen im Bereich der *Augmented Reality* geführt. Beispiele hierfür kommen unter anderem aus den Bereichen Sport oder Medizin. Entwicklungen wie Fitness-Armbänder erlauben in Kombination mit Smartphone-Applikationen die Erfassung von alltäglicher, menschlicher Aktivität auf Basis von Bewegung (z.B. Laufen, Sitzen Schlafen, etc.) oder messen menschliche Vitalfunktionen wie die Herzfrequenz und den Hautwiderstand. Ebenso gehören im Fitnesskontext auch das Messen von zurückgelegten Strecken oder die zeitliche Erfassung von Schlaf- und Wachphasen bereits zum Alltag [159, 167, 163].

Aus diesen neuen Möglichkeiten zur einfachen und unabhängigen Erfassung von menschlicher Aktivität ergeben sich ebenso zahlreiche, neue Herausforderungen und ungelöste Fragestellungen. Neben der reinen Messung physikalischer Parameter von Aktivität und deren anschließender Untersuchung zur Erkennung und Quantifizierung von Bewegungsereignissen könnten diese zum Beispiel auch genutzt werden, um tieferegehende Analysen von qualitativer Natur vorzunehmen. Fragestellungen im Kontext solcher Untersuchungen bilden im Folgenden den Hauptfokus der vorliegenden Arbeit, welche sich mit der Erfassung, Erkennung und qualitativen Analyse menschlicher Bewegungsaktivität als Teilgebiet der menschlichen Aktivitätserkennung beschäftigt.

1.1 Motivation und Zielsetzung dieser Arbeit

In diesem Abschnitt werden nun die Ziele dieser Arbeit definiert, um so in präziser Form die Grundlage für die Inhalte der folgenden Kapitel zu bereiten. Die Untersuchung menschlicher Bewegungsaktivität als Teilgebiet menschlicher Aktivitätserkennung ist Schwerpunkt einer Reihe verschiedener smarterer Anwendungen aus den Bereichen Sport und Sportmedizin [2, 111, 147, 137, 139, 42, 211]. Diese basieren oft aus einer Kombination von mobiler Sensorik und damit gekoppelten Anwendungen (bspw. auf Smartphones, Tablets und sonstigen Rechenplattformen) und bieten dem Nutzer damit eine Auswahl verschiedener Dienste an. Diese umfassen zum Beispiel die Ermittlung und Visualisierung von statistischen Parametern wie Informationen hinsichtlich zurückgelegter Strecken oder das Erkennen und Zählen durchgeführter Wiederholungen einer bestimmten sportlichen Übung [165, 162]. Andere Funktionalitäten erlauben die Erstellung von teilweise anspruchsvollen Trainingsplänen und bieten dazu auch Informationen und Anleitungen zu komplexen Bewegungsabläufen. So werden Anweisungen, welche sonst nur mithilfe von professionellen Trainern erhältlich sind, auch für Laien und Amateur-Athleten verfügbar. Darüber hinaus sollen Ranking-Funktionen, welche den Vergleich des eigenen Trainingspensums mit dem anderer Athleten ermöglichen, zur Motivationssteigerung dienen [160].

Wie verschiedene Studien zeigen, haben Applikationen mit Funktionen dieser Art rasante Wachstumsraten in Nutzung und Verkauf vorzuweisen und

begeistern sowohl Laien als auch regelmäßig aktive Sportler. Aber auch Dritte, wie zum Beispiel Versicherungen oder Krankenkassen, denken schon daran, solche technischen Neuerungen in ihr eigenes Geschäftsportfolio miteinzubinden [4, 175, 164].

Dieser Trend ist zunächst als generell positive Entwicklung zu werten. Einschlägige Studien und Untersuchungen zeigen, dass regelmäßige sportliche und körperliche Aktivität nicht nur erheblichen Einfluss auf die menschliche Psyche hat, sondern auch das Selbstvertrauen, besonders bei Heranwachsenden, steigert und die mentale Gesundheit entscheidend stärken kann [29, 114]. Andere Arbeiten beschäftigen sich mit den positiven Auswirkungen sportlicher Aktivität auf chronische Erkrankungen, wie beispielsweise die Alzheimer-Krankheit [157].

Diesen zunächst offensichtlichen Vorteilen durch erhöhte körperliche und sportliche Betätigung stehen im Kontext von deren Ausführung zusammen mit mobilen Fitness-Apps und Wearables aber die Probleme fehlender Aufsicht und die Absenz feingranularer Bewegungsanalyse entgegen. So bieten einschlägige Fitness-Applikationen zwar Anleitungen zu anspruchsvollen Übungen und Trainingsplänen, sie messen Distanzen oder zählen und erkennen wiederkehrende Ereignisse – die Überprüfung von deren korrekter Ausführung oder die Bereitstellung von direktem Feedback qualitativer Natur, wie bezüglich der fehlerhaften Stellung menschlicher Gliedmaßen, fehlt aber nahezu vollständig. Ein Beispiel hierzu ist die Erfassung der wiederholten Durchführung einer Fitnessübung wie der Kniebeuge: Das reine Erkennen und Zählen einzelner Wiederholungen erlaubt zwar Rückschlüsse auf das quantitative Trainingspensum eines Athleten, geht der Athlet aber nicht weit genug in die Knie, so ist sein Training weniger anstrengend im Vergleich zur korrekten Durchführung. Enthält die Bewegung darüber hinaus sogar Fehlstellungen, wie beispielsweise durch Krümmung des Rückens, so kann es sogar zu Zerrungen oder ernsthaften Verletzungen an der Wirbelsäule kommen. So zeigen auch unterschiedliche Arbeiten und Studien, dass gerade Fehlstellungen in Bewegungen nicht nur zu temporären Verletzungen sondern, auch zu ernsthaften chronischen Erkrankungen führen können [93, 120, 109, 2]. Weiterhin ist die qualitative Untersuchung menschlicher Bewegungsaktivität nicht nur sinnvoll zur Verhinderung von Verletzungen durch Fehlstellungen. Im Vorfeld kann sie durch präzise Einschätzung von sportlicher Leistung dabei helfen, dass Athleten ihre eigenen Fähigkeiten kritischer einschätzen und reflektieren. Somit soll Selbstüberschätzung verhindert und die Motivation zur Korrektur fehlerhafter Bewegungsabläufe erhöht werden. Zusätzlich kann eine solche Untersuchung auch Überlastungen einzelner Gliedmaßen oder Muskelpartien sowie damit verbundene Verletzungen adressieren.

Hauptziel dieser Arbeit ist es, genau solche Konzepte zur qualitativen, automatisierten und feingranularen Untersuchung von menschlicher Bewegung, mithilfe von Werkzeugen des maschinellen Lernens, zu entwickeln. Dazu werden verschiedene Ansätze zur Erfassung, Verarbeitung und Untersuchung von

zwei Unterarten menschlicher Bewegungsaktivität, nämlich solcher bestehend aus wiederkehrenden und solcher bestehend aus nicht vorhersehbaren Bewegungsereignissen, konzeptioniert und evaluiert. In diesem Kontext wurde je Unterart von Bewegungsaktivität ein eigenes Anwendungsszenario aus dem Bereich Sport ausgewählt – die dazu entwickelten Untersuchungskonzepte wurden anhand dieser Anwendungsszenarien in den entsprechenden Kapiteln validiert. Die Schwerpunkte liegen dabei nicht nur auf der ganzheitlichen Erfassung von menschlicher Bewegung mithilfe von Wearables und der Entwicklung neuer Verfahren zur adaptiven Extraktion von einzelnen Bewegungsereignissen. Eine weitere zentrale Frage ist auch die Formalisierung und Diskretisierung subjektiver Beobachtungen zur qualitativen Untersuchung von menschlicher Bewegung sowie die Entwicklung eines differenzierten und generischen Modells zur Beschreibung von physischer menschlicher Leistung im sportlichen Kontext.

Diesen und weiteren Fragestellungen sind die folgenden Kapitel gewidmet. Zuvor geht Abschnitt 1.2 detailliert auf die dieser Arbeit zugrundeliegenden Vorarbeiten und Vorveröffentlichungen ein. Danach gibt Abschnitt 1.3 einen Überblick über die Struktur und den inhaltlichen Aufbau.

1.2 Zugrundeliegende Vorarbeiten

Ein Großteil der Inhalte dieser Arbeit wurde bereits im Rahmen internationaler Fachkonferenzen veröffentlicht. Nachfolgend wird der Eigenanteil des Autors der vorliegenden Arbeit (im Folgenden nur noch als „Autor“ bezeichnet) an den für diese Arbeit relevanten und bereits publizierten Inhalten, im Einzelnen aufgeschlüsselt. Die einzelnen Vorveröffentlichungen werden dabei in Reihenfolge ihrer ersten inhaltlichen Erwähnung in dieser Arbeit aufgelistet. Zusätzlich zu dieser Übersicht wird zu Beginn eines jeden der betroffenen Inhaltskapitel 3, 4 und 5 noch einmal eine detaillierte Übersicht über die einzelnen, bereits veröffentlichten Inhaltsteile gegeben. Dabei wird auch auf solche Inhalte, welche über die bereits publizierten hinausgehen, explizit eingegangen.

- **Collision Detection for Parked Vehicles [50]** Diese Arbeit stellt ein Konzept zur Erkennung von Kollisionen an Fahrzeugen vor, welches rein auf Basis der Bewegungsinformationen einer mobilen, an einem Fahrzeug angebrachten Sensorplattform, arbeitet. Das in dieser und in [49] entwickelte Verfahren zur Segmentierung von multidimensionalen Bewegungsereignissen mithilfe statischer Fenstergrößen war Grundlage des in Kapitel 3.5.2.1 entwickelten Verfahrens zur statischen Segmentierung menschlicher Bewegungsaktivität. Die Hauptinhalte dieser Vorarbeit, insbesondere bezüglich der Idee, des theoretischen Konzepts sowie der Implementierung eines Prototypen und dessen Validierung stammen vom Autor. Sebastian Feld unterstützte bei der Einordnung und Abgrenzung des entwickelten Ansatzes sowie bei dessen Vergleich zu verwandten

Arbeiten. Martin Werner war an der Ausarbeitung der Veröffentlichung beteiligt.

- **Segmented and Directional Impact Detection for Parked Vehicles using Mobile Devices [49]** Diese Arbeit beschäftigt sich ebenso wie [50] mit der Erkennung von physischen Kollisionen an Fahrzeugen. Es stellt im Vergleich zu [50] aber ein erweitertes Konzept vor, welches ebenfalls rein auf Basis der Bewegungsinformationen einer mobilen, an einem Fahrzeug angebrachten Sensorplattform arbeitet. Das darin entwickelte Verfahren zur Segmentierung von multidimensionalen Bewegungsereignissen mithilfe statischer Fenstergrößen war Grundlage des in Kapitel 3.5.2.1 entwickelten Verfahrens zur statischen Segmentierung menschlicher Bewegungsaktivität. Die Inhalte und Konzepte darin resultierten aus der kontinuierlichen Weiterführung der eigenen Vorarbeit des Autors in [50]. Sebastian Feld und Florian Dorfmeister haben sich im Vorfeld bei der Literaturrecherche und durch kritische Diskussion des Konzepts eingebracht. Zu den Eigenanteilen des Autors gehören die grundsätzliche Idee und die Entwicklung des Konzepts sowie dessen Umsetzung, Validierung und die anschließende Veröffentlichung.
- **SensX: About Sensing and Assessment of Complex Human Motion [51]** Im Rahmen dieser Arbeit wird ein verteiltes Sensorsystem zur ganzheitlichen Erfassung und Untersuchung von menschlicher Bewegungsaktivität konzeptioniert, entwickelt und evaluiert. Die Konzepte und Ergebnisse dieser Arbeit bilden den Kerninhalt des Kapitels 3. Zusätzlich bilden die in dieser Vorveröffentlichung entwickelten Konzepte die Basis für die in den Kapiteln 3, 4 und 5 entwickelten Analyseansätze. Der im Rahmen von [53] veröffentlichte Datensatz wurde ebenfalls mit dem im Rahmen dieser Vorpublikation entwickelten Systeme erfasst. Chadly Marouane war an der Vorrecherche zu verwandten Arbeiten sowie bei der kritischen Diskussion und Ausarbeitung der unterschiedlichen Vorkonzepte beteiligt. Beim Design und der Ablaufplanung der Vorstudie zur Erhebung eines Evaluationsdatensatzes war Sebastian Steer im Rahmen seiner Bachelorarbeit beteiligt. Claudia Linnhoff-Popien hat bei Vorarbeiten und im Rahmen kritischer Diskussionen unterstützt. Marie Kiermeier war an der Ausarbeitung der Veröffentlichung beteiligt. Zu den Eigenanteilen des Autors dieser Arbeit gehören die Entwicklung der Idee, die Konzeption und Umsetzung des Systems sowie dessen Evaluation und die hauptsächliche Ausarbeitung der Veröffentlichung.
- **Qualitative Assessment of Recurrent Human Motion [46]** Kerninhalt dieser Vorveröffentlichung ist ein umfangreiches, feingranulares Konzept zur automatisierten, qualitativen Bewertung von wiederkehrender menschlicher Bewegungsaktivität. Die einzelnen Inhaltsteile sind dabei zugleich auch die Kerninhalte von Kapitel 4. Michael Till Beck, Lenz

Belzner und Andy Mattausch waren im Rahmen der Vorrecherche und Diskussion der Inhalte sowie an der Ausarbeitung der Veröffentlichung beteiligt. Claudia Linnhoff-Popien wirkte im Rahmen von Diskussionen zu vorgeschlagenen Konzepten mit. Kernidee, Umsetzung und Evaluation der entwickelten Konzepte stammen in Hauptsache vom Autor.

- **Automated Recognition and Difficulty Assessment of Boulder Routes [54]** Diese Vorveröffentlichung beschäftigt sich schwerpunktmäßig sowohl mit der qualitativen als auch der quantitativen, automatisierten Analyse von nicht vorhersehbarer, menschlicher Bewegungsaktivität. Die einzelnen Kerninhalte dieser Publikation sind dabei zugleich Teil der Kerninhalte von Kapitel 4. Kyrill Schmid, Chadly Marouane und Daniel Eichinger waren im Rahmen der Vorrecherche sowie an der Diskussion der Vorkonzepte und daraus entwickelter Inhalte, beteiligt. Claudia Linnhoff-Popien war durchgehend als kritische Diskussionspartnerin hinsichtlich Konzeptionierung und Ausarbeitung der Inhalte beteiligt. Zu den Eigenanteilen des Autors an dieser Arbeit gehören die grundsätzliche Idee und Entwicklung des Konzepts sowie dessen Umsetzung, Validierung und die anschließende Veröffentlichung der Ergebnisse.
- **An Open, Labeled Dataset for Analysis and Assessment of Human Motion [53]** Diese Vorveröffentlichung stellt einen umfangreichen, offenen Datensatz mit Bewegungsinformationen von über 11.000 individuellen Wiederholungen von 8 Körpereigengewichtsübungen, durchgeführt von 26 unterschiedlichen Studienteilnehmern, vor. Der Datensatz ist die Grundlage der Evaluationen in den Kapiteln 3.6 und 4.5. Vorarbeit des finalen Studiendesigns dieser Veröffentlichung war eine vom Autor und von Sebastian Steer konzeptionierte Vorstudie. Adrian Klein assistierte bei der Durchführung der Hauptstudie im Rahmen seiner Masterarbeit. Christian Ungnadner war im Rahmen eines Praktikums bei der Segmentierung von Bewegungsereignissen sowie an der Bereinigung des Datensatzes beteiligt. Chadly Marouane war als Diskussionspartner an der Ausarbeitung der Veröffentlichung beteiligt. Eigenteile des Autors sind die grundsätzliche Idee und die Konzeption sowie die Durchführung der Studie. Die anschließende Datenaufbereitung und Veröffentlichung wurden ebenso hauptsächlich vom Autor durchgeführt.

Nachfolgend werden nun die Kerninhalte dieser Vorarbeiten den einzelnen Inhaltskapiteln zugeordnet. Die Kerninhalte des Kapitels 3 wurden vom Autor bereits in [51] publiziert. Die grundlegenden Konzepte zur Segmentierung von Bewegungsereignissen wurden zuvor bereits in [50] und in [49] veröffentlicht. Der im Rahmen der Evaluation verwendete Datensatz wurde bereits in [53] veröffentlicht. Die Ergebnisse zur System-Laufzeit in Abschnitt 3.6.4 wurden noch nicht veröffentlicht. Die Abbildungen 3.1, 3.2 und 3.7 sind bereits in [51] enthalten.

Die Kerninhalte von Kapitel 4 wurden vom Autor bereits in [46] publiziert. Der zur Evaluation verwendete Datensatz in Abschnitt 4.5 sowie das Vorgehen zur Formalisierung von qualitativen Beobachtungen in Abschnitt 4.4.3 wurden bereits in [53] veröffentlicht. Die Untersuchungen zur System-Laufzeit in Abschnitt 4.5.4 unter sowie der Performanzvergleich verschiedener Segmentierungsansätze in Abschnitt 4.5.2 sind neu und wurden zuvor noch nicht veröffentlicht. Die Abbildungen 4.3, 4.5, 4.1 und 4.2 sind bereits in [46] enthalten. Abbildung 4.1 ist bereits in [51] enthalten.

Die Kerninhalte von Kapitel 5 zur automatisierten, qualitativen und quantitativen Untersuchung von nicht vorhersehbarer, menschlicher Bewegungsaktivität wurden in [54] veröffentlicht. Die damit verbundenen Abbildungen 5.1, 5.2, 5.3 und 5.4 wurden für diese Arbeit überarbeitet und in der vorliegenden Form noch nicht veröffentlicht. Bei dem Vorgehen zur Nutzung eines differenzierten Modells von Kernkompetenzen zur Beschreibung physischer, menschlicher Leistung sowie zum Feature-Engineering in Abschnitt 5.4.4 handelt es sich um eine Erweiterung der ursprünglichen Veröffentlichung. Auf dieser Erweiterung aufbauende Inhalte und Abbildungen wurden zuvor noch nicht veröffentlicht. Die Evaluation in Abschnitt 5.5 wurde auf Basis der Neustrukturierung und Erweiterung des Feature-Engineerings erneut durchgeführt.

Die Publikationen [52, 48, 47], bei welchen der Autor federführend beteiligt war sowie die Publikationen [128, 130, 129, 186, 131], bei welchen dieser als Koautor beteiligt war, sind inhaltlich kein Bestandteil der vorliegenden Arbeit.

1.3 Aufbau dieser Arbeit

Die vorliegende Arbeit ist wie folgt strukturiert: In Kapitel 2 wird zunächst auf Begrifflichkeiten und Definitionen, welche im Verlauf dieser Arbeit von besonderer Relevanz sind, eingegangen. Anschließend werden wichtige Grundlagen hinsichtlich der Vorverarbeitung und der Untersuchung von Sensorsignalen eingeführt. Darüber hinaus wird ein Einblick in Konzepte des maschinellen Lernens und deren Einsatz zur menschlichen Aktivitätserkennung im Rahmen einer generischen Prozesskette gegeben. Das anschließende Kapitel 3 befasst sich schwerpunktmäßig mit der Konzeption, Implementierung und Evaluation eines verteilten Sensorsystems, welches auf Basis von Bewegungssensorik menschliche Bewegungsaktivität erfassen und erkennen kann. Dieses System dient in den Folgekapiteln als Grundlage zur feingranularen, automatisierten Analyse menschlicher Bewegung. Im Kapitel 4 wird ein neuer Ansatz zur qualitativen Untersuchung von wiederkehrender, menschlicher Bewegungsaktivität mithilfe von adaptiver Ereignissegmentierung sowie eines Konzepts zur Diskretisierung von subjektiven Beobachtungen in menschlichen Bewegungsabläufen vorgestellt. Kapitel 5 befasst sich schließlich mit der Untersuchung nicht vorhersehbarer, menschlicher Bewegungsaktivität. Dazu wird ein differenziertes Modell zur Beschreibung der physischen Kernkompetenzen von athletischer Leistung entwickelt, welches deren Projektion in diskrete, von maschinellen Lernalgo-

rithmen lesbare Merkmalsvektoren, ermöglicht. Abschließend fasst Kapitel 6 die Ergebnisse dieser Arbeit zusammen und gibt einen Ausblick auf sich daraus ergebende, zukünftige Herausforderungen und Forschungsfragen.

2 Definitionen und Grundlagen

Dieses Kapitel beschäftigt sich mit Grundlagen und verwandten Arbeiten im Kontext der Erfassung und Untersuchung von menschlicher Bewegung. Zuerst sollen in Abschnitt 2.1 spezifische Begrifflichkeiten definiert werden, welche während des Verlaufs der Arbeit zum Verständnis von Relevanz sind. Anschließend liegt der Fokus in Abschnitt 2.2 auf Konzepten der Signalverarbeitung und deren Anwendung zur Vorbereitung von rohen Sensorinformationen für maschinelle Lernprozesse. Zusätzlich werden verschiedene Distanzmetriken und Kostenfunktionen zum Vergleich von Zeitreihen vorgestellt. Danach gibt Abschnitt 2.3 einen generellen Überblick über verschiedene Disziplinen des maschinellen Lernens. Diese kommen in den Folgekapiteln bei der Untersuchung von menschlicher Bewegung zum Einsatz. Abschnitt 2.4 befasst sich mit der Erkennung menschlicher Aktivität und der Platzierung der Untersuchung von Bewegungsaktivität innerhalb dieser Thematik. Weitergehend wird auch eine generische Prozesskette des überwachten, maschinellen Lernens zur Aktivitätserkennung auf Basis von Bewegungsinformation präsentiert, welche die Grundlage für die Untersuchungskonzepte der Kapitel 3, 4 und 5 bildet. Abschließend fasst Abschnitt 2.5 die Erkenntnisse aus diesem Kapitel zusammen und gibt einen Überblick über deren Anwendung in den Anschlusskapiteln.

2.1 Definitionen

Nachfolgend werden Begrifflichkeiten, welche zum Verständnis der anschließenden Kapitel dieser Arbeit notwendig sind, aufgeführt und definiert.

2.1.1 Wiederkehrende und nicht vorhersehbare Bewegung

Diese Arbeit befasst sich in Kapitel 3 schwerpunktmäßig mit der Erfassung und Erkennung von menschlicher Bewegungsaktivität. In Kapitel 4 liegt der Fokus auf der qualitativen Untersuchung von wiederkehrenden Bewegungsabläufen, in Kapitel 5 stehen vor allem Konzepte zur qualitativen und quantitativen Untersuchung von nicht vorhersehbarer Bewegungsaktivität im Vordergrund.

In diesem Kontext ist eine Bewegungsaktivität als wiederkehrend definiert, wenn sie aus Sequenzen von einzelnen, vorhersehbaren, individuellen Bewegungsereignissen besteht. Diese einzelnen Bewegungsereignisse weisen Periodizitäten zueinander auf und sind oft von symmetrischen Bewegungsmustern einzelner Gliedmaßen des menschlichen Körpers zueinander gekennzeichnet.

Bedingt durch diese Eigenschaften besitzen alle Bewegungsereignisse innerhalb einer Bewegungsaktivität eine zeitlich vergleichbare Dauer und können so in einzelne Ereignissegmente unterteilt werden.

Bewegungsaktivitäten von nicht vorhersehbarer Natur können ebenfalls aus einer Vielzahl von unterliegenden Einzelereignissen bestehen. Im Unterschied zu wiederkehrender Bewegung folgen diese jedoch bezüglich der Reihenfolge ihrer Ausführung meist keinem vorhersehbaren oder bereits bekannten Muster. Einzelne Bewegungsereignisse können zwar bereits bekannt sein, aber auf deren Reihenfolge, zusammen mit gegebenenfalls noch unbekanntem Ereignissen innerhalb einer kontinuierlichen Ereignissequenz, kann nicht geschlossen werden. Daraus folgt, dass Periodizitäten vorhanden sein, aber nicht generell zugrunde gelegt werden können und die genaue Abschätzung der Dauer von Einzelereignissen oder der gesamten Bewegungsaktivität im Vorfeld nicht zwingend möglich ist.

2.1.2 Qualitative und quantitative Untersuchung

Ein weiterer Fokus dieser Arbeit liegt auf der qualitativen Untersuchung von menschlicher Bewegung, welche über die rein quantitative Untersuchung hinausgeht. Quantitative Untersuchungen umfassen in diesem Kontext vor allem das reine Erkennen und Quantifizieren von Bewegungsaktivität und den dieser zugrundeliegenden Einzelereignissen. Beispiele hierfür sind das Zählen von Schritten, die Unterscheidung von Gehen und Treppensteigen oder das Messen von gelaufenen Strecken und der dafür benötigten Zeit.

Ziel der qualitativen Analyse ist es hingegen, Einzelereignisse oder gesamte Bewegungsaktivitäten nicht nur zu erkennen, sondern auch eine Bewertung derselben nach unterschiedlichen Gesichtspunkten vorzunehmen. Beispiel hierfür ist die Identifikation von Fehlstellungen in menschlichen Bewegungsabläufen bei sportlicher Aktivität oder die automatisierte Untersuchung der Komplexität eines Bewegungsmusters.

2.1.3 Einfache und komplexe Bewegungsereignisse

Untersuchungsgegenstand der vorliegenden Arbeit sind komplexe menschliche Bewegungsereignisse. Bei diesen ist, im Gegensatz zu einfachen Bewegungsereignissen, der Einsatz aller Extremitäten des menschlichen Körpers zu ihrer Umsetzung gleichzeitig notwendig. Ein Beispiel für komplexe Bewegung aus dem sportlichen Bereich ist die Ausführung von Kletterbewegung. Bei einer solchen sind im Verlauf eines Aufstiegs alle Extremitäten des menschlichen Körpers im Wechselspiel zueinander im Einsatz.

2.2 Grundlagen der Signalverarbeitung

Im Folgenden werden Grundlagen der Verarbeitung von Sensorsignalen in Form von ein- und mehrdimensionalen Zeitreihen beschrieben. Dabei wurde in der Auswahl ein besonderes Augenmerk auf Werkzeuge und Vorgehen mit Bezug auf die Verarbeitung von Signalen zur Beschreibung menschlicher Bewegung gelegt, da diese im weiteren Verlauf der vorliegenden Arbeit sowie für das Verständnis der Inhalte der Kapitel 3, 4 und 5 relevant sind.

2.2.1 Filtertechniken

Um Grundrauschen in analogen Signalen zu reduzieren, diese zu glätten oder Sampling-Fehler zu entfernen, können sie mit Glättungs- oder Bandpassfiltern bearbeitet werden. Im Kontext von Bewegungsinformationen wie Rotation oder Beschleunigung zur Beschreibung menschlicher Bewegung, eignen sich vor allem Tiefpass- und Glättungsfilter [28, 213, 184]. Dies ist damit zu

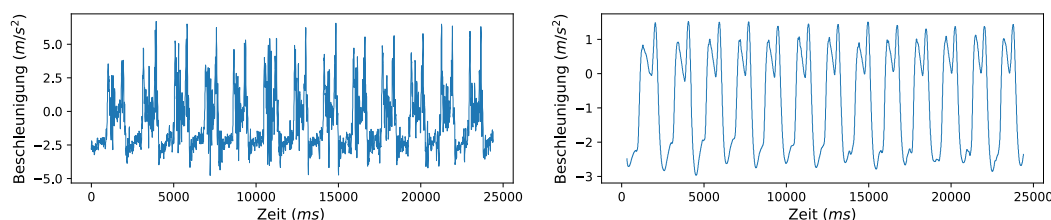


Abbildung 2.1: Beschleunigungsinformationen eines sich wiederholenden Bewegungsereignisses über die Zeit t in Millisekunden. Links sind die Rohdaten dargestellt, rechts sind diese nach Filterung mit einem Butterworth-Tiefpassfilter der 2. Ordnung zu sehen.

begründen, dass sich die relevanten Informationen menschlicher Bewegung im Verhältnis zu der hochfrequenten Erfassungskapazität aktueller Bewegungssensoren mit Sampling-Raten von weit über 100 Hz eher in niedrigerfrequenten Bereichen von $20\text{ Hz} - 60\text{ Hz}$ befinden [172, 117, 173, 177, 51, 40]. Das bedeutet in der Anwendung, dass höherfrequente Interferenzen und Signalrauschen den Filter nicht passieren können, während dies für tiefere Frequenzen, welche für die Bewegungsanalyse wichtige Informationen beinhalten, nahezu ungehindert zutrifft. Ein Filterbeispiel zur Glättung von menschlichen Beschleunigungsinformationen mit einem Butterworth-Tiefpass [28] ist in Abbildung 2.1 zu sehen. Ein wichtiger Parameter im Filterdesign ist dabei die Grenzfrequenz f_g (engl. Cutoff-Frequency, f_c), welche die Schwelle, ab welcher ein Signal den Filter nicht mehr ungehindert passiert und gedämpft wird, darstellt. Die Dämpfungscharakteristik, beispielsweise ob ein Filter das Signal mit einem steilen oder mit einem flachen Frequenzgang verändert, hängt maßgeblich vom individuellen Filter-Design ab. Im Folgenden werden verschiedene Filter zum Glätten von analogen Signalen kurz vorgestellt.

2.2.1.1 Butterworth-Filter

Der *Butterworth* Filter ist ein gängiger Filter im Bereich der Signal-Verarbeitung und soll in der Anwendung eine möglichst flache Übertragungsfunktion implementieren [28]. Das Hauptaugenmerk während des Filterprozesses liegt darauf, nicht nur ungewollte Frequenzen komplett aus einem Rohsignal herauszufiltern, sondern zusätzlich erwünschte Frequenzen mit einer gleichmäßigen Sensitivität passieren zu lassen. Wie Yu et al. zeigen, kann die für einen analogen Tiefpassfilter entwickelte Übertragungsfunktion von Butterworth auch auf ein digitales Filter-Design übertragen werden, um diese Anforderungen zu adressieren [213]. Weitergehend zeigte Butterworth, dass sein grundlegendes Filter-Design sich mit Modifizierungen auch zur Modellierung von Hochpass- und Bandpassfiltern sowie von Bandsperren eignet.

2.2.1.2 Savitzky-Golay-Filter

Der Savitzky-Golay-Filter ist ein Filter zum Glätten analoger Signale, welcher mithilfe von polynominaler Regression des k -ten Grades auf wenigstens $k+1$ äquidistanten Stützstellen einen geglätteten Wert für jeden in einer Zeitreihe enthaltenen Wert bestimmt [184]. Zur Implementierung eines Savitzky-Golay-Filters wird ein quadratisches Polynom angenommen, in welchem eine Gewichtung der einzelnen Signalwerte des Eingangssignals y_k der Länge k mithilfe der Filterkoeffizienten C_j vorgenommen wird:

$$y'_k = \frac{1}{f_{Norm}} \sum_{j=-m}^m C_j \cdot y_{k+j} \quad (2.1)$$

Der Normalisierungsfaktor f_{Norm} ergibt sich bei obestehender Filtergleichung aus der Summe aller Filterkoeffizienten C_j . Zusätzlich beeinflusst die Nutzung einer variablen Fensterbreite sowie variabler Glättungsfaktoren die Wirkung des Filters entscheidend. Durch Anpassung der Koeffizienten C_j kann er damit nicht nur eine Polynomialglättung, sondern ebenfalls eine gleitende Mittelwertbildung oder eine geglättete Ableitung realisieren [23].

2.2.2 Signalkorrelation

Die Korrelation von Signalen wird im Allgemeinen dazu genutzt, die Stärke und Richtung einer Abhängigkeit oder einer Verbindung zwischen zwei unterschiedlichen Signalen zu bestimmen. In der Aktivitätserkennung kann Korrelation auch dazu verwendet werden, um im eindimensionalen Raum zwischen Ereignissen zu unterscheiden und deren Länge oder Periodizitäten zu erkennen [178, 46].

2.2.2.1 Korrelationskoeffizienten

Zur Bestimmung des Grades der Korrelation zweier unterschiedlicher Signale werden Korrelationskoeffizienten für verschiedene Signaldimensionen berechnet. Dazu gibt es verschiedene, spezifisch nach Anwendung zu wählende, statistische oder geometrische Verfahren zur Koeffizientenbestimmung, welche unterschiedliche Signaleigenschaften differenziert gewichten können. Einer der meistgenutzten Koeffizienten ist der Korrelationskoeffizient nach *Pearson*, welcher als das Ratio der Kovarianz des Signals entlang der X - und Y -Achsen zum Produkt ihrer Standardabweichungen dargestellt wird [121]:

$$p_{x,y} = \frac{\text{cov}(x, y)}{\sigma_x \cdot \sigma_y} \quad (2.2)$$

Eine weitere Möglichkeit zur Bestimmung eines Korrelationskoeffizienten für normalisierte Datensätze sind geometrische Interpretationen. In [10, 89] nutzen die Autoren beispielsweise einen Ansatz basierend auf dem Ratio zwischen Punktprodukt und Länge des untersuchten Zeitfensters. Bei [105] hingegen wird der Koeffizient als Kosinus des Winkels zwischen zwei Vektoren, deren Produkt jeweils eine Dimension des von ihnen genutzten Sensormoduls repräsentiert, beschrieben.

2.2.2.2 Kreuzkorrelation und Autokorrelation

Die Kreuzkorrelation ist ein Maß zur Untersuchung von Ähnlichkeiten innerhalb zweier Signale. Sie wird zur Beschreibung von deren Korrelation bei einer Zeitverschiebung τ (dem sog. *Lag*) eingesetzt und ermöglicht damit die Identifikation eines bekannten oder ähnlichen Musters innerhalb zweier Signale $a(t)$ und $b(t)$. Bei zeitdiskreten Reihen wird $a(t)$ sequentiell in einzelnen Zeitschritten der Länge τ zu $b(t)$ komplett verschoben.

Ein Spezialfall der Kreuzkorrelation ist hierbei die Autokorrelation, bei welcher das Eingangssignal $a(t)$ mit einer zeitverzögerten Kopie von sich selbst kreuzkorreliert wird [144, 46]. Dies ermöglicht zum einen selbst in verrauschten Sensorinformationen das Erkennen von sich wiederholenden Mustern in $a(t)$, zum anderen können so auch Rückschlüsse auf Art und Anzahl von in $a(t)$ enthaltenen Periodizitäten sowie auf die Periodendauer Δt gezogen werden. Im Gegensatz zur Statistik, welche die Autokorrelation für stochastische Prozesse A_t als normierte Form der Autokovarianz berechnet, wird diese in der Signalverarbeitung als Faltung des zeitabhängigen Signals $a(t)$ mit sich selbst berechnet. Auch findet in der Signalverarbeitung keine Normalisierung durch Mittelwertsubtraktion und Varianzdivision statt. So lässt sich die Autokorrelation R für zeitdiskrete Signale mit dem Lag-Faktor l , welcher die diskrete

Verschiebung des Signals $a(t)$ zu seiner Kopie beschreibt, wie folgt definieren:

$$R_{a,a}(l) = \sum_{n=1}^n a_n a_{n-l} \quad (2.3)$$

Diese Notation ist allerdings nur für diskrete Signale von endlicher Energie und bekannter Fensterlänge, wie beispielsweise bei Bewegungsereignissen von definierter Dauer, anwendbar, während kontinuierliche Signale mit weitergehenden Definitionen als zufällige Prozesse beschrieben werden müssen. Durch

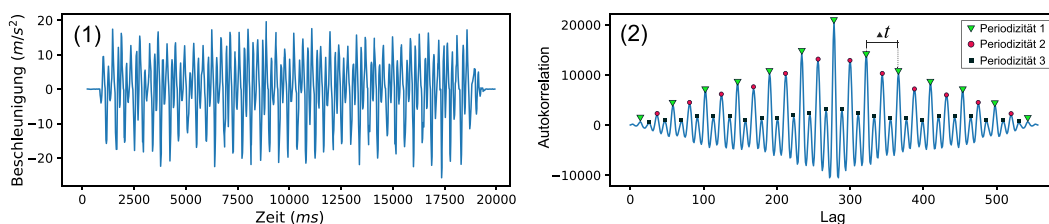


Abbildung 2.2: Ungefilterte Beschleunigung des Brustsensors eines Athleten in Y -Richtung (auf- und abwärts) (1) während der Durchführung eines Sets mit 20 Wiederholungen der Körpereigengewichtsübung *Hampelmann*. In (2) lässt die Autokorrelation eines Ausschnitts des Signals aus (1) Rückschlüsse auf enthaltene Periodizitäten und die Periodendauer Δt zu.

die komplette Verschiebung von $a(t)$ über sich selbst um dessen gesamte Laufzeit resultiert die Länge des autokorrelierten Signals als doppelt so lange, wie die des zugrundeliegenden Eingangssignals. Zusätzlich ist die Autokorrelationsfunktion eines periodischen Signals immer auch selbst periodisch.

Abbildung 2.2 (2) zeigt einen Ausschnitt des autokorrelierten Signals aus 2.2 (1), aufgenommen an der Brust eines Athleten während der Durchführung eines Sets von 20 Wiederholungen der Körpereigengewichtsübung *Hampelmann* in Y -Richtung (auf- und abwärts). Die Autokorrelation zeigt in Abbildung 2.2 (2), dass die Übung zu einer periodischen Beschleunigung an der Brust des Probanden führt, welche sich aus den individuellen Periodizitäten 1, 2 und 3 (vgl. Abbildung 2.2 (2)) zusammensetzt. Weitergehend lässt sich auch die jeweilige Periodendauer der einzelnen Wiederholungen Δt erkennen. Die höchste Korrelation zeigt sich im globalen Maxima der Autokorrelationsfunktion in der Mitte des Graphen, an welchem $a(t)$ und dessen Kopie genau deckungsgleich übereinander liegen.

2.2.3 Sliding-Windows

Vor einer weitergehenden Untersuchung kontinuierlicher Zeitreihen müssen diese in einzelne Sinnabschnitte unterteilt werden, welche anschließend sowohl unabhängig als auch abhängig voneinander betrachtet werden können. Dazu kann

das sogenannte *Sliding-Window* Verfahren angewandt werden, bei welchem ein Fenster der Länge Δt sequentiell über ein- oder multidimensionale Zeitreihen bewegt wird, um den jeweils aktuellen Inhalt für weitere Untersuchungen, wie beispielsweise zur Feature-Extraktion, in einzelne Segmente zu unterteilen.

Die initiale Fensterlänge $\Delta t = t_e - t_s + t_{safe}$ des Sliding-Windows kann für zeitlich genau definierte periodische Ereignisse beispielsweise mithilfe der Autokorrelation eines Signals bestimmt werden. Darüber hinaus ermöglicht sie auch Rückschlüsse auf deren Startzeitpunkt t_s und den Endzeitpunkt t_e sowie die Periodendauer [88, 46, 10]. Zusätzlich wird meist noch ein Sicherheitsabstand $\frac{t_{safe}}{2}$ an Start und Ende des initialen Fensters addiert, um mithilfe von Überlappung möglicherweise für ein Ereignis wichtige Informationen an Start und Ende des Segments nicht abzuschneiden.

Bei Ereignissen von unbestimmter Dauer oder in kontinuierlichen Zeitreihen lässt sich hingegen z.B. durch Errechnung der spektralen Energie (siehe Abschnitt 2.2.4) oder durch Bildung von Metriken wie der Standardabweichung oder dem Mittelwert eines Signals in Kombination mit Schwellwerten über die Länge des Fensters Aufschluss über den eigentlichen Fensterinhalt verschaffen. Anschließend können alle sequentiell überprüften Fensterlängen mit gleichen oder ähnlichen Charakteristiken demselben Aktivitätssegment zugeordnet werden [143, 111, 117, 10]. Die initiale Fensterlänge ist bei Ereignissen unbestimmter zeitlicher Länge also generell von untergeordneter Wichtigkeit, sollte aber nicht zu groß gewählt werden da es dadurch zu Segmentierungsfehlern an den Ereignisübergängen und damit zu fehlerhaften Vorhersagen in anschließenden Untersuchungen kommen kann.

Dadurch ergibt sich, dass die Wahl der Fenstergröße direkten Einfluss auf spätere Verarbeitungsschritte zur Untersuchung von Ereignissen hat und sowohl den Klassifizierungserfolg, als auch den Ressourcenverbrauch eines Systems entscheidend beeinflusst [88]. Zusätzlich wirkt sich die Fenstergröße auch auf mögliche Verzögerungen innerhalb einer Prozesskette zur Verarbeitung von Zeitreihen aus: Je größer die Fensterlänge gewählt wird, desto länger dauert es, bis ein Datensegment für den nächsten Verarbeitungsschritt zur Verfügung steht. Ist die initiale Fensterlänge hingegen zu klein gewählt, so werden die folgenden Verarbeitungsschritte umso öfter ausgeführt, wodurch die benötigte Rechenleistung ansteigt. Somit ergibt sich eine direkte Wechselwirkung zwischen der Präzision eines Segmentierungsprozesses, dem davon abhängigen Analyseerfolg und dem daraus resultierenden Berechnungsaufwand.

Ein im Kontext statischer Fensterlängen auftretendes Problem ist die fehlende Berücksichtigung der individuellen Länge der im Fenster beinhalteten Ereignisse. Vor allem im Bereich der Analyse wiederkehrender, menschlicher Bewegung ist die Länge unterschiedlicher Ereignisse innerhalb einer Klasse von Bewegungsaktivitäten selten exakt gleich, was bei steigenden Präzisionsanforderungen für deren Untersuchung und Bewertung dynamische Segmentierungsansätze notwendig macht [217]. Dazu wird in Kapitel 4 ein Ansatz zur adaptiven Segmentierung von menschlicher Bewegung entwickelt. Weitere

verwandte Arbeiten zur Segmentierung von menschlicher Bewegungsaktivität werden in Abschnitt 4.3.1 vorgestellt.

2.2.4 Signalstärke und spektrale Energie

Die Analyse der Energie eines Signals kann nicht nur zu dessen Segmentierung, sondern ebenso zur Erkennung von Ereignissen innerhalb einer Zeitserie, beispielsweise von Alltagsaktivitäten wie Radfahren, TV schauen verwendet werden [10]. Bao et al. berechnen dazu energetische Features als Summe der quadrierten Komponenten der Signalstärke mithilfe der diskreten *Fast Fourier Transformation* (FFT). Andere Ansätze nutzen die Summen der Flächen der Magnitude eines triaxialen Accelerometers, um den Energieaufwand bei täglichen, physischen Aktivitäten zu beschreiben [21, 62, 95, 133]. Dieser Energieaufwand, welcher sowohl zur Signalanalyse als auch zur Erkennung und Vorhersage von Aktivität nutzbar ist, wird als *Signal Magnitude Area* (SMA) bezeichnet, wobei die Eingangssignale $x(t)$, $y(t)$, und $z(t)$ mit der Dauer t in Bezug gesetzt sind:

$$SMA = \frac{1}{t} \left(\int_0^t |x(t)| dt + \int_0^t |y(t)| dt + \int_0^t |z(t)| dt \right) \quad (2.4)$$

Neben der SMA können auch die *Signal Vector Magnitude* und die *Differential Signal Vector Magnitude* [180, 95] genutzt werden, um Signale über ihre Energie zueinander in Bezug zu setzen und verschiedene Aktivitätsphasen innerhalb kontinuierlicher Zeitreihen zu isolieren [117, 111]. Weitere Ansätze errechnen die Energie eines Signals aus der quadrierten Summe von dessen spektralen Koeffizienten, welche über die Länge des untersuchten Fensters normalisiert werden. Auch hier wurde die Anwendbarkeit der Metrik durch Erkennen von Aktivitäten wie Fahrradfahren, Laufen und Gehen auf Basis von Bewegungsinformationen sowie durch Bestimmung verschiedener Nutzerkontexte wie das Gehen auf einer Straße oder das Führen einer Unterhaltung demonstriert [143, 101].

2.2.5 Interpolation und Resampling

Mithilfe der Interpolation von Zeitreihen können Signalausfälle und Messstörungen ausgeglichen oder durch unterschiedliche Sampling-Raten auftretende Asynchronitäten in multidimensionalen Folgen miteinander synchronisiert werden. Eine weitere Anwendung ist auch das Downsampling von hochfrequenten Signalfolgen, um den zur Weiterverarbeitung benötigten Rechenaufwand zu reduzieren oder deren Upsampling, um zum Beispiel fehlende Sampling-Punkte zu definierten Zeitpunkten zu approximieren. Dabei kommen beispielsweise die polynomiale Interpolation im univariaten oder die *Nearest Neighbour*-basierte Interpolation im multivariaten Fall zum Einsatz [185, 102].

Im univariaten Raum wird eine Interpolation meist über die Suche eines Polynoms realisiert, welches so exakt wie möglich durch die gemessenen Punkte eines erfassten Signals läuft und damit die Berechnung einer Näherung jedes einzelnen Punktes auf dem Graphen des Polynoms ermöglicht. Im multivariaten Raum wird dies im einfachen Fall zum Beispiel über die *Nearest Neighbour*-Interpolation implementiert, welche in einer oder mehreren Dimensionen anwendbar ist. Dazu nimmt der *Nearest Neighbour*-Algorithmus als approximierten Wert den des nächsten Nachbarn und ermöglicht eine stückweise, konstante Interpolation gesuchter Funktionswerte. Durch die fehlende Einbeziehung anderer benachbarter Punkte handelt es sich dabei um einen sogenannten naiven Approximationsansatz.

2.2.6 Vergleich von Zeitreihen

Um eindimensionale Signale ohne eine komplexe Prozesskette des maschinellen Lernens auf Ähnlichkeiten hin zu untersuchen, eignet sich der Einsatz von Distanzmetriken. Diese sind vor allem in Bereichen der Audio- und Sprachanalyse sowie in der Navigation und Trajektorienverarbeitung verbreitet, können aber unter Anpassungen auch für einfache Vergleiche von digitalen Signalen oder als Kostenfunktionen verwendet werden. Folgend werden ausgewählte Metriken und Algorithmen kurz vorgestellt und hinsichtlich ihrer Laufzeit und Komplexität eingeordnet. Die folgende Notation O bezieht sich dabei auf die Anzahl der elementaren Schritte, welche ausgehend von der Länge der Eingangssignale mit n Sample-Punkten benötigt werden, um die Distanz von Signal A zu Signal B zu berechnen.

2.2.6.1 Dynamic-Time-Warping (DTW)

Mithilfe des *Dynamic-Time-Warping* lassen sich zwei Zeitreihen $A = (a_1, \dots, a_n)$ und $B = (b_1, \dots, b_m)$ unterschiedlicher Längen n und m aufeinander abbilden [15]. Dazu wird zwischen A und B eine Matrix aufgespannt und mithilfe einer Kostenfunktion δ , beispielsweise der *Euklidischen Distanz* oder der *Mahalanobis Distanz*, werden die Kosten der Überführung eines jeden Parameters aus A in jeden Parameter aus B innerhalb der Matrix berechnet [100, 38]. Die Ähnlichkeit von A zu B kann anschließend durch Ermittlung des optimalen Überführungspfades $W = w_1, \dots, w_K$, welcher die minimalen Überführungskosten von A nach B aufweist, beschrieben werden. Dabei entspricht K der Länge des Überführungspfades, das k -te Element ist $w_k = (i, j)$, i steht für den Index eines Elements aus A und j für einen Index eines Elements aus B [15, 100]. Die globale Kostenmatrix D ist dabei wie folgt definiert:

$$D(i, j) = \delta(a_i, b_j) + \min \begin{cases} D(i, j - 1) \\ D(i - 1, j) \\ D(i - 1, j - 1) \end{cases} \quad (2.5)$$

Zusätzlich können bei der Suche des optimalen Überführungspfades Grenzbedingungen sowie Forderungen nach Kontinuität, Monotonität und andere gestellt werden. Um sicherzustellen, dass alle Elemente beider Zeitreihen bei der Suche miteinbezogen werden, muss jeder Überführungspfad bei $w_1 = (1, 1)$ beginnen und an den Endpunkten beider Zeitreihen $w_K = (i_k, j_k)$ enden, wobei $i_k = n$ und $j_k = m$ gilt. Damit ergibt sich folglich eine Komplexität von $O(nm)$. Um diese Komplexität zu verringern wurden in verwandten Arbeiten verschiedene Erweiterungen wie Beschränkungen mithilfe eines *Sakoe-Chiba*-Bandes oder eines *Itakura*-Parallelogramms, welche die zu berechnenden Felder innerhalb der aufgespannten Matrix begrenzen sollen sowie der angepasste *FastDTW*-Algorithmus, vorgestellt [90, 182, 183]. Alle DTW-basierten Distanzmetriken sowie damit verbundene Erweiterungen sind generell anfällig gegenüber Ausreißern.

2.2.6.2 Sum-of-Pair-Distance

Die *Sum of Pair*-Distanz (SoP) ermöglicht die Berechnung der Distanz zwischen zwei Zeitreihen A und B über die Summe der Distanzen δ aller ihrer, in sequentieller Reihenfolge zueinander korrespondierenden Wertepaare der Anzahl n , definiert als a_i und b_i [3]:

$$SoP(A, B) = \sum_{i=1}^n \delta(a_i, b_i) \quad (2.6)$$

Durch die Bedingung der Bildung von Punktepaaren ergibt sich die Notwendigkeit, dass beide Signale die gleiche Länge besitzen müssen. Da dies bei der sensorgestützten, elektronischen Signalverarbeitung aufgrund von Laufzeitverschiebungen, Verzögerungen und variierenden Sampling-Frequenzen in der Praxis selten der Fall ist, muss in der Regel eine Interpolation der Signale innerhalb eines Vorverarbeitungsschritts durchgeführt werden. Dadurch kann es zur ungewollten Verzerrung zeitlicher Information innerhalb der zugrundeliegenden Eingangssignale kommen. Durch seine im Vergleich zu den anderen vorgestellten Metriken geringe Komplexität $O(n)$ kann die SoP bei geeigneten Signalen gleicher Länge mit geringem Aufwand und Ressourcen-Effizient berechnet werden. Allerdings ist die SoP, wie auch der DTW-Algorithmus, anfällig gegenüber einzelnen Ausreißern; diese können die berechnete Distanz maßgeblich beeinflussen.

2.2.6.3 Hausdorff-Distanz

Die *Hausdorff*-Distanz H wird dazu genutzt, den Abstand $\delta(A, B)$ zwischen zwei endlichen Punktemengen $A = a_1, \dots, a_p$ und $B = b_1, \dots, b_q$ innerhalb eines metrischen Raums E zu messen und kann somit auch zur Messung des Abstandes zweier Signale verwendet werden. Dabei wird die Hausdorff-Distanz über das Maximum der minimalen Abstände von A zu B wie folgt definiert

[85]:

$$H(A, B) = \max(h(A, B), h(B, A)) \quad (2.7)$$

wobei

$$h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\| \quad (2.8)$$

und $\|\cdot\|$ eine zugrundeliegende Norm der Punkte aus A und B ist. Beispiele für hier zum Einsatz kommende Normen sind die L_2 - oder die *Euklidische Norm*. $h(A, B)$ ist die gerichtete *Hausdorff*-Distanz von A zu B und bestimmt den Punkt $a \in A$, welcher der am weitesten entfernte zu allen Punkten aus B ist und so den maximalen Minimalabstand von Signal A zu Signal B beschreibt. Vice versa beschreibt $h(B, A)$ den von B zu A . Das Maximum beider gerichteter Distanzen entspricht dann der *Hausdorff*-Distanz $H(A, B)$.

Die *Hausdorff*-Distanz besitzt aufgrund der Berechnung innerhalb einer Matrix eine Komplexität von $O(nm)$ und kann auch auf Signale unterschiedlicher Länge angewandt werden. Gegenüber Ausreißern verhält sich die *Hausdorff*-Distanz fehleranfällig.

2.2.6.4 Fréchet-Distanz

Die *Fréchet*-Distanz F ist ein Maß für die Ähnlichkeit, welche die Position und Ordnung der Punkte innerhalb der Zeitreihen A und B miteinbezieht. Die eigentliche Distanz F zwischen A und B wird hierbei definiert als das Infimum aller Reparametrisierungen α und β von $[0, 1]$ über das Maximum für alle $t \in [0, 1]$ der Distanz S zwischen $A(\alpha(t))$ und $B(\beta(t))$ [210, 56]:

$$F(A, B) = \inf_{\alpha, \beta} \max_{t \in [0, 1]} \{\delta(A(\alpha(t)), B(\beta(t)))\} \quad (2.9)$$

Bei der diskreten *Fréchet*-Distanz handelt es sich um eine Approximation der ursprünglichen *Fréchet*-Distanz, welche nur Punktpositionen betrachtet, deren Endpunkte sich an Eckpunkten der beiden polygonalen Zeitreihen befinden und niemals innerhalb einer Kante. Durch diese Limitierung in der Betrachtung möglicher Punktepaare lässt sich die Komplexität wie bei der *Hausdorff*-Distanz auf eine Komplexität von $O(nm)$ senken; in ihrer Empfindlichkeit gegenüber Ausreißern entspricht sie ebenfalls in etwa der *Hausdorff*-Distanz.

2.3 Maschinelles Lernen

Im Rahmen dieses Abschnitts werden verschiedene Disziplinen des maschinellen Lernens und damit verwandte Vorgehensweisen zur Validierung und Teilung von Datensätzen sowie zum Feature-Engineering und zu deren Selektion vorgestellt. Der Schwerpunkt liegt dabei vor allem auf der detaillierten Beschreibung von Konzepten im Kontext des überwachten maschinellen Lernens, da diese im Verlauf dieser Arbeit als Hauptwerkzeuge zur Untersuchung von menschlicher Bewegung zum Einsatz kommen. Zusätzlich wird auch ein

Überblick über verwandte Disziplinen und Algorithmen, wie beispielsweise das *Reinforcement Learning* oder die Architektur neuronaler Netze, gegeben.

2.3.1 Bereiche des maschinellen Lernens

Im klassischen, maschinellen Lernen wird zwischen zwei großen Teilgebieten des Lernens, dem überwachten Lernen (*Supervised Learning*) und dem nicht überwachten Lernen (*Unsupervised Learning*), unterschieden. Bei ersterem liegt der Schwerpunkt der Untersuchung vor allem auf dem Erkennen und der Vorhersage von Ereignissen oder Klasseninstanzen, während das Ziel des nicht überwachten Lernens das Verstehen und Ordnen von gegebenen Daten ist. Als weitere Disziplin hat das *Reinforcement Learning* in den letzten Jahren große Aufmerksamkeit in der Forschung erhalten. Im Folgenden werden diese Gebiete im Detail vorgestellt und einige einschlägige Algorithmen des maschinellen Lernens mit Anwendungsschwerpunkt Aktivitäts- und Bewegungserkennung genauer betrachtet. Ein zusätzlicher, eigener Unterabschnitt beschäftigt sich mit der Thematik neuronaler Netze, welche vor allem in den letzten Jahren für große Aufmerksamkeit auf dem Gebiet der künstlichen Intelligenz gesorgt haben. Diese können je nach Struktur und Netzkonfiguration sowohl Problemstellungen des überwachten als auch des nicht überwachten maschinellen Lernens lösen [181, 76, 141, 78].

2.3.1.1 Überwachtes maschinelles Lernen

Überwachte Lernansätze arbeiten mit gelabelten Eingangsinformationen. Typische Lernprobleme sind beispielsweise Klassifizierungen oder Regressionsanalysen. Dabei handelt es sich im Allgemeinen um eine Anzahl von Ereignisinstanzen oder Entitäten, welche jeweils über einen Merkmalsvektor im Merkmalsraum abgebildet und oft innerhalb einer Matrix organisiert werden. Jeder Instanz wird eine spezifische Klasse innerhalb des Raumes, ein sogenanntes Label, zugewiesen. Diese diskreten Instanzen in Vektorform werden als Eingangsinformationen für einen Klassifikator genommen, welcher als Ausgangsinformation einen einzelnen, diskreten Wert ausgibt, die sogenannte Klasse. Dazu wird mithilfe eines Lernalgorithmus und eines Trainingsdatensatzes mit Beispielinstanzen (X_i, y_i) ein Vorhersagemodell trainiert, wobei $X_i = (x_i^1, \dots, x_i^n)$ dem beobachteten Merkmalsvektor und y_i dem korrespondierenden Ausgangswert entspricht (vgl. Kapitel 2.4.2). Dieses Modell dient anschließend als Klassifikator. Im folgenden Schritt, der Klassifizierung, werden nun Testdaten, welche entweder zuvor während eines Unterteilungsschrittes in Test- und Trainingsdaten aus dem Gesamtdatensatz herausgenommen wurden (siehe Kapitel 2.3.2) oder online gesammelte, neue Instanzen mithilfe des Klassifikators ihren korrespondierenden Klassen y_t zugewiesen [181, 44]. Folgend werden einige relevante Algorithmen zur Klassifikation im Kontext des überwachten Lernens vorgestellt.

Random Forest Der *Random Forest* (RF) ist ein Klassifikations- und Regressionsverfahren, welches auf mehreren nicht miteinander korrelierten Entscheidungsbäumen basiert und damit das traditionelle Entscheidungsbaum-Modell erweitert [22]. Entscheidungsbäume sind in diesem Kontext als geordnete, gerichtete Graphen definiert, welche der Darstellung von Entscheidungsregeln dienen. Im Fall des *Random Forest* beruht jeder Entscheidungsbaum auf den Werten eines Zufallsvektors, welcher zwar unabhängig von allen Vektoren anderer Entscheidungsbäume, aber mit derselben Verteilung gesampelt wird. Das führt dazu, dass jeder Entscheidungsbaum des Waldes eine individuelle, unabhängige Entscheidungsstruktur repräsentiert. Für eine Klassifikation gibt jeder Baum innerhalb des Waldes auf Basis dieser Struktur eine Stimme für eine Klasse ab; die Klasse mit den meisten Stimmen wird als Klassifikationsergebnis angenommen. Verglichen mit anderen Lernalgorithmen bietet der *Random Forest* eine ganze Reihe an Vorteilen. Zum einen kann er sowohl zur Regression als auch zur Klassifizierung eingesetzt werden. Zum anderen tendiert der *Random Forest* bei einer geringen Anzahl von Entscheidungsbäumen zwar zu Overfittung und Generalisierungsfehlern (vgl. Kapitel 2.3.2.1), bei ausreichender Menge an Entscheidungsbäumen wird das Verfahren aber zunehmend robuster. Weitere Vorteile ergeben sich aus dem Aufbau unabhängiger Entscheidungsbäume, womit eine Parallelisierung des Algorithmus und damit ein verhältnismäßig schnelles Modelltraining realisiert werden kann. Die Trainingszeit steigt dabei linear mit der Anzahl der aufgebauten Entscheidungsbäume an [22, 122].

C 4.5 Bei dem *C4.5* handelt es sich ebenso wie bei dem *Random Forest* um einen Algorithmus zum Aufbau von Entscheidungsbäumen, welcher von Ross Quinlan auf Basis seines zuvor vorgestellten *ID3*-Algorithmus entwickelt wurde [155]. Dabei werden im *C4.5* Einschränkungen des *ID3*, bezüglich Anforderungen an fehlerfreie, vollständige Eingangdatensätze sowie bezüglich diskreter Merkmalswerte, behoben. Er implementiert außerdem zusätzliche Pruning-Methoden und optimierte Vorgänge zur Auswahl von Klassenattributen, sodass Features mit vielen Ausprägungen keine Bevorzugung gegenüber solchen mit weniger Ausprägungen erfahren. Weiterhin implementiert er über eine Fenstertechnik das mehrfache Erzeugen und Beschneiden von Bäumen. In der Anwendung generiert der *C4.5* aus gegebenen Trainingsdaten Entscheidungsbäume über das Konzept des mittleren Informationsgehalts. Der erste Knoten innerhalb eines Baumes selektiert das Attribut des Inputvektors $X_i = (x_i^1, \dots, x_i^n)$ als Entscheidungsmerkmal, mit welchem sich die gegebenen Eingangsinstanzen am effektivsten in Teilmengen unterteilen lassen. Diese Teilmengen können wiederum einer der möglichen Klassen zugeordnet werden. Als Unterteilungskriterium wird jenes mit dem höchsten normalisierten Gewinn an Information ausgewählt, um die Klassenentscheidung zu treffen [81]. Dabei spricht man von der sogenannten *Kullback Leibler*-Divergenz. Anschließend wandert der *C4.5* mit dem gleichen Verfahren rekursiv über die

entstandenen Teilmengen. Ist keine weitere Unterteilung mehr möglich, so entsteht keine weitere Verästelung des Baumes, sondern ein Blatt welches die Zielklassifikation enthält [155, 181].

Naive Bayes (NB) *Naive Bayes*-Klassifikatoren basieren auf dem *Bayesschen Theorem* und eignen sich vor allem zur Anwendung auf hochdimensionale Eingangsinformationen [31]. Durch seine Einfachheit realisiert er im Vergleich zu anderen Algorithmen eine schnelle Ausführungszeit bei verhältnismäßig guter Klassifizierungsperformanz. Dazu ordnet der *Naive Bayes* jeder Klassifizierung Kosten über ein Kostenmaß zu und minimiert anschließend diese für alle Klassifizierungen. Der *Naive Bayes* geht dabei von der Grundannahme aus, dass alle Attribute eines gegebenen Merkmalsvektors voneinander unabhängig sind. Obwohl dies in der Anwendung nur selten zutrifft, da einzelne beschreibende Attribute einer Klasse meist Korrelationen zueinander aufweisen, erreicht er bei der Klassifizierung für Attribute mit moderater Korrelation zueinander oft sehr gute Ergebnisse [181, 140].

Hidden Markov Model (HMM) Das *Hidden Markov Model* ist ein stochastisches Modell, welches auf *Markov Ketten* beruht und einen Spezialfall der dynamischen *Bayesschen Netze* darstellt [67, 132, 55]. Es basiert zwar auf Markov-Ketten, aber im Gegensatz zu diesen sind die Wahrscheinlichkeiten, welche den zufälligen Übergang von Systemzuständen beschreiben, nicht sichtbar. Lediglich die Ausgabewerte sind, abhängig vom Zustand, als sogenannte Emissionen zu beobachten. *Hidden Markov*-Modelle finden vor allem im *Reinforcement Learning* sowie bei der Untersuchung von Zeitreihen Anwendung, beispielsweise zur Sprachsynthese und bei der Erkennung von Gesten oder von Handschrift [181].

Support Vector Machine (SVM) Die *Support Vector Machine* ist ein Klassifikations- und Regressionsalgorithmus [79]. Dabei wird jede Instanz aus einem bereitgestellten Satz von Trainingsdaten einer von zwei verfügbaren Kategorien zugeordnet. Daraus wird anschließend ein Modell trainiert, welches auch neue Instanzen genau einer Kategorie zuweist. Es handelt sich also um einen nicht probabilistischen, binären und linearen Klassifikator. Durch den sogenannten Kernel-Trick, bei welchem innere Produkte von Features zueinander statt deren Koordinaten im gesamten Merkmalsraum berechnet werden, ist es möglich, den SVM-Algorithmus auch nicht-linear auf hochdimensionale Daten anzuwenden [187]. Die Anlehnung an den Namen Support Vector geht dabei auf das Vorgehen zurück, welches Klassen von Instanzen mithilfe einer Hyperebene voneinander trennt. In einem trainierten Modell sind dabei alle Instanzen X_i , welche einem n -dimensionalen Vektor entsprechen, einer Klasse welche von einer $(n - 1)$ -dimensionalen Hyperebene von anderen Klassen getrennt ist, zugeordnet. Dabei sind viele Möglichkeiten zum Ansetzen der Hyperebene möglich, es soll jedoch die als Grenze gewählt

werden, welche die vorhandenen Klassen als Separator mit dem größten möglichen Abstand zueinander trennt. Die Hyperebene wird also so gewählt, dass die Distanz zu den nächsten Instanzen auf beiden Seiten (zu beiden Klassen) maximiert wird. Wird eine solche Hyperebene gefunden, so nennt man diese *Maximum-Margin-Hyperplane* oder auch *Perceptron of optimal Stability*. Über Erweiterungen kann eine SVM auch auf nicht gelabelte Daten im Bereich des nicht überwachten Lernens angewendet werden, man spricht dann vom sogenannten *Support Vector Clustering* [181, 196].

K-Star (K*) Der *K-Star*-Algorithmus ermöglicht das Auffinden der k kürzesten Pfade zwischen einem festgelegten Paar von Scheitelpunkten in einem gerichteten und gewichteten Graphen. Dabei implementiert der K* zwei Vorteile gegenüber anderen k -kürzester-Weg-Algorithmen. Zum einen verarbeitet er Graphen *on-the-fly*, was bedeutet dass nicht der gesamte Graph aufgebaut werden muss, sondern nur die gerade zur Verarbeitung benötigte Teile. Zum anderen kann der K* durch den Einbezug von heuristischen Funktionen optimiert werden, was sich in erheblichen Laufzeitverbesserungen niederschlagen kann [5].

k-Nearest-Neighbour (kNN) Bei dem *k-Nearest-Neighbour*-Algorithmus handelt es sich um eine nicht parametrisierte Methode zur Klassifizierung und Regression. Im Falle der Klassifizierung ist das Ergebnis des kNN die Zugehörigkeit einer gegebenen Instanz zu einer bestimmten Klasse, welche über ein Mehrheitsvotum seiner benachbarten Instanzen bestimmt wird. Die zuzuordnende Instanz wird dabei der Klasse zugewiesen, welcher ihre k -nächsten Nachbarn angehören. Im Falle von $k = 1$ wird das Objekt der Klasse des einzigen nächsten Nachbarn zugeordnet [97].

2.3.1.2 Nicht überwachttes maschinelles Lernen

Hauptziel des nicht überwachten maschinellen Lernens ist das Verstehen von nicht gelabelten Daten durch Suchen von versteckten oder nicht sofort sichtbaren Strukturen und Mustern, beispielsweise im Rahmen einer Cluster-Analyse. Anwendungen finden sich bei der Identifikation von gleichen Objekten oder Mustern in der Bildverarbeitung, in der Erkennung von Anomalien sowie der Komprimierung von Daten durch Cluster-Bildung oder via Principal-Component-Analysis (vgl. Abschnitt 2.3.3.2) [208]. Die Eingangsinformation eines nicht überwachten Lernalgorithmus ist ebenso wie bei überwachtem Lernen eine Menge von Merkmalsvektoren $X_i = (x_i^1, \dots, x_i^n)$. Diese ist im Gegensatz dazu aber keinem Label y_i zugeordnet. Dem überwachten Lernalgorithmus ist also unbekannt, welcher Klasse eine Eingangsinstanz angehört. Im Folgenden werden einige relevante Algorithmen des nicht überwachten maschinellen Lernens vorgestellt [76].

k-Means Der *k-Means* ist ein Clustering-Algorithmus, welcher auf Ba-

sis von Vektorquantisierung arbeitet, indem er n Beobachtungen eines Ereignisses in k Cluster unterteilt. Die Menge der Cluster k muss zuvor festgelegt werden, der Algorithmus bevorzugt bei der Suche von Cluster-Zentren die Bildung von Gruppen mit geringer Varianz und möglichst ähnlicher Größe. In seiner Standardimplementierung mit einer gegebenen Anzahl von k zufälligen Cluster-Zentren (Means) $m_1^{(1)}, \dots, m_k^{(1)}$ durchläuft der k -Means Algorithmus abwechselnd zwei Verarbeitungsschritte, 1) den Zuweisungsschritt und 2) den Update-Schritt. In 1) wird jede gegebene Beobachtung dem Cluster zugewiesen, dessen Cluster-Varianz durch Zuordnung der neuen Beobachtung am geringsten erhöht wird, anschließend wird in 2) für jeden Cluster der aktuelle Mittelwert berechnet und ein neuer Centroid bestimmt. Diese zwei Schritte werden so lange sequentiell durchgeführt, bis sich die Zuordnungen zu den unterschiedlichen Clustern nicht mehr ändern [123, 202].

DBSCAN Bei dem *Density-based spatial clustering with noise* (DBSCAN) handelt es sich um einen dichte-basierten Clustering-Algorithmus. Bei einem gegebenen Set von Punkten gruppiert der DBSCAN Punkte, welche dicht beieinander liegen und markiert Punkte, welche in Regionen mit geringerer Dichte positioniert sind als Ausreißer. Sukzessive werden alle im Datensatz vorhandenen Punkte entweder als Kernpunkte, als dichte-verbundene (in ϵ -Reichweite zu den Kernpunkten) liegende Punkte oder als Ausreißer klassifiziert. Dabei gilt: Punkt p fungiert dann als Kernpunkt, wenn mindestens $minPts$ Punkte innerhalb seiner ϵ -Nachbarschaft $N_\epsilon(p)$ gefunden werden können. Ein Punkt q ist dichte-verbunden zu p , wenn er innerhalb der Distanz ϵ liegt und p als Kernpunkt fungiert. Alle Punkte, welche nicht von einem anderen Punkt aus erreichbar sind gelten als Ausreißer, wobei Punkt q von p aus erreichbar ist, wenn es einen Pfad p_1, \dots, p_n mit $p_1 = p$ und $p_n = q$ gibt und jedes p_{i+1} direkt von p_i aus erreichbar ist. Dabei müssen alle Punkte auf einem Pfad mit Ausnahme von q Kernpunkte sein [58].

OPTICS Der *Ordering Points To Identify Clustering Structure* (OPTICS) ist ein dichte-basierter Algorithmus zur Cluster-Analyse, dessen Grundidee der von DBSCAN ähnelt. Im Gegensatz zu diesem adressiert er aber auch eines von dessen Hauptproblemen, nämlich die Erkennung von bedeutsamen Clustern in Datensätzen mit variabler Auflösung. Dazu werden Punkte im gegebenen Datensatz solchermaßen geordnet, dass sie bei räumlicher Nähe auch in ihrer Reihenfolge nahe zueinander sortiert sind. Als Initialparameter benötigt OPTICS zum einen die zu berücksichtigende maximale Entfernung ϵ zweier Punkte, zum anderen die minimale Anzahl an Punkten $minPts$, welche zur Bildung eines Clusters notwendig ist. Ein Punkt p fungiert dann als Kernpunkt eines Clusters, wenn mindestens $minPts$ Punkte innerhalb seiner ϵ -Nachbarschaft $N_\epsilon(p)$ gefunden werden können [7].

2.3.1.3 Reinforcement Learning

Beim *Reinforcement Learning* handelt es sich um eine Disziplin des maschinellen Lernens, bei welcher Software-Agenten in einer ihnen bereitgestellten Umgebung selbstständig Maßnahmen ergreifen, um Belohnungen, welche aus der Ausführung der von ihnen zur Verfügung gestellten Aktionen resultieren, kumulativ zu maximieren. Diese Belohnungen können auch negativer Natur sein und sind als direktes Feedback auf realisierte Aktionen zu sehen. In verwandten Arbeiten wird das *Reinforcement Learning* in diesem Kontext auch als approximative dynamische Programmierung oder als neurodynamische Programmierung bezeichnet [16].

Grundlegend wird *Reinforcement Learning* dabei als *Markov Entscheidungsprozess* modelliert, welcher eine Menge aus möglichen Status S für die Umwelt und den Agenten und eine Menge von möglichen Aktionen A enthält. Die Wahrscheinlichkeit P_a des Übergangs von Status s in s' bei Ausführung der Aktion a , welche mit der Wahrscheinlichkeit des Erhaltens einer Belohnung Pr gleichzusetzen ist, wird dabei wie folgt definiert:

$$P_a(s, s') = Pr(s_{t+1} | s_t = s, a_t = a) \quad (2.10)$$

Die sofortige Belohnung, welche nach Übergang von s in s' durch Aktion a erfolgt, wird mit $R_a(s, s')$ angegeben. Zusätzlich werden meist stochastische Regeln benötigt, welche die Beobachtungen, typischerweise basierend auf zuvor erhaltenen Belohnungen, eines Agenten beschreiben.

Während des Lernprozesses agieren die Agenten mit ihrer Umwelt in diskreten Zeitschritten, wobei sie zu jedem Zeitpunkt t eine Beobachtung o_t , welche auch die Belohnung r_t enthält, bekommen. Auf Basis dieser Beobachtungen wird sukzessive die nächste Aktion a_t aus der diskreten Menge möglicher Aktionen ausgewählt und ausgeführt. Das System schreitet nun zum nächsten Zeitschritt s_{t+1} voran und die diesem zugeordnete Belohnung r_{t+1} , welche mit dem Übergang (s_t, a_t, s_{t+1}) assoziiert ist, wird ermittelt. Grundziel des Agenten ist dabei, die erhaltene kumulative Belohnung zu maximieren, wodurch optimale Verhaltensmuster für hochdimensionale oder komplexe Probleme erlernt werden sollen [27, 198, 17].

2.3.1.4 Künstliche Neuronale Netze

Künstliche Neuronale Netze bestehen aus künstlichen Neuronen und sind von der biologischen Hirnstruktur und deren neuronaler Netzwerkarchitektur inspiriert [78]. Da diese je nach Konfiguration und unterliegender Netzarchitektur sowohl für überwachte als auch für nicht überwachte Lernprozesse sowie in kombinierten Konzepten des *Reinforcement Learnings* zum Einsatz kommen können, werden sie hier gesondert aufgeführt.

Bei den verwendeten, künstlichen Neuronen handelt es sich um miteinander verbundene Knoten, welche über den menschlichen Synapsen entsprechen-

de Verbindungen Signale in Form von reellen Zahlen zu anderen Knoten zur Verarbeitung und Weiterleitung senden können [78]. In üblichen Implementierungen wird die Ausgangsinformation jedes Neurons über eine nichtlineare Funktion der Summe seiner Eingangsinformationen berechnet.

Der Lernfortschritt eines Netzes wird über die Gewichtung $w_{i,j}$ der Verbindungen zwischen den Neuronen während der Trainingsphase eines Modells angepasst. Diese Gewichtung bestimmt, ob die Stärke eines Signals zwischen zwei Neuronen i und j erhöht oder gedämpft wird. Zusätzlich kann mit der Implementierung von Schwellwerten das Weiterleiten von Signalen ganz verhindert werden. Innerhalb der Netzarchitektur sind die Neuronen dabei in Schichten organisiert, welche an ihren Eingängen verschiedene Arten von Signaltransformationen realisieren können. Je nach Implementierung können einzelne Schichten auch mehrfach von einem Signal durchlaufen werden. Während der Trainingsphase eines Netzes können neue Verbindungen zwischen Neuronen entwickelt, existierende Verbindungen gelöscht oder die Gewichte $w_{i,j}$ einzelner Verbindungen zwischen Neuron i zu Neuron j modifiziert werden. Auch ganze Neuronen können gelöscht oder der Schwellwert θ_j für die Aktivierung gesetzt werden. Zusätzlich können hier auch Modifikationen an der Aktivierungs-, der Propagierungs- und der Ausgabefunktion vorgenommen werden [78].

Ein künstliches Neuron j , das die Eingangsinformation $p_j(t)$ von seinen Vorgängerneuronen empfängt, besteht im Allgemeinen aus folgenden Komponenten: 1) einer Aktivierung $a_j(t)$, welche vom diskreten Zeitparameter t abhängig ist, 2) falls gesetzt einem Schwellwert θ_j welcher nur von der Lernfunktion verändert werden kann, 3) einer Aktivierungsfunktion f welche die neue Aktivierung des Neurons zu gegebener Zeit $t + 1$ aus der Aktivierung $a_j(t)$, dem Schwellwert θ_j und der Eingangsinformation des Netzes $p_j(t)$ berechnet. Die Ausgangsinformation $o_j(t)$ aus der Aktivierung wird von 4) der Ausgangsfunktion $f_{out}(a_j(t)) = o_j(t)$ berechnet [215].

2.3.1.5 Template-basierte Klassifizierung

Bei der Template-basierten Klassifizierung handelt es sich ebenfalls um eine Form des überwachten Lernens, da hierbei Musterschablonen, so genannte Templates, welche eindeutig einer bestimmten Instanzklasse zugeordnet sind, zum Einsatz kommen. Es wird also nicht auf ein mithilfe eines Lernalgorithmus trainiertes Vorhersagemodell zurückgegriffen. Die dazu benötigten Templates werden beispielsweise auf Basis von vorverarbeiteten Zeitreihen erstellt und zum späteren Vergleich über ein Distanzmaß abgelegt. Dabei können für eine Klasse mehrere Templates mit verschiedenen Ausprägungen angenommen und anschließend sequentiell abgeglichen werden. Andere Ansätze erstellen gemittelte Templates über den Median oder das arithmetische Mittel unterschiedlicher Ausprägungen und nutzen diese für die anschließende Klassifizierung. Letzteres ist auf Grund der reduzierten Anzahl an abzugleichenden Templates deutlich Ressourcen-schonender, Grenzfälle können aber gegebenenfalls schlechter ermittelt werden. Mögliche Distanzmaße und Kostenfunk-

tionen zum Vergleich klassischer Zeitreihen sind in Kapitel 2.2.6 aufgeführt, andere Ansätze verfolgen beispielsweise die Alphabetisierung von Zeitreihen mit anschließendem Vergleich zueinander über Metriken für Zeichenketten, wie der Levensthein-Distanz [214]. Nachteile Template-basierter Methoden ergeben sich, je nach eingesetzter Vergleichsmetrik, aus der Notwendigkeit zur gleichen Skalierung von Templates und zu klassifizierenden Daten. Weiterhin ist das Verfahren unflexibel bei der Integration neuer Entitäten und dem Einsatz von klassischen Distanzmetriken auf Zeitreihen. Aus diesem folgt ein oft Ressourcen-intensiver Klassifizierungsprozess und für jede neue, in eine Klassifizierungskette integrierte Entität steigt der benötigte Rechenaufwand linear an, wodurch das Template-basierte Klassifizieren vor allem in hochdimensionalen Datensätzen mit vielen unterschiedlichen Klassen selten praktikabel ist [150, 20, 50, 49].

2.3.2 Datensätze zum maschinellen Lernen

Eines der wichtigsten Ziele des maschinellen Lernens ist es, unterschiedliche Entitäten und deren Ausprägungen mit möglichst generischen Aussagen zu beschreiben. Die Qualität von trainierten Vorhersagemodellen kann dabei über ihre Fähigkeit zur Generalisierung von Mustern in Daten, auf welchen sie nicht trainiert wurden, evaluiert werden [158, 174]. Ein wichtiger Aspekt ist hierbei auch das Zusammenspiel von Bias und Varianz im Kontext des Overfittings, welches im folgenden Abschnitt behandelt wird.

Um diese Thematik zu adressieren, wurden im Rahmen verwandter Arbeiten bereits verschiedene Vorgehen entwickelt, nach welchen Datensätze in Test- und Trainingsdaten unterteilt werden. Diese beschäftigen sich beispielsweise mit der direkten Teilung von Instanzen nach vordefinierten Ratios, der Anwendung von k -fach gefalteten Kreuzvalidierungsverfahren oder auch der Leave-one-out Cross-Validation (LOOCV). Ergebnis dieser verschiedenen Methoden zur Unterteilung von Datensätzen sind immer ein Trainingsdatensatz und ein Testdatensatz. Der Trainingsdatensatz wird anschließend dazu benutzt, um einen Klassifikator zu trainieren, wobei dessen unterliegendes Vorhersagemodell an die individuellen Parameter des Datensatzes, beispielsweise über variable Gewichtung, angepasst wird. Der Testdatensatz ist je nach Methode der Teilung als unabhängig vom Trainingsdatensatz zu betrachten, er folgt aber ebenso dessen Wahrscheinlichkeitsverteilung.

Eine gängige Methode zur Teilung ist das direkte Aufteilen in Test- und Trainingsdatensatz, wobei meist ein Ratio von 70% – 80% des Datensatzes als Trainingsdaten und 20% – 30% als Testdaten verwendet werden [158, 174]. Eine Alternative zum direkten Aufteilen nach vordefinierten Ratios ist die k -gefaltete Kreuzvalidierung. Dabei wird ein Datensatz zufällig in k -Teile unterteilt, welche den Faltungen der Kreuzvalidierung entsprechen. Anschließend wird genau k mal ein Vorhersagemodell trainiert, wobei für jedes ein anderer individueller Datenanteil $\frac{1}{k}$ als Testdatensatz beiseite gelegt wird und die

restlichen $k - 1$ Anteile den Trainingsdatensatz bilden. Einen speziellen Fall der k -gefalteten Kreuzvalidierung bildet die LOOCV, bei welcher der Datensatz nicht zufällig in k Anteile unterteilt wird, sondern k beispielsweise der Anzahl unterschiedlicher Teilnehmer einer Studie entspricht. Der Datensatz wird also in die Datenanteile seiner einzelnen Probanden unterteilt, anschließend wird auf Basis dieser Anteile im Vorgehen der Kreuzvalidierung getestet und validiert. So kann das Verhalten eines trainierten Modells gegenüber Ereignisinstanzen getestet werden, welche unabhängig von den zum Training verwendeten Entitäten sind [106, 142, 33].

2.3.2.1 Generalisierungsfehler und Overfitting

Overfitting und Generalisierungsfehler beschreiben die Tendenz von Modellierungsmethoden, sich bei Modellen an Trainingsdaten zu stark anzupassen und während eines Lernprozesses auch in Daten enthaltene Fehler oder unabhängige, für eine Klasse nicht relevante Merkmale als für deren Klassifikation relevant zu interpretieren. So ist es möglich, dass Klassifikatoren einen im Trainingsdatensatz enthaltenen Trend nicht erkennen, sondern diesen inklusive möglichem Rauschen oder Fehlern auswendig lernen und anschließend nur Daten dieses Datensatzes mit einer hohen Genauigkeit klassifizieren können. Eine Generalisierung und Anwendung auf andere, ähnliche Instanzen hingegen ist durch eine zu gute Anpassung an den Trainingsdatensatz nicht mehr gegeben [41, 195, 96].

Zum besseren Verständnis lassen sich Generalisierungsfehler in die zwei Unterkomponenten Bias und Varianz zerlegen. Bias beschreibt dabei die Tendenz eines Lernalgorithmus, Fehler zu lernen und das unterliegende Modell auf diese auszurichten, da angenommen wird, dass diese eine expressive Beschreibung einer Instanz repräsentieren. Varianz beschreibt in diesem Kontext die Eigenschaft des Lernens von Rauschen und Zufallsmerkmalen, während expressive und wichtige Features ignoriert oder nur unzureichend berücksichtigt werden. Lineare Klassifikatoren neigen dazu einen hohen Bias aufzuweisen, da sie nicht fähig, sind eine Grenze zwischen zwei Instanz-Klassen im Merkmalsraum zu ziehen, wenn es sich bei dieser nicht um eine Hyperebene handelt. Algorithmen basierend auf Entscheidungsbäumen hingegen, wie beispielsweise Random Forest oder C4.5, zeigen sich eher robust gegenüber Bias-bedingten Fehlern, da sie jede Art von Boolescher Funktion implementieren können. Dafür sind sie aber anfällig gegenüber Varianz-bedingten Fehlern. Dadurch können Entscheidungsbaummodelle sehr unterschiedlich modelliert sein, obwohl sie auf unterschiedlichen Testdatensätzen des gleichen Ursprungs trainiert wurden und die resultierenden Entscheidungsmodelle eigentlich ähnlich sein sollten [44, 43].

Zur Vorbeugung gegen oben beschriebene Fehler gibt es verschiedene Maßnahmen. So muss eine ausreichende Anzahl von Trainingsdaten, welche möglichst viele Instanzen einer Klasse in unterschiedlichsten Ausprägungen enthält, vorhanden sein. Im Kontext der Analyse menschlicher Bewegung bedeutet dies beispielsweise, dass Informationen über ein Bewegungsereignis von möglichst

vielen, unterschiedlichen Probanden erfasst werden sollten. Dabei ist auch auf eine hohe Varianz bezüglich Alter, Geschlecht, Mobilität, u.a. in den Daten zu achten, denn nur so kann anschließend ein generisches Vorhersagemodell trainiert werden [51, 44]. Eine weitere Möglichkeit ist die Überprüfung eines Modells mithilfe einer Kreuzvalidierung (siehe auch Kapitel 2.3.2), welche es ermöglicht, die optimale Größe eines Entscheidungsbaums auszumachen [142]. Im Bereich der neuronalen Netze gibt es Methoden wie den Dropout von einzelnen Knotenpunkten zusammen mit ihren Verbindungen zu anderen Knoten, um so eine Überanpassung an ein vorgegebenes Datenset zu verhindern [195].

2.3.2.2 Class-Imbalancement

Eine weitere wichtige Voraussetzung zur Verhinderung von Overfitting und Generalisierungsfehlern (vgl. 2.3.2.1) ist die Sensibilität bei dem Design von Validierungsverfahren gegenüber nicht ausbalancierten Klassenverteilungen. Diese beschreiben das Problem von nicht ausbalancierten Anteilen unterschiedlicher Aktivitätsklassen innerhalb eines Datensatzes. So kommen Aktivitätsklassen wie beispielsweise Schlafen oder Arbeiten in einem Datensatz mit täglichen Aktivitäten einzelner Personen ggf. vergleichsweise öfter und länger vor, als das Trinken von Kaffee [19]. Klassifiziert ein Modell im gleichen Kontext nun neue, unbekannte Aktivitäten als zugehörig zur Klasse Arbeiten oder Schlafen, so ist dieses Ergebnis mit deutlich höherer Wahrscheinlichkeit zutreffend. Um solche Probleme zu umgehen, sollen Klassen in Datensätzen zum maschinellen Lernen möglichst gleich verteilt sein. Dies kann zum einen durch vermehrtes Erfassen von Instanzen unterrepräsentierter Ereignisse realisiert werden, andere Ansätze beschäftigen sich mit dem Generieren synthetischer Trainingsdaten oder dem Oversampling vorhandener Informationen [26]. Generell sollten Klassen innerhalb eines Datensatzes zum maschinellen Lernen im Optimalfall mit einer Gleichverteilung repräsentiert sein.

2.3.3 Feature-Engineering

Die Anzahl und Auswahl von expressiven Features ist bei Untersuchung von Ereignissen im Kontext des maschinellen Lernens von essentieller Bedeutung und hat sowohl direkte Auswirkungen auf den Erfolg einer Klassifizierung als auch auf den dazu notwendigen Berechnungsaufwand. Generell gilt, je höher die Länge eines Merkmalsvektors und damit die Dimensionalität des Merkmalsraums, desto mehr Zeit wird benötigt, um die Parameter für ein Vorhersagemodell während der Trainingsphase zu bestimmen und zu konfigurieren. Der Begriff expressiver Features beschreibt dabei Merkmale eines Ereignisses, welche einen besonders hohen Informationsgehalt bezüglich dessen individueller Gestalt beinhalten.

2.3.3.1 Kategorisierung von Features zur Beschreibung von Bewegung

Im Kontext verschiedener Arbeiten zur Untersuchung von menschlicher Aktivität auf Basis von Sensorinformationen lassen sich verschiedene Kategorien von genutzten Features identifizieren. Diese Kategorien ermöglichen eine Übersicht über verschiedene Möglichkeiten des Feature-Engineerings und bestimmen das Vorgehen zur Extraktion von Features aus einem Datensatz und somit den Aufbau des resultierenden Merkmalsvektors maßgeblich. Eine Aufstellung von Bulling et al. unterscheidet dabei zwischen den folgenden Feature-Gruppen [24].

Signal-basierte Features Diese basieren meist auf statistischen Größen wie Mittelwert, Varianz, Standardabweichung oder Kurtosis einer Zeitreihe. Dabei sind sie vor allem auf Grund ihrer Einfachheit und der hohen Performanz bei der Lösung verschiedener Probleme der Aktivitätserkennung beliebt [88, 143, 178]. Aber auch spektrale und Energie-basierte Features, auf deren Gewinnung bereits in Kapitel 2.2.4 eingegangen wird, zählen zur Gruppe der Signal-basierten Features.

Körpermodell-basierte Features Auf Körpermodellen basierende Features werden unter Zuhilfenahme von approximierten oder nachgebildeten 3D-Modellen des menschlichen Skeletts errechnet. Die Bewegungen eines Probanden werden dabei mithilfe von Beschleunigungssensorik, Tiefensensorik oder visueller Sensorik (vgl. Kapitel 3.3) auf ein 3D-Modell übertragen, welches dann die Rekonstruktion von Bewegung und damit die Extraktion von Features ermöglicht. Ein Beispiel ist das Nachzeichnen der Bewegungstrajektorien von Extremitäten über deren zeitliche Beschleunigungs- und Rotationsinformationen. Daraus können diskrete Features, wie beispielsweise die auftretende Steigung oder das Gefälle zwischen variierenden Positionen unterschiedlicher Körperteile, zurückgelegte Strecken oder geometrische Kennzahlen zwischen einzelnen Körperteilen, extrahiert werden [148, 217, 19].

Ereignis-basierte Features Diese Art von Features wird auf Basis von zuvor bereits mit anderen Merkmalsvektoren erkannten Teilereignissen konstruiert. So zum Beispiel durch Nutzung der wiederholten Bewegung des menschlichen Augapfels zum Erkennen komplexerer Ereignisinstanzen wie Unaufmerksamkeit oder der Wiedererkennung von Objekten [25].

Multilevel-Features Bei diesem Ansatz wird der zu untersuchende Datensatz zunächst mit einem Clustering-Verfahren bearbeitet, beispielsweise mit dem *Estimation Maximization*-Algorithmus oder dem darauf basierenden *k-Means*-Algorithmus (vgl. 2.3.1.2), mit welchem einzelne Cluster innerhalb eines untersuchten Fensters gebildet werden [138]. Enthaltene temporale Informationen, Frequenzen und Korrelationen von Werten ermöglichen anschließend die Bildung expressiver Merkmalsvektoren [87, 18, 216].

2.3.3.2 Hauptkomponentenanalyse

Die Hauptkomponentenanalyse (engl. *Principal Component Analysis*, (PCA)) ist ein Vorgehen, mit dem versucht wird, signifikante Muster und Features in gegebenen Eingangsinformationen zu identifizieren, um so beispielsweise einen Merkmalsvektor für eine spätere Klassifizierung aussagekräftiger zu gestalten. Dabei kommen meist Werkzeuge des nicht überwachten maschinellen Lernens, aber auch die Bildung von Mittelwerten und Varianzen zum Einsatz. Hauptmotivation der PCA ist dabei zum einen, die Dimensionalität eines Merkmalsvektors durch Pruning zu reduzieren und so Rauschen oder unwichtige Features zu entfernen, zum anderen aber auch die Aussagekraft des ursprünglichen Merkmalsvektors so weit wie möglich zu erhalten. Es handelt sich also um einen Ansatz, welcher die Aussage umfangreicher Datensätze mit einer Vielzahl statistischer Variablen durch wenige Hauptkomponenten näherungsweise ausdrücken soll. Dazu werden multidimensionale Daten in einen niederdimensionalen Raum projiziert. So können beispielsweise Klassifizierungsergebnisse verbessert und Laufzeiten sowie der Ressourcenverbrauch einer anschließenden Klassifikation verringert werden [78, 208, 92]. Weitere Ansätze zur Reduktion von Dimensionalität sind auch die *Singular Value Composition* oder das *Sammon's Mapping* [203, 39].

2.3.3.3 Ranking und Auswahl von Features

Im Unterschied zur in Kapitel 2.3.3.2 beschriebenen PCA behandeln das Ranking und die Auswahl von Features nicht die Reduzierung von Dimensionalität. Beide Ansätze haben zwar zum Ziel, einen Merkmalsvektor auf expressive Features zu reduzieren, aber bei der Auswahl von Features werden nur solche in einen Merkmalsvektor übernommen, die auch vorher schon in einem Datensatz vorhanden waren. Im Rahmen der PCA hingegen können einem Merkmalsvektor auch neue Kombinationen und Projektionen von Attributen hinzugefügt werden. Ziel von Ranking und Auswahl ist aber ebenso das Erreichen einer besseren Vorhersagegenauigkeit durch Kompression von Information und Verwendung von Feature-Subsets. Dazu werden unnötige, redundante oder solche Features mit Störinformationen identifiziert und aussortiert [72, 149, 193]. Dies kann beispielsweise mit einem der folgenden drei Schemata realisiert werden: dem *Filter-Schema*, dem *Wrapper-Schema* und dem *eingebetteten Schema*.

Das **Filter-basierte Schema** zur Feature-Auswahl basiert auf der Bewertung einzelner Features nach statistischer Methodik oder Korrelation (z.B. via Korrelationskoeffizient nach *Pearson*). Je nach punktueller Bewertung werden einzelne Features geordnet und zur weiteren Einbehaltung im Datensatz ausgewählt oder daraus entfernt. Diese Art von Schemata sind meist univariat und behandeln einzelne Features unabhängig voneinander oder beziehen abhängige Variablen in den Bewertungsprozess mit ein [75, 136, 70].

Das **Wrapper-Schema** behandelt die Auswahl einer Menge von Features als Suchproblem, wobei unterschiedliche Feature-Kombinationen erstellt, evaluiert und miteinander verglichen werden. Durch Training verschiedener Vorhersagemodelle mit unterschiedlichen Feature-Kombinationen und deren anschließender Evaluation kann jeweils ein Genauigkeits-Score zugewiesen und zum Ranking einzelner Features genutzt werden. Der Prozess zur Suche verschiedener Feature-Kombinationen kann dabei methodisch (z.B. *best-first* Suche) oder stochastisch aufgebaut sein (z.B. *Random Hill Climbing*-Algorithmus) sowie auch heuristische Verfahren nutzen [75, 136, 107].

Das **eingebettete Schema** arbeitet lernbasiert und versucht Features, welche am meisten zur Vorhersagegenauigkeit eines Modells beitragen, während der Erstellung eines Modells zu identifizieren. Am weitesten verbreitet sind dabei Methoden der Regularisierung. Diese werden oft auch Penalty-Methoden genannt und integrieren zusätzliche Bedingungen in den Optimierungs- und Trainingsprozess eines Vorhersagemodells. So kann beispielsweise ein Bias zugunsten von Modellen mit geringerer Komplexität induziert und damit das Training eines Modells mit weniger Parametern forciert werden [36, 136].

2.3.4 Hyper-Parameter-Optimierung

Die *Hyper-Parameter-Optimierung* (HPO) beschreibt im maschinellen Lernen das Problem der Auswahl von optimalen Hyper-Parametern zur Feinjustierung eines Lernalgorithmus während der Modellbildung und der anschließenden Klassifizierungsphase. Der Begriff Hyperparameter umfasst dabei alle Initialparameter wie Gewichte, Distanzen, Arten von Distanzmaßen oder Lernraten, welche vor Einsatz eines Lernalgorithmus für diesen festgelegt werden müssen. Diese dienen als Stellschrauben zu dessen Anpassung an spezifische Anwendungszwecke [14].

Nicht alle Algorithmen benötigen dabei initiale Eingangsparameter, so beispielsweise der kNN (vgl. Kapitel 2.3.1.1) als nicht parametrisierter Klassifikator. Ein *Hidden Markov*-Modell hingegen benötigt die Matrizen A und B mit Informationen zu den Übergängen unterschiedlicher Zustände, den Ausgangswahrscheinlichkeiten jedes Zustands und den Parameter θ für die initiale Zustandswahrscheinlichkeit. Ziel der HPO ist es, das optimale Tupel an Hyperparametern zu finden, mit welchem sich die vordefinierte Verlustfunktion eines trainierten Modells für einen unabhängigen Validierungsdatensatz minimieren lässt. Dies kann mit verschiedenen Ansätzen realisiert werden [24].

Eine klassische und sehr aufwendige Variante ist zum Beispiel der *Grid Search*, bzw. das *Parameter Sweeping*, bei welchem verschiedene Parameter-tupel manuell kreiert und parallel getestet werden, um ein optimales Ergebnis zu erlangen. Zuvor müssen Parametergrenzen definiert werden, in deren

Rahmen alle verschiedenen Kombinationen systematisch zu evaluieren sind, bis alle Möglichkeiten durchlaufen wurden. Dabei handelt es sich um trivial parallele Optimierungsaufgaben, welche getrennt voneinander gleichzeitig abgearbeitet werden können. Allerdings ist dieser *Brute Force*-ähnliche Ansatz äußerst Ressourcen-intensiv und im hochdimensionalen Raum vom *Curse of Dimensionality* betroffen, weshalb verschiedene Alternativen dazu konzipiert wurden [14].

Eine Alternative ist der sogenannte *Random Search*, bei welchem manuell erstellte Kombinationen von Hyper-Parametern durch deren zufällige Auswahl ersetzt werden. Dabei ist dieses Vorgehen sowohl auf kontinuierliche als auch diskrete und hybride Räume anwendbar. Vor allem für eine geringere Anzahl an gesuchten Hyper-Parametern ist der *Random Search* dem *Grid Search* sowohl hinsichtlich der benötigten Zeit zu deren Identifikation als auch bezüglich der schlussendlichen Performanz des damit konfigurierten Lernalgorithmus oft überlegen [13]. In diesem Fall spricht man davon, dass das Optimierungsproblem eine geringe intrinsische Dimensionalität besitzt [205]. Auch beim *Random Search* handelt es sich um ein trivial paralleles Szenario; zusätzlich kann zur Suchoptimierung auch Vorwissen, wie zum Beispiel über die Verteilung eines Datensatzes, in die Suche miteinbezogen werden.

Weitere Methoden zur Identifikation von optimalen Hyper-Parametern sind das Konzept der *Bayesschen Optimierung*, die gradientenbasierte Optimierung und die Anwendung evolutionärer Algorithmen [119, 14, 86, 191, 35, 145].

2.4 Aktivitätserkennung

Die Aktivitätserkennung ist aktuell eines der bedeutendsten Forschungsfelder wenn es um die Entwicklung von Kontext-sensitiven Anwendungen und Applikationen im Umfeld des *Internet of Things* (IoT) sowie im Bereich *Smart Home* geht [192, 166]. Dabei lassen sich durch das Erkennen spezifischer Aktivitäten eines oder mehrerer Akteure nicht nur quantitative und qualitative Aussagen zur Aktivität selbst oder zur möglichen Beziehung der Akteure zueinander ableiten. Weitergehend wird versucht, auf Basis von historischen Informationen Prognosen hinsichtlich zukünftiger Aktivitäten zu treffen oder Trends vorherzusagen. Durch neue Möglichkeiten der sensorischen Erfassung von Rohdaten sowie der Notwendigkeit zur Analyse von riesigen Mengen an Nutzer- und Metadaten, erfahren Konzepte zur Extraktion und Formalisierung von Kontext mithilfe von Anwendungen des maschinellen Lernens und der *Big Data*-Analyse erhöhte Aufmerksamkeit. Typische Anwendungsgebiete sind der medizinische Bereich und die Altenpflege, Anwendungen aus dem Bereich des *Affective Computing*, die Analyse und Erfassung menschlicher Bewegung, die Untersuchung von Kaufverhalten und Nutzerströmen, Anwendungen aus dem Bereich der Sicherheit oder auch die Analyse und Steuerung von Interaktionen und Verhalten in sozialen Medien [211, 148, 45, 124, 30].

Nachfolgend werden verschiedene Konzepte und Begrifflichkeiten zur Unter-

suchung von Bewegungsaktivität als Teil der Aktivitätserkennung präsentiert. Zusätzlich wird eine generische Prozesskette zur Untersuchung von Bewegungsaktivität vorgestellt, welche als Grundlage für die im Rahmen dieser Arbeit entwickelten und angewendeten Konzepte dient.

2.4.1 Untersuchung von Bewegungsaktivität

Die Erfassung von menschlicher Bewegung ist die Grundvoraussetzung der Analyse menschlicher Bewegungsaktivität, welche als Teildisziplin der Aktivitätserkennung zu sehen ist. Generell lässt sich Bewegungsaktivität dabei aus einem Makro- und einem Mikro-Blickwinkel betrachten, welche beide auf zu meist unterschiedlichen technischen Grundlagen basieren. So lässt sich menschliche Aktivität, wie zum Beispiel die Art der Fortbewegung, aus einem Makro-Blickwinkel mithilfe von Satelliten-gestützter Sensorik, wie dem *Global Positioning System* (GPS), via Signalstärken, mithilfe akustischer Signale oder auch über das Funknetz eines mobilen Endgeräts extrahieren [125, 11]. Um menschliche Aktivität hingegen aus einem Mikro-Blickwinkel betrachten zu können, beispielsweise zur Erkennung und Untersuchung spezifischer Bewegungsabläufe des menschlichen Körpers, werden meist Tiefensensoren, visuelle Sensorik oder Bewegungssensoren in Form von Wearables genutzt [139, 127, 9]. Ein detaillierter Überblick zur diesbezüglich anwendbaren Sensorik und den damit verbundenen Verfahren zur Aktivitätserkennung wird im Rahmen der verwandten Arbeiten in Kapitel 3.3 gegeben.

Für diese Arbeit ist die Klasse einer Bewegungsaktivität dabei als übergeordnete Aktivitätskategorie definiert, welche aus einzelnen Bewegungsereignissen besteht. Dies ist beispielsweise die Bewegungsaktivität Laufen, welche von einer Sequenz aus unterliegenden individuellen Einzelereignissen, den einzelnen Schritten, definiert wird. Die feingranulare Untersuchung von individuellen Bewegungsereignissen innerhalb einer übergeordneten Bewegungsaktivität sowie deren automatisierte, qualitative Untersuchung bilden den Hauptfokus dieser Arbeit.

2.4.2 Prozesskette zur Untersuchung von Bewegungsaktivität

Eines der Hauptwerkzeuge, welche in den folgenden Kapiteln zur Analyse von menschlicher Bewegungsaktivität auf Basis von Bewegungsinformation eingesetzt wird, ist das überwachte maschinelle Lernen. Daran angelehnt wird eine generische Prozesskette zur Untersuchung von Bewegungsaktivität definiert, welche in den anschließenden Kapiteln als Grundlage für die im Rahmen dieser Arbeit entwickelten Analysekonzepte dient. Die nachfolgend beschriebenen Verarbeitungsschritte lassen sich dabei je nach Abstraktionslevel für unterschiedliche Arten von Eingangsinformationen nutzbar machen. Im vorliegenden Fall wird davon ausgegangen, dass es sich bei den Eingangssignalen um

Bewegungsinformationen in Form von multidimensionalen Zeitreihen handelt. Einen Überblick über die fünf Schritte dieser Prozesskette gibt Abbildung 2.3:



Abbildung 2.3: Überblick über eine generische Prozesskette zur Aktivitätserkennung mithilfe von Prozessen des maschinellen Lernens.

(1) die *Erfassung von Bewegungsinformation*, (2) deren *Vorverarbeitung*, (3) die *Segmentierung* in einzelne Ereignissegmente, (4) das *Feature-Engineering* und (5) die anschließende *Klassifizierung* zuvor segmentierter Ereignisse. Diese Schritte lassen sich in ähnlicher oder abgewandelter Form auch in verwandten Arbeiten finden, welche als Grundlage für die folgenden Definitionen dienen [150, 117, 8, 24, 103]. Ziel ist es dabei, die individuellen Ausprägungen einzelner Verarbeitungsschritte auf ein möglichst generisches, allgemein anwendbares Modell zur Untersuchung von Bewegungsaktivität zu projizieren.

2.4.2.1 Schritt 1: Erfassung von Bewegungsinformation

Die *Erfassung von Bewegungsinformation* ist der erste Schritt innerhalb der Prozesskette aus Abbildung 2.4. Er beschreibt die Erfassung von Bewegung,

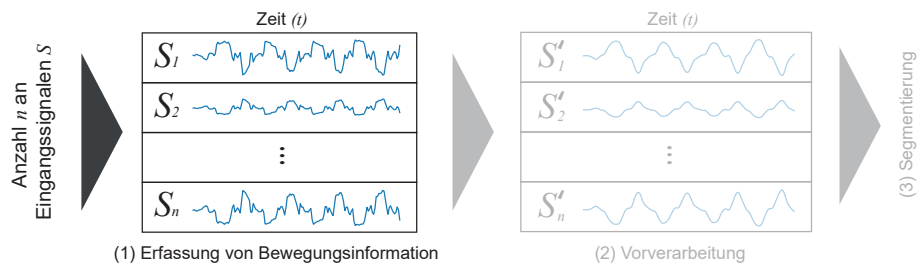


Abbildung 2.4: Schematische Darstellung von Schritt (1) *Erfassung von Bewegungsinformation* innerhalb der Prozesskette aus Abbildung 2.3. Die Eingangsinformationen setzen sich aus n Eingangssignalen zusammen.

welche typischerweise mit einer der in Kapitel 3.3 beschriebenen Sensorsysteme realisiert wird. Im Fall von Bewegungssensoren handelt es sich bei den eingehenden Signalen um multidimensionale Zeitreihen S , deren Anzahl n von der Menge der eingesetzten Sensoren und den zu erfassenden räumlichen Dimensionen abhängig ist. Bei inertialen Messeinheiten sind diese oft X -, Y -, und Z -Richtung der Rotation und der Beschleunigung eines Körpers. Ein beliebiges Einzelsignal S_v an Position v lässt sich somit als Zeitreihe der Länge k beschreiben, wobei k die Anzahl aller erfassten Samples zum jeweiligen

Zeitpunkt t eines individuellen Signals ist:

$$S_v = (s_v^{t_1}, s_v^{t_2}, \dots, s_v^{t_k}) \quad (2.11)$$

Aufgrund von variierenden Sampling-Raten (siehe Kapitel 2.4.2.2) einzelner Sensoren sowie möglicher Synchronisationsfehler können einzelne Zeitreihen in diesem ersten Schritt der Prozesskette nur getrennt voneinander betrachtet und nicht zusammengefasst als Matrix verarbeitet werden. Ausgabeinformation dieses Arbeitsschrittes, welche an Schritt (2) *Vorverarbeitung* weitergegeben wird, ist also eine Menge von n einzelnen, unbearbeiteten Signalen S , welche über die Zeitstempel t ihrer Erfassungszeit zueinander in Bezug gesetzt werden können.

2.4.2.2 Schritt 2: Vorverarbeitung

In Schritt (2) *Vorverarbeitung* der Prozesskette werden die in Schritt (1) erfassten, unbearbeiteten Signale bereinigt und gefiltert, um sie für eine anschließende Segmentierung und das Feature-Engineering vorzubereiten. Die Gründe hierfür sind vielfältig.

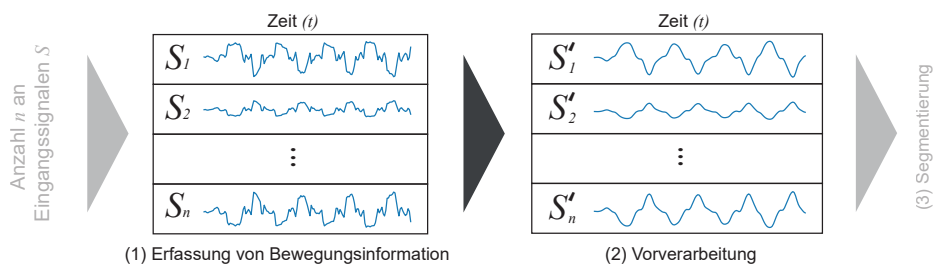


Abbildung 2.5: Überführung der im ersten Schritt erfassten Bewegungsinformation in einen im Rahmen von Schritt (2) *Vorverarbeitung* geglätteten und zur weiteren Untersuchung geeigneten Zustand.

Ein grundsätzliches Problem bei der Verwendung von Bewegungssensorik ist das sogenannte Basisrauschen, welches meist kleine, hochfrequent aufgezeichnete Messfehler in der Amplitude eines Signals über die Zeit beschreibt [190]. Ursprung dafür können beispielsweise verwendete Hardware und Materialien, Umwelteinflüsse wie Temperatur oder Vibration, aber auch Änderungen des Messverhaltens eines Sensors über einen Zeitraum (z.B. durch Materialalterung oder Aufwärmphasen) sein. Um durch Basisrauschen bedingte Verzerrung oder Fehlinterpretation von Signalen auszuschließen sowie unnötige Sample-Punkte zu entfernen und den damit verbundenen Verarbeitungsaufwand zu minimieren, können Rohsignale beispielsweise mit einem Tiefpass- oder Glättungsfilter bearbeitet werden [213] (vgl. Kapitel 2.2.1). Des Weiteren können auch Ausfälle, Messstörungen, variierende Sampling-Raten einzelner Sensoren oder unterschiedliche Sampling-Raten für individuelle Sensorplattformen eine Vorverarbeitung notwendig machen. Solche Störungen oder Varianzen können

z.B. mithilfe von polynominaler Interpolation im univariaten oder mit einer Nearest-Neighbour-basierten Interpolation im multivariaten Fall ausgeglichen werden (vgl. Abschnitt 2.2.5). Eine zusätzliche Fehlerquelle, welche die Weiterverarbeitung und Untersuchung von Signalen erheblich erschweren kann, sind Synchronisationsfehler. Diese entstehen beispielsweise durch die Verteilung von Sensoren und damit einhergehenden Latenzen in den notwendigen Kommunikationswegen oder durch das Fehlen einer Initialsynchronisation. Um dieser Problematik entgegenzuwirken, kann die Ausrichtung von Signalen zueinander Schwellwert-basiert oder anhand markanter Merkmale, z.B. zueinander in Bezug stehender Extrema oder Frequenzen, korrigiert werden [49]. Eine weitere Möglichkeit ist die nachträgliche Synchronisation von Informationen durch eine zentrale Recheneinheit während des Erfassungsvorgangs [51].

Abbildung 2.5 veranschaulicht die Überführung der im ersten Schritt erfassten, rohen Sensorinformationen in einen im Rahmen von Schritt (2) *Vorverarbeitung* geglätteten und aufbereiteten Zustand. Nach Glättung aller zu einem Ereignisset gehörenden Einzelsignale $S'_v = G(S_v)$ der Anzahl n und deren Interpolation auf die gleiche Länge t mit anschließender Synchronisation kann dieses Set S' nun in Matrizenform dargestellt werden:

$$S' = \begin{pmatrix} s'_1{}^{t_1} & s'_1{}^{t_2} & \dots & s'_1{}^{t_k} \\ \dots & \dots & \dots & \dots \\ s'_n{}^{t_1} & s'_n{}^{t_2} & \dots & s'_n{}^{t_k} \end{pmatrix} \quad (2.12)$$

Damit stellt S' zugleich die finale Ausgangsinformation aus Schritt (2) *Vorverarbeitung* dar und ist die Eingangsinformation von Schritt (3) *Segmentierung*. Anzumerken ist hierbei, dass keiner der beschriebenen Vorverarbeitungsschritte zwingend erforderlich ist. Vielmehr ergeben sich notwendige Vorverarbeitungsschritte aus den individuellen Anforderungen nachfolgender Verarbeitungsschritte.

2.4.2.3 Schritt 3: Segmentierung

Die Extraktion von individuellen Ereignissen aus kontinuierlichen, multidimensionalen Zeitreihen ist Aufgabe von Schritt (3) *Segmentierung*, welcher in Abbildung 2.6 verdeutlicht wird. Der Parameter i gibt dabei die Position eines Ereignisses e_i innerhalb einer kontinuierlichen Folge von Ereignissen an. Dabei gibt es zwei Hauptaufgaben, welche im Rahmen eines erfolgreichen Segmentierungsprozesses gelöst werden müssen. Diese sind einerseits die Detektion von einzelnen Bewegungsereignissen innerhalb einer Ereigniskette, andererseits deren Extraktion in individuelle Ereignissegmente.

Zur Detektion von Ereignissen innerhalb kontinuierlicher Zeitreihen können verschiedene Vorgehensweisen zum Einsatz gebracht werden, wobei der erste Schritt meist die Implementierung eines *Sliding Window*-Verfahrens ist (siehe Kapitel 2.2.3) [74, 50, 111, 201]. Dabei wird der Inhalt der einzelnen Fenster nach deren Bildung sequentiell auf beinhaltete Ereignisse oder Indikatoren

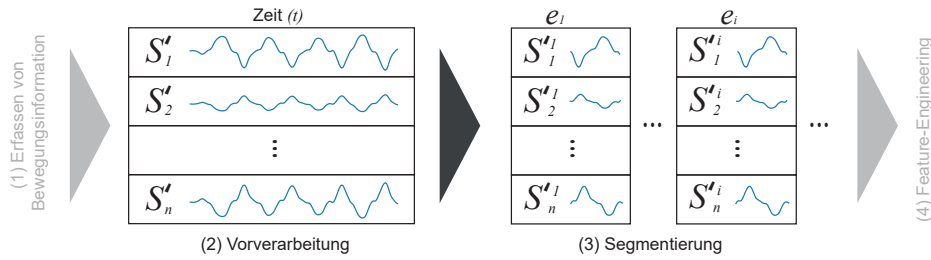


Abbildung 2.6: Schritt (3) *Segmentierung* befasst sich mit der Extraktion von individuellen Ereignissegmenten e_x aus dem im zweiten Schritt vorverarbeiteten Set S' von Zeitreihen.

für diese überprüft. Ein einfaches Beispiel kann mithilfe einer Schwellwertbasierten Prüfung auf Maxima innerhalb eines Signals realisiert werden. Sobald ein Maximum entdeckt wird werden alle Informationen um dieses herum in ein Ereignissegment e_i mit vordefinierter zeitlicher Dauer Δt von Ereignisbeginn t_s bis Ereignisende t_e extrahiert. Handelt es sich bei den Eingangsinformationen um multidimensionale Zeitreihen, so müssen für ein Segment alle Einzelsignale innerhalb des Intervalls Δt extrahiert und dem Segment hinzugefügt werden [50, 49].

Handelt es sich aber um Ereignisse mit generell geringer Signaldynamik oder von unbekannter Länge, so ist das zuvor genannte einfache Verfahren zur Detektion von Ereignissen nicht mehr anwendbar. Statt dessen können beispielsweise Energie-basierte Verfahren zum Einsatz kommen [117, 201, 71]. Diese basieren auf der Annahme, dass unterschiedliche Bewegungsaktivitäten mit unterschiedlicher Bewegungsintensität ausgeführt werden, wobei sich Intensitätsunterschiede direkt in unterschiedliche Energiestufen der Eingangssignale übersetzen lassen [24]. Das Energielevel ϵ eines Signals s_n lässt sich über das Integral $\epsilon = \int_{-\infty}^{\infty} |s_n(t)|^2 dt$ ermitteln. Mithilfe von Schwellwerten für ϵ kann das Segment anschließend einer spezifischen Aktivität zugeordnet werden.

Unabhängig von der Wahl des Verfahrens zur Ereignisdetektion ist die Ausgangsinformation des Schrittes (3) *Segmentierung* eine Menge E von einzelnen Ereignissen:

$$E = \{e_1, e_2, e_i, \dots\} \quad (2.13)$$

Dabei sind alle e_i innerhalb der Ereignismenge E in der Reihenfolge ihres ursprünglichen Eintretens geordnet und die Menge E oder einzelne Elemente e_i daraus bilden die Eingangsinformationen für den nächsten Schritt (4) *Feature-Engineering*. Jedes Ereignis e_i hat eine genau definierte Dauer $\Delta t = t_e - t_s$ innerhalb der Ereignisgrenzen t_s und t_e . Einzelereignisse können sich je nach Wahl des Detektions- und Extraktionsverfahrens sowie in Abhängigkeit von eventuell geforderten Sicherheitsabständen auch überschneiden.

2.4.2.4 Schritt 4: Feature-Engineering

Nach Durchführung der in der Abbildung 2.3 dargestellten Schritte 1-3, werden nun alle individuell extrahierten Bewegungsereignisse e_i innerhalb ihres jeweiligen Segments der Länge Δt an Schritt (4) *Feature-Engineering* übergeben. Je nach angewendetem Segmentierungsansatz kann Δt für jedes Einzelereignis eine statische oder auch eine variierende Segmentlänge umfassen. Hauptaufgabe des vierten Schrittes ist die Ermittlung von aussagekräftigen

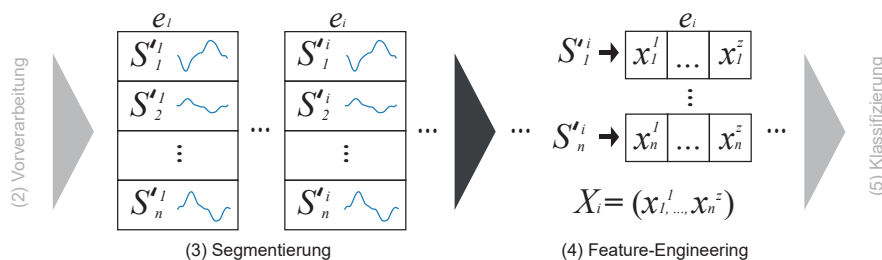


Abbildung 2.7: Schritt (4) *Feature-Engineering* verdeutlicht die Überführung der im dritten Schritt extrahierten Einzelsegmente e_i in jeweils einen korrespondierenden Merkmalsvektor X_i .

Merkmale, sogenannten Features, welche ein Ereignis innerhalb eines Segments über Bildung eines diskreten Merkmalsvektors so genau wie möglich beschreiben sollen. Dazu können beispielsweise physikalische Sensorinformationen wie Zeit, Rotation, oder Beschleunigung quantisiert und mithilfe einer Feature-Extraktionsfunktion $F(e_i)$ in diskrete Merkmale x überführt werden. Die Gesamtheit dieser Features bildet den Merkmalsraum, aus welchem für jedes Ereignis e_i der korrespondierende Merkmalsvektor X_i (vgl. Abbildung 2.7) konstruiert wird:

$$X_i = F(e_i) = (x^i_1, \dots, x^i_z) \quad (2.14)$$

In Darstellung 2.14 ist n die Gesamtanzahl der Eingangssignale und z die Anzahl aller Features je Signal. Einzelne Features x^i_z können hierbei automatisch berechnet oder auch manuell mithilfe von Expertenwissen bestimmt werden [153]. Wichtig ist, dass sie sich direkt auf spezifische temporale und physische Eigenschaften des Ereignisses e_i abbilden lassen und diese repräsentieren. Im Idealfall lassen sich ähnliche Ereignisse anhand ihrer Merkmalsvektoren auch im Merkmalsraum zu Clustern zusammenführen, während weniger ähnliche Ereignisse weiter voneinander entfernt sind.

Im Kontext der Untersuchung von Bewegungsaktivität müssen ausgewählte Features aber nicht nur aussagekräftig sein, sondern auch robust gegenüber einer heterogenen Menge an Probanden oder gegenüber Varianzen innerhalb verschiedener Ereignisklassen. Dazu kann die Aussagekraft und Robustheit eines Merkmalsvektors mithilfe von Pruning und selektiver Feature-Auswahl (vgl. Kapitel 2.3.3.1), durch den Einsatz der PCA (vgl. Kapitel 2.3.3.2) und durch Verhinderung von Generalisierungsfehlern (vgl. Kapitel 2.3.2.1) erheb-

lich gesteigert werden. Zur Verbesserung von Vorhersageergebnissen während des Klassifizierungsschrittes ist die sorgfältige Selektion und Reduktion von Features jedoch nicht von großer Relevanz [149, 72]. Je höher die Dimensionalität des aufgespannten Merkmalsraums ist, desto mehr Daten werden zum Training eines Klassifizierungsmodells benötigt und desto Ressourcenintensiver wird der eigentliche Schritt der Klassifizierung [24].

Ausgabeinformation von Schritt (4) *Feature-Engineering* der Prozesskette zur Untersuchung von Bewegungsaktivität ist eine Menge von Instanzen, welche jeweils genau einem Ereignis e_i entspricht. Im Merkmalsraum wird jede Instanz von einem Merkmalsvektor X_i beschrieben.

2.4.2.5 Schritt 5: Klassifizierung

Die *Klassifizierung*, zusammen mit dem notwendigen Training eines Vorhersagemodells, ist der letzte Schritt innerhalb der zu beschreibenden Prozesskette nach Abbildung 2.3. Während der Trainingsphase wird zunächst ein Vorhersagemodell trainiert, auf dessen Basis es später möglich ist, Ereignisse einer spezifischen Ereignisklasse zuzuordnen. Dies geschieht unter Zuhilfenahme eines definierten Sets von gelabelten Ereignisinstanzen $T = \{(X_i, l_i)\}_{i=1}^N$, wobei N die Anzahl aller Paare von Merkmalsvektoren X_i mit ihren korrespondierenden Klassenlabels l_i aus der Menge aller im Merkmalsraum definierten Klassen $L = \{l_1, l_2, \dots, l_k\}$ mit der Größe k ist. Die Anzahl N der Merkmalsvektor-Label-Paare konvergiert dabei direkt mit der Anzahl der Ereignisinstanzen, welche in Schritt (3) segmentiert wurden. Über das automatisierte oder manuelle Vergeben von Labels werden die nötigen Grundvoraussetzungen für einen überwachten Lernprozess geschaffen. Das Merkmalsvektorset T wird dafür nach einem zuvor festgelegten Verhältnis (zum Beispiel 80:20-Split nach Ng oder via 10fCV [158, 174]) in einen Testdatensatz $T_{test} = \{(X_i, l_i)\}_{i=1}^{N_{test}}$ und einen Trainingsdatensatz $T_{train} = \{(X_i, l_i)\}_{i=1}^{N_{train}}$ unterteilt, wobei gilt $N_{test} = N - N_{train}$. T_{train} ist zum Training eines Vorhersagemodells bestimmt, T_{test} zu dessen anschließender Validierung.

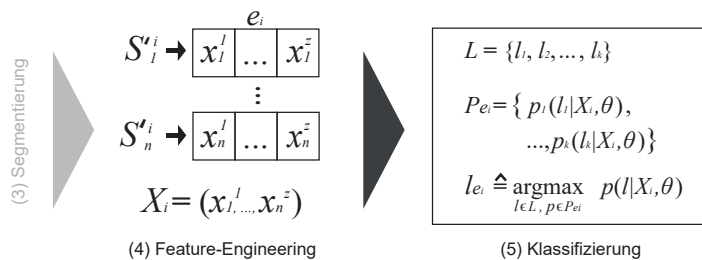


Abbildung 2.8: Untersuchung der im vierten Schritt kreierten Merkmalsvektoren X_i im Rahmen des letzten Schrittes (5) *Klassifizierung* mithilfe eines Prozesses des überwachten, maschinellen Lernens.

Anschließend kann mithilfe eines der in Kapitel 2.3.1.1 dargestellten Algorithmen zum überwachten Lernen ein Vorhersagemodell auf Basis von T_{train}

und unter Einbezug der Modellparameter θ trainiert werden. Die initialen Parameter θ sind für jeden Klassifikator von unterschiedlicher Art und Anzahl und dienen als Stellschrauben zur Anpassung des zugrundeliegenden Klassifizierungsalgorithmus an verschiedene, durch den zu untersuchenden Datensatz vorgegebene, Rahmenbedingungen. Dadurch können über geschickte Parameterwahl Klassifizierungsfehler minimiert und die Vorhersagequalität optimiert werden.

Für ein *Hidden Markov*-Modell definieren sich diese Parameter beispielsweise über $\theta = \{\pi, A, B\}$ mit der Matrix A , welche mit den Übergängen zu unterschiedlichen Zuständen korrespondiert, der Matrix B , welche die Ausgangswahrscheinlichkeiten jedes Zustandes beschreibt sowie π für die initialen Zustandswahrscheinlichkeiten [156, 200, 104, 63]. Andere Algorithmen wie der kNN sind nicht parametrisiert und nutzen nur das gelabelte Trainingsset $\theta = (T_{train})$ als Eingangsinformation [24]. Optimale Eingangsparameter können zum Beispiel via manuellem Vorgehen, via Erwartungsmaximierung über Training eines separaten Vorhersagemodells für jede spezifische Klasse im Merkmalsraum, via diskriminativem Ansatz zur Fehlerminimierung durch Gradientenverfahren oder via automatisierter HPO (vgl. Abschnitt 2.3.4) bestimmt werden [156, 199, 63].

Nach erfolgreichem Training eines Vorhersagemodells kann der finale Schritt innerhalb der Prozesskette, die eigentliche Klassifizierung, erfolgen (vgl. Abbildung 2.8). Dazu wird für alle Zuordnungen von vorhandenen Merkmalsvektoren im Merkmalsraum X_i zu definierten Ereignisklassen, welche über ihr Label L identifiziert werden können, eine Menge von Konfidenzwerten $P_i = \{p_i^1, p_i^2, \dots, p_i^k\}$ errechnet. Diese werden meist von Distanzmetriken oder Wahrscheinlichkeiten repräsentiert und mit einer vom Klassifikator abhängigen Interferenz-Methode errechnet:

$$I(X_i, \theta) = p_i(l|X_i, \theta), \text{ für } l \in L \quad (2.15)$$

In einem Folgeschritt können diese zuvor errechneten Konfidenzwerte nun benutzt werden, um die wahrscheinlichste Klasse für eine spezifische Ereignisinstanz e_i zu ermitteln. Für Wahrscheinlichkeiten als Konfidenzgrößen kann im einfachsten Fall der Maximalwert aus der Menge P_i ermittelt und das korrespondierende Klassenlabel l_i als Ausgabe der Klassifikation genutzt werden:

$$l_i = \arg \max_{l \in L, p \in P_i} p(l|X_i, \theta) \quad (2.16)$$

Über die Definition von Schwellwerten können Konfidenzwerte weitergehend auch dazu benutzt werden, um unbekannte Ereignisse zu erkennen und ein Vorhersagemodell automatisiert erneut zu trainieren, um so auch solche Ereignisse zu erlernen. Zusätzlich kann so auch eine relative Toleranz gegenüber Ausreißern implementiert werden.

2.5 Zusammenfassung

In diesem Kapitel wurden Definitionen und Begrifflichkeiten eingeführt, welche in den anschließenden Kapiteln dieser Arbeit zur Anwendung kommen. Darüber hinaus wurde ein Überblick über Grundlagen und verwandte Arbeiten mit Fokus auf der Untersuchung von menschlicher Bewegungsaktivität unter Zuhilfenahme von Sensorinformationen gegeben.

Im weiteren Verlauf dieser Arbeit sind vor allem die folgenden Erkenntnisse von besonderer Relevanz:

- Abschnitt 2.1 befasste sich mit der Definition von wichtigen Begrifflichkeiten, welche im Kontext dieser Arbeit genutzt werden. Diese dienen der Präzisierung von deren Fokus und ermöglichen eine differenzierte Unterscheidung unterschiedlicher Ansätze zur Untersuchung menschlicher Bewegungsaktivität.
- Die beschriebenen Konzepte der Signalverarbeitung aus Abschnitt 2.2 dienen in den Anschlusskapiteln 3, 4 und 5 als Grundlage zur Entwicklung von Verfahren zur Vorverarbeitung von Eingangsinformationen, zur Identifikation von Ereignissen innerhalb kontinuierlicher Zeitreihen sowie zu deren Extraktion in individuelle Ereignissegmente. In Kapitel 3 wird auch ein Vorgehen zur Realisierung von Merkmalsvektoren auf Basis von Resampling und Interpolation vorgestellt.
- In allen der drei nachfolgenden Kapitel 3, 4 und 5, werden zur Untersuchung und Extraktion von Wissen aus roher Sensorinformation Werkzeuge des maschinellen Lernens genutzt, welches in Abschnitt 2.3 vorgestellt wird. Ein besonderer Schwerpunkt liegt hierbei auf Algorithmen und Vorgehensweisen des überwachten maschinellen Lernens.
- Im Rahmen des Abschnitts 2.4 wurde eine Einführung in die Thematik der Aktivitätserkennung gegeben. Dabei wurde die Untersuchung von menschlicher Bewegungsaktivität als eine von deren Teildisziplinen definiert. Zusätzlich wurde eine generalisierte Prozesskette zur Untersuchung von menschlicher Bewegung auf Basis von Sensorinformationen sowie mithilfe von Werkzeugen des überwachten maschinellen Lernens, präsentiert. Diese dient in den folgenden Kapiteln 3, 4 und 5 als Grundlage zur Entwicklung der jeweils vorgestellten Untersuchungskonzepte.

Im Anschluss wird in Kapitel 3 mit dem SensX Sensorsystem eine Grundlage zur Erfassung und Erkennung von menschlicher Bewegungsaktivität geschaffen. Nachfolgend beschäftigt sich Kapitel 4 mit der automatisierten Untersuchung und qualitativen Bewertung von wiederkehrender, komplexer menschlicher Bewegungsaktivität. Kapitel 5 präsentiert Vorgehensweisen zur quantitativen und qualitativen Untersuchung von nicht vorhersehbarer, komplexer menschlicher Bewegungsaktivität auf Basis eines feingranularen Feature-Modells zu deren Beschreibung.

3 Erfassen und Erkennen von menschlicher Bewegung

In diesem Kapitel wird ein verteiltes Sensorsystem namens SensX zur ganzheitlichen Erfassung von menschlicher Bewegungsaktivität auf Basis von Bewegungssensorik vorgestellt. Dieses soll autark von externer Infrastruktur und ohne Einschränkungen durch gegebene Laborbedingungen arbeiten, um die Erfassung und Erkennung von Bewegungsereignissen in Echtzeit zu ermöglichen.

Nach den Vorveröffentlichungen in Abschnitt 3.1 und der Motivation zum vorliegenden Kapitel in Abschnitt 3.2 werden in Abschnitt 3.3 zunächst verwandte Arbeiten zur Erkennung menschlicher Aktivität vorgestellt, um anschließend auf dieser Grundlage in Abschnitt 3.4 die Anforderungen an SensX und damit dessen grundlegende Systemarchitektur zu definieren. Nachfolgend wird im Rahmen von Abschnitt 3.5 die Umsetzung des Systems, untergliedert in die Hardware-Ebene und die logische Ebene, im Detail beschrieben. Hier werden auch Einblicke in das Vorgehen zur Vorverarbeitung roher Bewegungsinformationen, die Segmentierung einzelner Bewegungsereignisse und den anschließenden Aufbau von diskreten Merkmalsvektoren für einzelne Ereignisinstanzen gegeben. Abschnitt 3.6 gibt mit der Evaluation einen Überblick über eine umfangreiche Studie zur Validierung des vorgestellten Konzepts sowie zu den Ergebnissen hinsichtlich der Klassifizierung von Bewegungsereignissen. Darüber hinaus wird auch der Einfluss unterschiedlicher Sensorkonfigurationen auf das Klassifizierungsergebnis und die Systemlaufzeit untersucht. Schwerpunkt ist dabei die Untersuchung der erfassten Bewegungsinformationen hinsichtlich ihrer Tauglichkeit zur quantitativen Erkennung einzelner, komplexer Bewegungsaktivitäten (vgl. Kapitel 2.1.3).

Das aus diesem Kapitel resultierende Konzept ist dabei nicht nur fähig, menschliche Bewegungsaktivität in quantitativer Manier zu erfassen und zu erkennen, es liefert mit der ganzheitlichen Erfassung eines Athleten auch die benötigte Informationsgrundlage für die qualitativen Untersuchungen im Rahmen der anschließenden Kapitel 4 und 5.

3.1 Vorveröffentlichungen

Die Kerninhalte dieses Kapitels wurden vom Autor bereits in [51] publiziert. Die in diesem Rahmen zum Einsatz kommenden, grundlegenden Konzepte zur Segmentierung von Bewegungsereignissen aus kontinuierlichen, multidimensionalen Zeitreihen wurden vom Autor zuvor bereits in [49] und [50] veröffentlicht.

Der im Rahmen der Evaluation verwendete Datensatz wurde ebenfalls bereits in [53] veröffentlicht. Wie in Kapitel 1.2 dargestellt, stammen die in den Vorveröffentlichungen und im Folgenden präsentierten Inhalte bezüglich der Idee, des Konzepts und der Evaluation vom Autor der vorliegenden Arbeit. Die Evaluation von SensX wurde in der Vorveröffentlichung auf Basis eines im Rahmen einer Vorstudie erfassten Datensatzes durchgeführt. Um nun eine Vergleichbarkeit zu den Folgekapiteln zu ermöglichen, wurde die Evaluation in diesem Kapitel auf Basis des Datensatzes aus [53] erneut durchgeführt; dabei kamen die gleichen Validierungsverfahren wie in der ursprünglichen Veröffentlichung zum Einsatz. Die Ergebnisse zur Systemlaufzeit in Abschnitt 3.6.4 wurden zuvor noch nicht veröffentlicht. Die Abbildungen 3.1, 3.2 und 3.7 sind bereits in [51] enthalten.

3.2 Grundidee und Motivation

Die Erfassung und Untersuchung menschlicher Bewegung, sowohl quantitativ als auch qualitativ, ist vor dem Hintergrund der rasanten Entwicklung von Fitness- und eHealth-Produkten auf dem IoT-Markt sowie im medizinischen Umfeld von essentieller Bedeutung. Zahlreiche Entwicklungen wie Fitness-Armbänder auf Basis von Bewegungssensorik bieten Funktionen zum quantitativen Tracking von täglicher menschlicher Aktivität [159, 163]. Im Bereich Sport sind diese Wearables oft mit Smartphone-Applikationen gekoppelt, welche Informationen übersichtlich visualisieren und darüber hinaus anspruchsvolle Trainings- und Zeitpläne bereitstellen [165, 167, 162]. Durch die bislang meist quantitative Analyse von Sensorinformationen können so Bewegungsereignisse erkannt oder gezählt werden; objektive und datengetriebene Rückschlüsse auf die Qualität dieser, beispielsweise im Sinne von Korrektheit der Durchführung, werden aber bislang selten getroffen. Dabei zeigen einschlägige Arbeiten im sportlichen Umfeld, dass gerade die Qualität und Korrektheit der Ausführung von menschlichen Bewegungsabläufen einen signifikanten Einfluss auf das persönliche Leistungempfinden und damit verbunden auf den Trainingserfolg und die Motivation haben. Diese wiederum sind auch positive Faktoren zur Steigerung des psychischen Wohlbefindens, des Selbstvertrauens und der mentalen Gesundheit des Menschen [29, 157]. Darüber hinaus können Fehlstellungen in Bewegungsabläufen aber nicht nur zu vermindertem Trainingserfolg, sondern auch zu temporären Verletzungen oder sogar zu ernsthaften chronischen Erkrankungen führen [93, 120, 109, 2].

Eine fundierte Überprüfung von Bewegungsabläufen zu deren Optimierung und zur Verhinderung von ungewollten Fehlstellungen ist mit aktuellen Konzepten, wie zum Beispiel den oben genannten Wearables in Kombination mit Smartphone-Anwendungen oder stationären Auswertungssystemen für einen Anwender selbst zur Zeit nicht möglich. Dies ist zum einen bedingt durch bislang fehlende Konzepte zur ganzheitlichen sensorischen Erfassung des menschlichen Körpers bedingt, welche benötigt werden, um auch komplexe Bewe-

gungsabläufe, unter Einbezug der gleichzeitigen Koordination verschiedener menschlicher Extremitäten, zu untersuchen. Zum anderen fehlen auch logische Konzepte, welche die Untersuchung von Bewegungsereignissen unter qualitativen Aspekten in Echtzeit umsetzen.

Um die Grundlagen zur Adressierung dieser Herausforderungen zu schaffen, wird in diesem Kapitel die Konzeption, Implementierung und Evaluation des SensX-Sensorsystems vorgestellt. Dieses arbeitet auf Basis von Infrastruktur-unabhängigen Bewegungssensoren und ist fähig zur ganzheitlichen Erfassung und Erkennung von menschlicher Bewegung in Echtzeit. Durch die Erfassung aller Extremitäten des menschlichen Körpers soll damit die Untersuchung komplexer Bewegungsereignisse ermöglicht werden. Zusätzlich ist das Konzept hoch mobil und auch in widrigen Szenarien ohne zusätzliche Assistenz durch Dritte einsetzbar. Nach der Evaluation von SensX bezüglich dessen Fähigkeiten zur quantitativen Erfassung und Erkennung von menschlicher Bewegungsaktivität im Rahmen dieses Kapitels dient es anschließend in den Kapiteln 4 und 5 als Grundlage zur feingranularen, qualitativen Untersuchung von menschlicher Bewegungsaktivität.

3.3 Verwandte Arbeiten

In diesem Abschnitt werden verschiedene Arbeiten zu Erfassung und Erkennung von menschlicher Aktivität vorgestellt. Generell können diese in drei unterschiedliche Kategorien eingeordnet werden: 1) visuell arbeitende Systeme, 2) Systeme auf Basis von Tiefeninformation und 3) Systeme auf Basis von Bewegungsinformationen.

3.3.1 Arbeiten auf Basis von visueller Information

Im Folgenden werden verschiedene Arbeiten zur visuellen Bewegungserfassung und -untersuchung präsentiert.

Elegantal et al. stellen hierzu ein markerloses, visuelles Überwachungssystem auf Basis von Vorder- und Hintergrund-Modellierung vor [57]. Dazu wird zunächst eine statistische Repräsentation einer Szenerie erstellt, welche sensitiv auf sich bewegende Objekte reagiert. Die Autoren beschreiben das System als robust gegenüber natürlichen und künstlichen Störinformationen, wie beispielsweise Lichteinflüssen oder Bildrauschen. Nachfolgend wird eine zweite statistische Repräsentation des Vordergrundes kreiert, welche die Erfassung und Verfolgung von sich bewegenden Objekten unter Beachtung der Umgebungsverdeckung realisiert. Da die Wahrscheinlichkeitsdichtefunktion des Vorder- und Hintergrunds von Bild zu Bild variiert und keine generell bekannte parametrisierte Form haben kann, entwickeln die Autoren einen generischen, nicht-parametrisierten Kerndichteschätzer zur Bildung von statistischen Modellen, welche die Bestimmung der Wahrscheinlichkeitsdichtefunktion ohne Kenntnis einer zugrundeliegenden Verteilung erlauben.

Auch Corozza et al. stellen ein markerloses System zur Bewegungserfassung vor, welches im klinischen Umfeld zur Untersuchung der Biomechanik von Muskeln und des menschlichen Skeletts genutzt werden kann [37]. Dazu wird der Körper eines Probanden a priori mit einem visuellen Hüllmodell rekonstruiert und anschließend ein heuristisches Approximationsverfahren, das *Simulated Annealing*, auf dieses Modell angepasst. Damit können Bewegungsabläufe wie beispielsweise das Gehen, aber auch Bewegungen einzelner Körperteile sowie Hüft- und Knieflexion oder -erweiterung erfasst werden.

Chan et al. stellen ein markerbasiertes System zum Training von Tanzbewegungen vor, mit welchem Tänzer auf eine Wand projizierte Bewegungen nachahmen können. Ähnlich dazu ist auch ein System von Komura et al. zum Trainieren von Kampfsportarten [108]. Beide stützen sich dabei auf das Erfassen von Bewegungen mithilfe eines Anzuges mit angebrachten Markern, welche dreidimensional von Kameras in einem Aufnahmerraum gefilmt werden. Ein weiterer großer Anwendungsbereich für markerbasierte Aufnahmesysteme in Kombination mit visuellen Erfassungssystemen ist auch die Entwicklung von Animationen für Filme und Videospiele [154, 134].

Gravila et al. bieten einen umfangreichen Überblick über zahlreiche weitere Systeme zur visuellen Bewegungserfassung [68]. Sie ordnen diese dabei in drei Kategorien: 1) zweidimensionale Ansätze ohne Generierung expliziter Körperformen, 2) zweidimensionale Ansätze mit Generierung von Körperformen und 3) die dreidimensionale Erfassung.

Wang et al. präsentieren mit dem *Trajectory-pooled Deep-convolutional Descriptor* TDD ein visuelles, markerloses Verfahren zur Erkennung von menschlicher Aktivität in Videos [204]. Dabei verbinden sie Features, welche mit einem herkömmlichen Ansatz manuell extrahiert wurden mit solchen, welche beim *Deep Learning* zum Einsatz kommen. Dazu werden im manuellen Ansatz mithilfe des SURF-Algorithmus temporale Trajektorien aus markanten Bildpunkten innerhalb zusammenhängender Bewegungssequenzen extrahiert [12]. Diese werden anschließend zusammen mit generischen Feature-Maps, welche unter Zuhilfenahme eines vortrainierten *Convolutional Neural Networks* (CNN) extrahiert werden, in einen Feature-Vektor überführt. Anschließend nutzen die Autoren einen linearen SVM-Klassifikator zur Aktivitätserkennung. Im Rahmen der Evaluation, welche die Autoren auf Grundlage von zwei in verwandten Arbeiten vielgenutzten und frei verfügbaren, gelabelten Videodatensätzen (UCF101[194] und HMDB51[116]) durchführen, erreichten sie eine durchschnittlich korrekte Klassifizierungsrate von 91,5% für den UCF101-Datensatz und 65,9% für HMDB51-Datensatz.

Marouane et al. präsentieren ein umfangreiches, markerloses Konzept zur Schritt- und Aktivitätserkennung sowie einen Kompass auf Basis von visuellen Informationen, welche mit einer am Körper getragenen Kamera aufgezeichnet werden [127, 130, 128]. Hervorzuheben ist hierbei, dass das Konzept im Gegensatz zu anderen Ansätzen nicht auf Basis von Bildinformationen aus einer Dritte-Person-Perspektive, sondern solchen aus der Erste-Person-Perspektive

arbeitet. Haupteingangsinformation in den Prozess der Feature-Extraktion sind dabei temporale Trajektorien, welche mithilfe des SURF-Algorithmus aus den Scale- und Orientation-Informationen zeitlich geordneter Bildfolgen konstruiert werden.

Zusammenfassend ist zu vermerken, dass die vorgestellten und auf Basis von visuellen Bildinformationen arbeitende Ansätze teilweise sehr gute Resultate in puncto Rekonstruktion und Nachzeichnung von menschlicher Bewegung ermöglichen. Aufgrund stationärer Bindung eignen sie sich aber nicht für den in diesem Kapitel fokussierten Anwendungszweck, da sie zu ihrem Einsatz entweder eine umfangreiche technische Infrastruktur benötigen und damit nicht mobil sind oder einen Anwender aus einer zur Bewegungsanalyse untauglichen Position erfassen. Zusätzlich ist der Einsatz der meisten Konzepte, bedingt durch den Bedarf von Fachpersonal zu ihrer Bedienung, außerhalb von Laborbedingungen nahezu unmöglich und äußerst kostspielig. Ein weiteres Problem ergibt sich durch Verdeckungseffekte einzelner Körperteile beim Einsatz von zweidimensional arbeitenden Systemen. Dieser Effekt kann mit mehrdimensionalen Ansätzen behoben werden, was aber ebenso zu erhöhter Komplexität und geringerer Mobilität führt. Der in diesem Kapitel geforderte Anwendungszweck fordert aber konträr dazu eine hohe Mobilität, robuste Sensorik mit einfachem Handling und einen Aufbau aus konventionellen, preiswerten Komponenten. Zusätzlich ist auch die generische Untersuchung unterschiedlichster Bewegungsereignisse in unterschiedlichen Anwendungsszenarien und unter verschiedenen Umwelteinflüssen notwendig. Diese Anforderungen sind mit visuell arbeitenden Konzepten nicht vollständig zu realisieren.

3.3.2 Arbeiten auf Basis von Tiefeninformationen

Spätestens mit der Einführung des *Kinect*-Systems durch Microsoft¹ hat der Einsatz von Tiefensensorik auch den Einzug in den breiten Markt der Consumer-Elektronik geschafft [152].

Ein Anwendungsbeispiel zur Bewegungsuntersuchung mithilfe von Tiefensensorik liefern Parisi et al. mit der Vorstellung eines Systems zur Bewertung und zum Training von Gewichthebeübungen. Dazu werden Bewegungen eines Gewichthebers mithilfe des infrarotbasierten Sensors einer Kinect erfasst und unter Zuhilfenahme eines rekurrenten neuronalen Netzwerks klassifiziert [148].

Pfister et al. vergleichen die Leistungsfähigkeit von Tiefensensorik mit derer von herkömmlichen visuellen Verfahren zur Bewegungserfassung [152]. Dazu wurde eine größere Menge von Studienteilnehmern während des Gehens und des Joggens in drei Geschwindigkeiten aufgezeichnet und die gleichzeitige Hüft- und Knie-Flexion sowie das Schritt-Timing gemessen. Im Vergleich zu traditioneller, visueller Analyse wurden Flexion und Erweiterung von Hüfte und Kniebewegung bei Erfassung mithilfe des Tiefensensors während des Untersuchungsprozesses unterschätzt bzw. überschätzt und waren damit nicht tauglich

¹<https://developer.microsoft.com/de-de/windows/kinect>

für fundierte, klinische Beurteilungen.

Diese Ungenauigkeiten, zusammen mit den bereits für visuelle Systeme genannten Problematiken der Verdeckung von Extremitäten und dem Zwang zur stationären Anwendung, machen Systeme, welche auf Tiefensensorik basieren, für den im Rahmen dieser Arbeit angedachten Anwendungszweck unbrauchbar. Zwar gibt es auch hier Konzepte, welche die Verdeckungsproblematik durch Verbindung mehrerer Sensorplattformen zu adressieren versuchen, allerdings verstärkt dies auch hier das stationäre Design der Architektur und macht diese damit weniger mobil [9].

3.3.3 Arbeiten auf Basis von Bewegungsinformationen

Durch den Trend hin zur Entwicklung kleinerer, flexiblerer und unabhängiger Sensorplattformen hat die Nutzung von Wearables mit Bewegungssensorik in den letzten Jahren einen beständigen Aufwärtstrend erlebt.

In diesem Kontext präsentieren Dernbach et al. einen rein auf der Bewegungssensorik eines Smartphones basierenden Ansatz zur Erkennung von menschlicher Bewegungsaktivität. Dazu verarbeiten sie gesammelte Informationen sequentiell mit einem Sliding-Window-Verfahren und extrahieren für jedes Fenster einen Merkmalsvektor mit 30 Features, bestehend aus Mittelwert, Median, Minima, Maxima, Nulldurchgängen und anderen. Danach klassifizieren Sie diese Parameter mithilfe von überwachten Lernalgorithmen und erkennen sogenannte simple Aktivitäten wie Laufen mit einer Rate von über 90%, komplexe Aktivitäten mit etwa 35% - 50%. Mit komplexen Aktivitäten umschreiben die Autoren Tätigkeiten wie beispielsweise das Gießen von Blumen.

Strohrmann et al. untersuchen die Bewegungsinformationen von 21 Läufern, welche mit mobilen Bewegungssensoren, sog. ETHOS-Einheiten, aufgenommen wurden [197]. Dabei handelt es sich um an den Füßen der Läufer angebrachte, tragbare Sensorplattformen mit integriertem Gyroskop, Accelerometer und Magnetfeldmesser. Diese zeichnen in einem Offline-Szenario, außerhalb von Gebäuden, Bewegungsinformationen auf ihrem internen Speicher zur späteren Auswertung auf. Ziel der Arbeit war es, das Fitnesslevel der Probanden unter Berücksichtigung der Dauer des Bodenkontakts der Füße, der Art des Auftretens und dem Heben der Fersen als kinematische Parameter, zu evaluieren und so Überbelastungen zu erkennen. Durch die fehlende Möglichkeit zur Echtzeitberechnung konnte allerdings kein direktes Feedback an die Probanden übermittelt werden.

Chang et al. beschäftigten sich mit dem Erkennen von Gewichthebeübungen auf Basis von Beschleunigungsinformation [34]. Dazu trugen die Probanden ihrer Studie zwei Beschleunigungssensoren: einen integriert in einem Handschuh und einen in Form eines Clips am Gürtel. Zum Segmentieren einzelner, unterliegender Ereignisse entwickelten und testeten die Autoren zwei Algorithmen. Einer arbeitet auf Basis der Maxima in den Signalen, der zweite basiert auf dem

Algorithmus von Viterbi und wird mit einem Hidden-Markov-Model kombiniert [65]. Die unterschiedlichen Übungen konnten mithilfe eines Naive-Bayes-Klassifikators mit einer Vorhersagewahrscheinlichkeit von über 90% korrekt voneinander unterschieden werden.

Mit Recofit stellen Morris et al. ein System zur Erkennung von menschlicher Bewegungsaktivität mit unterschiedlichen Komplexitäten vor [139]. Zum Erfassen von Bewegung wird ein Sensorarmband mit integriertem Gyroskop und Accelerometer am rechten Unterarm angebracht und via Bluetooth mit einem stationären Rechner verbunden. Für jedes der sechs aufgenommen Bewegungssignale (2 Sensoren, 3 Dimensionen in X -, Y - und Z -Richtung) extrahieren die Autoren 28 Features zur anschließenden Segmentierung, u.a. Maxima, Wendepunkte, Korrelationen und erhalten damit einen Merkmalsvektor mit 168 Features zur Beschreibung von Bewegungsereignissen. Die Segmentierung selbst wird anschließend via überwachtem maschinellen Lernen unter Zuhilfenahme statischer Sliding-Windows realisiert. Die Klassifizierung zur Bewegungserkennung selbst findet auf Basis von Autokorrelationsmerkmalen, der RMS-Amplitude der Signale, sowie Merkmalen extrahiert aus den Frequenzbändern der Signale und Metriken wie der Standardabweichung statt, welche je Fenstergröße in einen 120-stelligen Merkmalsvektor gekapselt werden. Dazu kommt eine modifizierte SVM als Klassifikator zum Einsatz. Je nach Komplexität des Evaluationsszenarios konnten 97% der Bewegungsereignisse korrekt segmentiert und anschließend, je nach Validierungsmetrik, im besten Ergebnis mit einer Erfolgsrate von durchschnittlich 98,2% für die Untersuchung einer gesamten Aktivitätssequenz klassifiziert werden.

Chang et al. stellen mit *DCNN+* einen Ansatz vor, welcher menschliche Bewegungsaktivität mit Bewegungssensorik erfasst und diese anschließend in einen visuellen Untersuchungsprozess überführt [91]. Dazu ordnen sie Rotations- und Bewegungssignale zunächst in einem Signalabbild und überführen die beinhalteten Frequenzen mithilfe der *Diskreten Fourier Transformation* anschließend in ein Aktivitätsabbild. Dieses dient als Eingangsinformation für eine CNN-Architektur. Die Ausgangsinformationen aus diesem, in Form einer Wahrscheinlichkeitsverteilung für die unterschiedlichen Aktivitäten, werden anschließend noch einmal durch die Verarbeitung mit einem SVM-Klassifikator optimiert. Evaluationsgrundlage waren anschließend einfache menschliche Bewegungsaktivitäten wie Gehen, Stehen und Liegen. Im Rahmen dieses Szenarios erreichten die Autoren mit ihrem Ansatz im besten Ergebnis eine durchschnittliche korrekte Erkennungsrate von 99,93%.

Kosmolla et al. erfassen Kletterbewegung mithilfe von zwei Accelerometern, welche an den Handgelenken des Kletterers befestigt werden. Anschließend extrahieren sie aus den Bewegungsinformationen für einzelne Kletterrouten Features und übersetzen diese mithilfe eines Mappingverfahrens in Zeichenketten, um sie sukzessiv mithilfe der Levenshtein-Distanz, einer Distanzmetrik für Zeichenketten, im Rahmen eines Template-basierten Klassifizierungsverfahrens wiederzuerkennen [111].

Zusammenfassend ist zu vermerken, dass alle vorgestellten Systeme auf Basis von Bewegungssensorik in der Theorie unabhängig von Infrastruktur und ohne Laborbedingungen oder zusätzliches Personal zur Erfassung von Bewegung eingesetzt werden können. Allerdings ist bis auf den Ansatz von Dernbach et al. keines der Systeme fähig, Informationen in Echtzeit zu klassifizieren und Feedback an den Nutzer zu liefern. Auch ist kein System dazu fähig, die Bewegungen eines Menschen in der Gesamtheit und unter Einbeziehung aller Extremitäten zu erfassen.

3.3.4 Weitere Ansätze

Neben den zuvor vorgestellten Arbeiten gibt es auch Ansätze, welche mit Sensorik, die im vorliegenden Kontext weniger gebräuchlich ist, arbeiten. So zum Beispiel FEMO von Ding et al. [42]. Hierbei handelt es sich um eine Plattform zur Überwachung von Freihandtraining, welche auf dem Einsatz von RFID-Tags basiert. Allerdings lassen sich damit keine konkreten Bewegungen von Extremitäten nachvollziehen. Andere Sensorik, wie die Anwendung von GPS zum Tracken von Laufstrecken ermöglicht zwar das Nachvollziehen menschlicher Bewegung; diese wird aber nur aus einem Makro-Blickwinkel heraus betrachtet und es besteht keine Möglichkeit Rückschlüsse hinsichtlich der tatsächlichen Bewegungsabläufe einzelner Körperteile zu ziehen [11].

3.3.5 Bestandsaufnahme

In den vorhergehenden Abschnitten wurde ein Überblick über visuell arbeitende Systeme, über Tiefensensor-basierte Systeme und über mit Bewegungsinformationen arbeitende Konzepte zur Erfassung von menschlicher Bewegungsaktivität gegeben. Visuelle und Tiefensensor-basierte Systeme sind aufgrund ihrer komplexen Handhabung sowie der Notwendigkeit zu einer aufwendigen und immobilen Infrastruktur nicht für den im Rahmen dieser Arbeit adressierten Anwendungszweck geeignet. Auch haben beide Kategorien Defizite im Bereich der Erfassung von verdeckten Gliedmaßen. Zusätzlich sind visuelle Systeme im Allgemeinen kostspielig und meist nicht von einzelnen Personen ohne Assistenz Dritter bedienbar. Nur Konzepte aus dem Bereich der Bewegungssensorik erfüllen grundlegend die Forderungen nach Mobilität, stationärer Unabhängigkeit und Echtzeitfähigkeit. Allerdings konnte zum Zeitpunkt der Entwicklung des, im Rahmen dieser Arbeit entstandenen, SensX-Sensorsystems keine der verwandten Arbeiten die Bewegung des menschlichen Körpers ganzheitlich, unter Berücksichtigung aller Extremitäten und in Echtzeit betrachten. Dennoch zeigen die Arbeiten, dass der Einsatz von Bewegungssensorik, trotz vieler noch offener Fragestellungen, ein solides Fundament zum Aufbau eines unabhängigen, Echtzeit-fähigen Sensorsystems zur ganzheitlichen Untersuchung menschlicher Bewegungsaktivität bietet.

3.4 Konzeption einer verteilten Sensorplattform

In diesem Abschnitt werden die grundsätzlichen Anforderungen an SensX definiert sowie eine Basisarchitektur zur Erfassung von menschlicher Bewegung und drei generelle Komponenten zu deren Untersuchung spezifiziert. Dazu werden diese Komponenten entweder der Hardware-Ebene oder der logischen Ebene zugeordnet (vgl. Abbildung 3.1).

3.4.1 Anforderungsanalyse

Um die Entwicklung von SensX fokussiert und zielgerichtet umsetzen zu können, wurden die folgenden Anforderungen und Ziele vor dem eigentlichen Entwicklungsprozess identifiziert. Diese sollen im Nachgang auch zur belastbaren, qualitativen Evaluation des Systems herangezogen werden können.

1. **Unabhängigkeit und Mobilität** Das System soll unabhängig von stationärer Infrastruktur arbeitsfähig und über längere Zeiträume unter widrigen Bedingungen in unterschiedlichsten Umgebungen und Anwendungsszenarien mobil einsetzbar sein. Dabei muss es energieeffizient arbeiten und ohne externe Energieversorgung autark über einen längeren Zeitraum operieren können.
2. **Ganzheitlichkeit** Eine weitere Anforderung ist die Fähigkeit zur sensorischen Erfassung des gesamten menschlichen Körpers, um Bewegungen einzelner Extremitäten auch unabhängig von einander untersuchen zu können. Nur durch ganzheitliche Erfassung ist es möglich, komplexe Bewegungsabläufe zu rekonstruieren und sowohl in quantitativer als auch in qualitativer Hinsicht zu untersuchen. Dazu muss eine ausreichende Anzahl einzelner Sensoren sinnvoll am menschlichen Körper platziert werden. Diese müssen über eine angemessene Abtastrate verfügen.
3. **Generalisierbarkeit und Erweiterbarkeit** Das System muss auf technischer Ebene Bewegungsinformation generisch bereitstellen, welche anschließend auf der logischen Ebene mit unterschiedlichen Konzepten der Vorverarbeitung und Analyse individuell und je nach nach Nutzungsszenario weiterverarbeitet und untersucht werden können. Das System soll dabei ohne großen Aufwand an neue Szenarien, beispielsweise bislang unbekannte Bewegungsereignisse, angepasst werden können und dazu generische Schnittstellen zwischen der Hardware-Ebene und der logischen Ebene bereitstellen. Zur Vorhersage und Untersuchung von Ereignissen müssen unterschiedliche Klassifikatoren und Vorhersagemodelle ohne komplexe Adaption in die Verarbeitungskette integriert und gelernt werden können.
4. **Einfachheit und Robustheit** Da das System in erster Linie Nutzer im sportlichen und medizinischen Bereich adressiert und dort über längere

Zeiträume zu Studienzwecken getragen werden soll, darf es Probanden in ihrer Bewegungsfreiheit nicht einschränken oder Bewegungsabläufe beeinflussen. Es soll einer Person alleine bedienbar sein und auch für den Einsatz in widrigen Szenarien, wie beispielsweise während bewegungsintensiver Sportarten außerhalb von Indoor- und Laborbedingungen, ein robustes Design aufweisen, welches Umwelteinflüssen wie Feuchtigkeit oder Stößen bis zu einem gewissen Maß widersteht.

5. **Echtzeitfähigkeit** Zur Realisierung von Bewegungsanalysen in Echtzeit sowie der Möglichkeit zum Bereitstellen von Feedback und zur dynamischen Anpassung des Systems an neue Bewegungsabläufe muss eine Recheneinheit integriert sein, welche Untersuchungen und Rekonfiguration in Echtzeit ermöglicht. Zusätzlich dient diese auch zur Synchronisation der verteilten Sensorik sowie zur Speicherung erfasster Rohinformationen für längerfristige Auswertungen.
6. **Konventionalität** Haupteinsatzzweck des Systems ist die Untersuchung von menschlicher Bewegung für den Normalverbraucher. Es soll nicht auf teurer Spezial-Hardware aufbauen, sondern auf handelsübliche und preiswerte Komponenten zurückgreifen. Auch die verwendeten Algorithmen und Software-Werkzeuge zur Verarbeitung der Sensorinformationen sollen frei und offen verfügbar sein.

3.4.2 Spezifikation einer Grundarchitektur

Die im Rahmen dieser Arbeit entwickelte Grundarchitektur von SensX organisiert sich in zwei Hauptblöcke, die Hardware-Ebene und die logische Ebene, welche in Abbildung 3.1 dargestellt sind. Erstere umfasst alle technischen

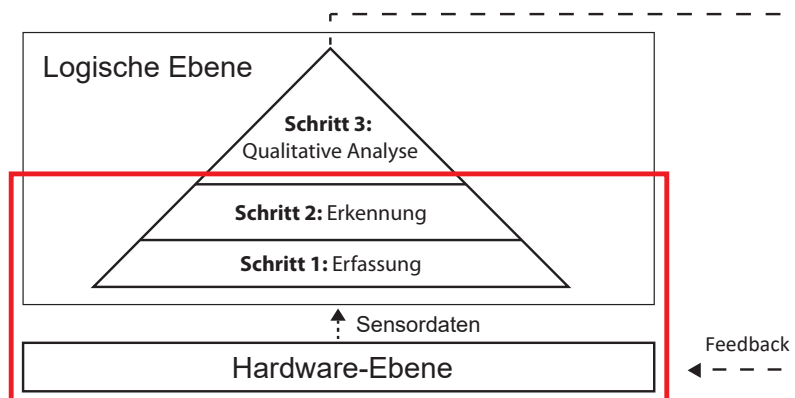


Abbildung 3.1: Übersicht über das Architekturkonzept von SensX, unterteilt in die Hardware-Ebene und die logische Ebene : Die Hardware-Ebene umfasst den technischen Aufbau, die logische Ebene die abstrahierten Schritte 1. Erfassung, 2. Erkennung und 3. qualitative Analyse von menschlicher Bewegung.

Aspekte, wie die einzusetzende Sensorik und Hardware-Komponenten sowie deren Kommunikation untereinander. Bei der einzusetzenden Sensorik handelt es sich im Fall von SensX pro Sensorplattform um ein triaxiales Gyroskop und ein triaxiales Accelerometer zur Erfassung von Beschleunigungs- und Rotationsinformationen des menschlichen Körpers. Die logische Ebene kapselt die einzelnen Schritte zur Vorverarbeitung und zur anschließenden Untersuchung unterschiedlicher Bewegungsereignisse. Zusätzlich zu den in diesem Kapitel, behandelten Schritten *1. Erkennung* und *2. Erfassung* ist auch Schritt *3. Qualitative Analyse* im Architekturschaubild mitaufgeführt, welche im Detail erst in den Kapiteln 4 und 5 dieser Arbeit adressiert wird. Die drei Schritte zur *1. Erfassung*, *2. Erkennung* und *3. Qualitative Analyse* menschlicher Bewegung sind dabei als elementare, generische Entitäten zu sehen, welche adaptiv mit unterschiedlichen logischen Bausteinen umgesetzt werden können, um das jeweils vorgesehene Ziel eines Schrittes für den jeweiligen Anwendungsfall zu erreichen. Die Ziele der einzelnen Schritte sind:

- **Schritt 1: Erfassung** Im Rahmen dieses Schrittes sollen individuelle Ereignisse innerhalb eines kontinuierlichen Flusses von Sensordaten erfasst werden. Dabei werden die von der Hardware-Ebene gelieferten Bewegungsinformationen mit jeweils für den Anwendungszweck geeigneten Vorverarbeitungs- und Segmentierungsansätzen verarbeitet, um anschließend Sequenzen von segmentierten Bewegungsereignissen an den nächsten Schritt, die Erkennung, weiterzugeben.
- **Schritt 2: Erkennung** Dieser Schritt realisiert die automatisierte Erkennung, Vorhersage und Unterscheidung von Ereignissen, beispielsweise durch den Vergleich von Ähnlichkeiten oder die Anwendung von Vorhersagemodellen. Die Ausgangsinformation des zweiten Schrittes an den dritten ist ein klassifiziertes Ereignis.
- **Schritt 3: Qualitative Analyse** Die Hauptaufgabe dieses Schrittes ist die tatsächliche Bewertung von Ereignissen, welche über das reine, aus verwandten Arbeiten bekannte Erkennen oder Zählen von Ereignisinstanzen hinausgeht. Nur auf Basis einer qualitativen Analyse und Bewertung kann einem Nutzer individuelles Feedback bereitgestellt und ein System adaptiv anhand dessen persönlicher Fähigkeiten justiert werden.

Die Umsetzung der Schritte 1 und 2 wurde zum Zeitpunkt der Entwicklung von SensX in dieser Form und mit den in Abschnitt 3.4.1 definierten Zielen in verwandten Arbeiten noch nicht betrachtet. Diese soll im Folgenden auf Basis der in Abschnitt 3.3.3 vorgestellten Grundkonzepte und verwandten Arbeiten realisiert werden. In puncto Ganzheitlichkeit und Echtzeitfähigkeit erfüllen die darin vorgestellten Vorarbeiten zwar nicht die Anforderungen dieser Arbeit, doch sie liefern umfangreiche Informationen zu technischen Details und zu einschlägigen Verfahren der Vorverarbeitung und des Einsatzes von

überwachten Klassifikatoren zur Erkennung von Bewegungsaktivität. Schritt 3. *Qualitative Analyse* ist in verwandten Arbeiten ein bis dato nur spärlich behandeltes Forschungsfeld. Es existieren nahezu keine Konzepte zur automatisierten, qualitativen Bewertung von menschlicher Bewegungsaktivität. Die Behandlung dieser zwei Aspekte ist in Abbildung 3.1 mit Schritt 3 vorgesehen und wird als Schwerpunkt in den anschließenden Kapiteln 4 und 5 für sowohl wiederkehrende als auch für nicht vorhersehbare Bewegungsereignisse thematisiert.

Das Hauptaugenmerk dieses Kapitels hingegen liegt auf den in Abbildung 3.1 rot hervorgehobenen Aspekten: der Entwicklung und Umsetzung einer technisch verteilten Sensorplattform sowie den notwendigen logischen Konzepten zur Erfassung und zur Erkennung von menschlicher Bewegungsaktivität auf Basis der Bewegungsinformationen.

3.5 Umsetzung von SensX

Dieser Abschnitt gibt Einblicke in die konkrete Umsetzung der in 3.4 definierten Architektur. Dabei unterteilt sich die Umsetzung in zwei Ebenen, die Hardware-Ebene, welche alle technischen Aspekte zur Erfassung von Bewegungsinformationen adressiert, und die logische Ebene, welche sich mit der Vorverarbeitung, der Segmentierung, dem Feature-Engineering und der Klassifikation von Bewegungsereignissen beschäftigt.

3.5.1 Hardware Ebene

Wie zuvor schon angedeutet, ergeben sich die Hardware-seitigen Anforderungen bereits aus den in Abschnitt 3.4.1 für SensX definierten Zielen. Die wichtigsten sind hierbei die individuelle Erfassung der einzelnen Gliedmaßen des menschlichen Körpers und die Realisierung der Untersuchung von Bewegungsaktivität in Echtzeit. Zusätzlich soll die verwendete Hardware auf frei erhältlichen und preiswerten Standardkomponenten bestehen. Um alle diese Anforderungen zu erfüllen, setzt SensX auf vier externe Sensorplattformen, welche jeweils an den Armen und den Beinen eines Probanden befestigt sind. Zusätzlich ist als zentrale Recheneinheit ein Smartphone, welches an der Brust getragen wird und zusätzlich auch die Beschleunigung und Rotation des Probanden misst, in das System integriert. Abbildung 3.2 zeigt das SensX-Sensorsystem, getragen von einem Probanden während einer der Evaluationsstudien. Folgend werden die technischen Spezifikationen von SensX sowie die zu dessen Betrieb notwendigen logischen Komponenten im Detail beschrieben.

3.5.1.1 Zentrale Recheneinheit

Die *zentrale Recheneinheit* wurde Mithilfe eines handelsüblichen Smartphones, welches an der Brust eines Probanden befestigt wird, realisiert. Dabei wur-

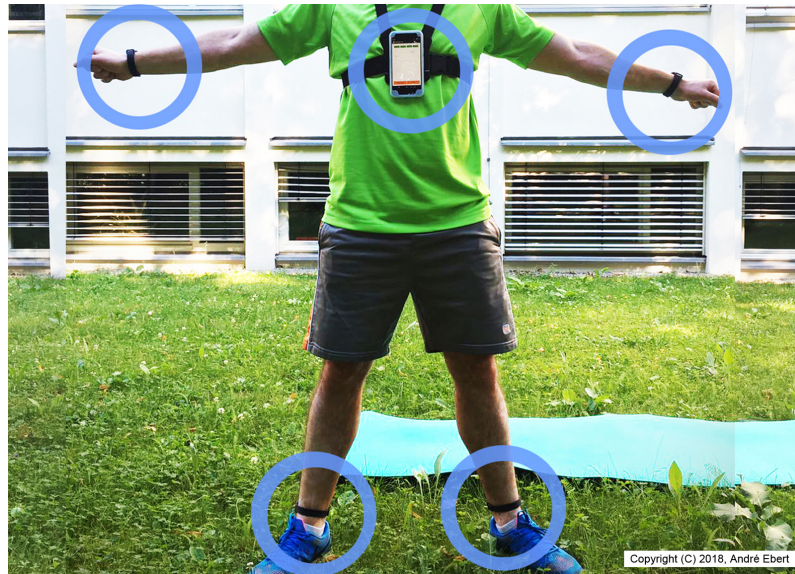


Abbildung 3.2: SensX getragen von einem Probanden: die vier externen Sensoren sind an den einzelnen Extremitäten befestigt, während das Smartphone als zentrale Recheneinheit und Sensor zugleich an der Brust getragen wird.

den Experimente mit zwei unterschiedlichen Gerätetypen durchgeführt: zum einen mit einem *Google Nexus 5x* mit *Android 6.0.1 Marshmallow* als installiertes Betriebssystem, zum anderen mit einem *HTC One (M7)* mit *Android 5.0 Lollipop*. Hauptgrund für das Arbeiten mit zwei unterschiedlichen Geräten sind unterschiedliche Abtastraten bei der simultanen Initialisierung mehrerer Bluetooth 4.0 (BLE) Verbindungen, worauf in Abschnitt 3.5.1.3 näher eingegangen wird. Während der Aufnahme von Trainingsdaten zur Entwicklung der logischen Ebene wurde auf den Geräten eine Android-Applikation installiert, welche das Aufnehmen von Beschleunigungs- und Rotationsinformationen sowie deren Labeling nach Art des durchgeführten Bewegungsereignisses ermöglicht. Die Bewegungsinformationen wurden zur anschließenden Entwicklung eines Analysekonzepts zeilenweise mit Zeitstempel und strukturiert nach jeweiligem Ereignis auf dem internen Speicher der Geräte abgelegt. Zusätzlich implementierte die Applikation auch die Möglichkeit zum manuellen und automatisierten Pausieren von Messungen, um so auch komplexere Abfolgen von Bewegungsereignissen strukturiert erfassen zu können. Die Abtastrate der in den Smartphones verbauten Bewegungssensorik war für die benutzten Modelle unterschiedlich hoch. Während das Gyroskop und das Accelerometer des LG Nexus 5x mit jeweils etwa 150 Hz operierten, erreichte das HTC je Sensor nur Abtastraten von etwa 100 Hz . Mit Aktivierung der in der Android API integrierten Funktion zur automatischen Bestimmung der linearen Beschleunigung A_{lin} , welche durch Herausrechnen der Erdbeschleunigung g mithilfe eines Hochpassfilters realisiert wird, fiel die Abtastrate durch den zusätzlichen

Rechenaufwand für beide Geräte auf etwa 50 Hz [170].

3.5.1.2 Externe Sensorik

Als externe Sensorik wurden vier *MBIENTLAB Meta Wear CPRO*² Sensorplattformen via BLE mit der zentralen Recheneinheit gekoppelt. Die Sensorplattformen, die zu ihrem Schutz genutzten Plastikkappen und die Gummibänder zur Befestigung sind in Abbildung 3.3 zu sehen. Das Sensormodell CPRO



Abbildung 3.3: Abbildung der externen Sensorplattformen mit Schutzhülle und Gummiband zur Befestigung. Zur autarken Energieversorgung dienen handelsübliche CR 2032 Knopfzellen.

ist eine Multisensor-Plattform und kann nicht nur Beschleunigung und Rotation, sondern auch Temperatur, Luftdruck, die magnetische Flussdichte und das Umgebungslicht messen. Im Gegensatz zu in verwandten Arbeiten verwendeter Sensorik ist die CPRO-Plattform zusätzlich fähig, gemessene Sensorinformationen in Echtzeit via BLE an die zentrale Recheneinheit zur Weiterverarbeitung zu senden. Die theoretisch mögliche Abtastrate des Gyroskops liegt zwischen 25 Hz - 3200 Hz während mit dem Accelerometer Raten von 0.78 Hz - 1600 Hz realisiert werden können. Durch die Übertragung der Informationen über einen einzigen BLE-Kanal wird die mögliche Datenrate in der Realität allerdings erheblich eingeschränkt (vgl. Abschnitt 3.5.1.3). Zur lokalen Speicherung von Sensorinformationen direkt auf der Plattform steht ein Speicher von 256 kB zur Verfügung. Je nach Menge der genutzten Sensoren und der angeforderten Abtastrate ist dieser Speicher aber nur zum Zwischenspeichern von Daten für wenige Minuten ausreichend. Insgesamt werden im Rahmen von SensX drei Dimensionen (X -, Y -, und Z -Richtung) von zwei Sensoren (Accelerometer und Gyroskop) und 5 Sensorplattformen (vier externe CPRO-Einheiten plus die zentrale Recheneinheit), also insgesamt 30 Signale in Echtzeit zur Analy-

²<https://mbientlab.com/product/metawearc/>

se von menschlicher Bewegung bereitgestellt. Die Laufzeit der Batterien der CPRO-Einheiten betrug bei durchgehender Sensoraktivität etwa eine Woche.

In dieser Hinsicht haben sich das kompakte Design der externen CPRO-Einheiten, die Möglichkeit zur Aktivierung unterschiedlichster Sensortypen und die Übertragung von Messwerten in Echtzeit sowie der günstige Anschaffungspreis (je nach Ausstattung und Modell etwa 50 – 90 USD pro Stück³) als essentiell für die erfolgreiche Entwicklung eines verteilten Sensorsystems, welches zur ganzheitlichen Erfassung des menschlichen Körpers fähig ist, herausgestellt. Vor allem große und den Träger behindernde Sensorik ohne Echtzeitfähigkeit war bislang eines der Haupthindernisse bei der Entwicklung eines Systems auf Basis von Bewegungssensorik, welches für vergleichbare Einsatzmöglichkeiten geeignet ist. Diese Hindernisse wurden mit dem vorgestellten System-Design umgangen.

3.5.1.3 Datentransfer und Kommunikation

Wie bereits erwähnt, gibt es zwei Möglichkeiten zum Organisieren aufgenommener Sensordaten: erstens das lokale Speichern von Sensorinformationen auf den jeweiligen Plattformen (256 kB für die externen Plattformen, je nach Modell und SD-Karte variabel für Smartphone) und zweitens die Übertragung via BLE an die zentrale Recheneinheit zur direkten Synchronisation, Ablage und Weiterverarbeitung. Im Fall von SensX wurde nur der zweite Fall implementiert, da der interne Speicher der CPRO-Einheiten nicht zum Erfassen von länger andauernden Bewegungsaktivitäten ausreicht und die Informationen in Echtzeit verarbeitet werden sollen.

In Abschnitt 3.5.1.2 wurde bereits auf die theoretisch möglichen Abstrakten der externen CPRO-Einheiten eingegangen. In der praktischen Nutzung sind diese allerdings durch die Nutzung einer einzigen BLE-Verbindung zur Integration aller vier Sensorplattformen in den Kommunikationskreis mit der zentralen Recheneinheit erheblich eingeschränkt. Laut dem von der *Bluetooth Special Interest Group* (SIG) verabschiedeten Standard IEEE 802.15.1⁴ müssen BLE-Empfänger mit einer Abtastrate von mindestens 100 Hz übertragen. Dieser Kanal muss allerdings zwischen allen mit dem System verbundenen Sensorplattformen aufgeteilt werden, was bei acht externen Sensoren (jeweils Accelerometer und Gyroskop je Sensorplattform) nur noch eine mögliche Rate von 12,5 Hz pro Sensor ermöglicht. Allerdings stellte sich während der im Laufe der Arbeit durchgeführten Experimenten heraus, dass verschiedene Hersteller ihre BLE-Empfänger nach unterschiedlichen Spezifikationen fertigen: während auf dem LG Nexus 5x Datenraten von 25 Hz pro externem Sensor realisiert werden konnten, wurden für das HTC One (M7) sogar 40 Hz erreicht, was einer Gesamtkapazität des BLE-Kanals von 200 Hz bzw. 320 Hz entspricht. Wie verwandte Arbeiten in Abschnitt 3.3 und die Ergebnisse in Ab-

³<https://mbientlab.com/store/>

⁴<https://www.bluetooth.com/specifications/bluetooth-core-specifications>

schnitt 3.6 sowie in den Kapiteln 4 und 5 zeigen, ist dies mehr als ausreichend zur fundierten Untersuchung von menschlicher Bewegungsaktivität.

3.5.1.4 Anbringung des Systems

Wie in Abbildung 3.2 dargestellt, wird SensX mithilfe von Gummibändern direkt am Körper der Probanden befestigt. Die externen CPRO-Plattformen werden dabei zum zusätzlichen Schutz in Plastikabdeckungen platziert und an den Hand- und Fußgelenken angebracht. So können die Bewegungen aller Gliedmaßen eines Probanden aufgezeichnet werden. Allerdings können Streckungen und Beugungen von Arm- und Kniegelenken, bedingt durch fehlende Abdeckung jeder Gliedmaße mit zwei Sensorplattformen (beispielsweise am Ober- und Unterarm platziert), hier nur bedingt rekonstruiert werden.

Die Gummiarmbänder und die zugehörigen Plastikabdeckungen sind in Abbildung 3.3 im Detail zu sehen. Um die Sensorarmbänder voneinander unterscheiden zu können und ihre Anbringung in immer gleicher Ausrichtung sicherzustellen, wurde diese auf der Plastikabdeckung markiert; zur Unterscheidung wurden die Kürzel der individuellen MAC-Adressen auf den Armbändern vermerkt.

Das Smartphone als zentrale Recheneinheit und zusätzliche Sensorplattform wurde mithilfe eines Brustgurts an der Brust der Probanden befestigt, wobei die Z -Achse der Bewegungssensorik von hinten (Rücken) nach vorne zeigt (Bauch), die Y -Achse von unten (Füße) nach oben (Kopf) und die X -Achse von der linken zur rechten Körperseite.

3.5.2 Logische Ebene

Auf der logischen Ebene werden nun die zuvor auf der Hardware-Ebene erfassten Sensorinformationen gesammelt, synchronisiert und mithilfe einer Prozesskette des maschinellen Lernens verarbeitet, um einzelne Bewegungsereignisse zu extrahieren und anschließend zu klassifizieren. Abbildung 3.4 projiziert die Aufnahme von Rohdaten sowie die Schritte 1 und 2 aus Abbildung 3.1 auf eine angepasste Version der bereits in Kapitel 2.4.2 vorgestellten, generischen Prozesskette und verdeutlicht so den nun folgenden Arbeitsablauf. In Grün ist in der Hardware-Ebene mit (1) die Aufnahme von rohen Sensordaten angeordnet. Die logische Ebene ist im blauen Bereich dargestellt und besteht aus den Unterschritten *Schritt 1: Erfassung* und *Schritt 2: Erkennung*, welche sich in (2) Vorverarbeitung, (3) Segmentierung, (4) Feature-Engineering und schließlich in die (5) Klassifizierung aufteilen. Nachfolgend wird die Umsetzung dieser einzelnen Teilaspekte zur Bewegungserkennung mit SensX im Detail erläutert, einzelne Arbeitsschritte werden hier teilweise zusammengefasst. Im Grundsatz orientiert sich die logische Architektur von SensX an einem klassischen Modell des maschinellen Lernens, wie es zuvor bereits in Kapitel 2.4.2 vorgestellt wurde.

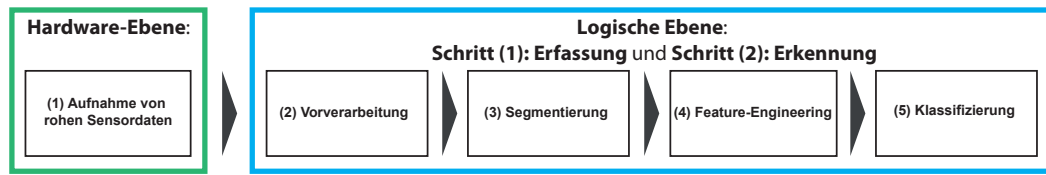


Abbildung 3.4: Zuordnung einzelner Schritte der Prozesskette des maschinellen Lernens aus Abbildung 2.3 zur Hardware-Ebene und zur logischen Ebene von SensX.

3.5.2.1 Vorverarbeitung und statische Segmentierung

Im Rahmen der Vorverarbeitung werden die 30 Eingangssignale des SensX-Sensorsystems zunächst mithilfe eines Tiefpassfilters geglättet (vgl. Kapitel 2.2.1), um so im Signal enthaltenes Grundrauschen zu reduzieren und enthaltene Sampling-Fehler zu minimieren. Dazu wurde ein Butterwort-Tiefpassfilter mit der Cutoff-Frequenz $f_c = f_s * cf$ definiert, wobei f_s jeweils der individuellen Sampling-Frequenz des Sensorsignals entspricht und cf dem mit dieser zu multiplizierenden Cutoff-Faktor. Die Größe des Cutoff-Faktors gibt dabei maßgeblich an, wie stark der Filter das Signal glättet und ist für jede eingehende Sampling-Frequenz individuell zu bestimmen. Im vorliegenden Fall wurde der optimale Bereich des Cutoff-Faktors Δcf durch Abgleich der Ergebnisse des im Anschluss präsentierten Algorithmus (vgl. Abbildung 3.5) zur Segmentierung von Bewegungsereignissen empirisch ermittelt. Dazu wird die Anzahl von extrahierten Bewegungsereignissen n_e innerhalb einer multidimensionalen Zeitreihe mit der zuvor bekannten tatsächlich enthaltenen Anzahl n_t für schrittweise vergrößerte cf -Konfigurationen abgeglichen. Δcf beschreibt schließlich den Bereich, in welchem $n_e = n_t$ gilt (vgl. Kapitel 4.3.1). Wird cf zu groß gewählt, so werden die Signale zu stark geglättet und zur Segmentierung benötigte Informationen sind nicht mehr enthalten. Ist der Faktor hingegen zu klein gewählt, so resultiert dies in einer zu geringen Glättungseigenschaft, wodurch Rauschinformationen oder auch lokale Extrema vom Segmentierungsalgorithmus fälschlicherweise als Indikatoren für individuelle Bewegungsereignisse gedeutet werden können und den Segmentierungsvorgang stören. Das Ergebnis einer solchen Glättung ist im vorhergehenden Kapitel in Abbildung 2.1 am Beispiel eines Sets von sich wiederholenden Bewegungsereignissen in Form der Körpereigengewichtsübung *Crunches* nachzuvollziehen.

Nach der Signalglättung im Rahmen des Vorverarbeitungsschrittes müssen nun individuelle Bewegungsereignisse aus den eingehenden Sensorinformationen extrahiert werden. Dazu wurde ein vom Autor dieser Arbeit im Rahmen einer Vorveröffentlichung entwickelter, Peak-basierter Segmentierungsalgorithmus mit statischer Fenstergröße angepasst und erweitert [49, 50]. In seiner ursprünglichen Form ist dieser zur Segmentierung von bei Fahrzeugen auftretenden Kollisionsereignissen, auf Basis zweidimensionaler Beschleunigungsinformation, gedacht. Ein Überblick über eine angepasste Variante, welche

```

1   $S = \{S_1, \dots, S_n\}$  // Menge  $S$  aller Eingangssignale  $S_i$  der Anzahl  $n$ 
2   $S_{MMS} = \text{getMMS}(S)$ ; // Ermittlung des Most-Meaningful-Signals  $S_{MMS}$  aus allen  $S$ 
3
4   $\text{getMMS}(S)$ :
5  for each element in  $S$  do
6  |    $STDs.add(std(element))$ 
7  end
8   $idx = \text{pos}(\max(STDs))$ ;
9  return  $S_{idx}$ 
10
11  $\text{processData}(S_{MMS})$ :
12 for  $cnt := 1$  to  $\text{length}(S_{MMS})$  do
13 |    $isPeak, lastPosVal, lastNegVal = \text{detectPeak}(s_{cnt}, lastPosVal, lastNegVal)$ ;
14 |   if ( $isPeak$ ) then
15 |       |    $\text{createEventFrame}(cnt, l)$ ; // Extrahiere Ereignis in ein Fenster der Länge  $l$ 
16 |       |    $cnt = lastExtractedSamplePos$ ; // Setze Zähler auf Position des letzten Samples
17 |       |   innerhalb des zuletzt extrahierten Ereignisses
18 |   end
19 end
20  $\text{detectPeak}(s_{cnt}, lastPosVal, lastNegVal)$ 
21 if ( $s_{cnt} \geq posThresh$ ) then
22 |   if ( $s_{cnt} \geq lastPosVal$ ) then
23 |       |    $lastPosVal = s_{cnt}$ ;
24 |       else
25 |           |    $isPeak = \text{true}$ ; // lokales Maximum wurde erkannt
26 |           end
27 |   else if ( $s_{cnt} < posThresh$  and  $lastPosVal \geq posThresh$ ) then
28 |       |    $isPeak = \text{true}$ ; // lokales Maximum wurde erkannt
29 |   end
30 |   if ( $s_{cnt} \leq negThresh$ ) then
31 |       |   if ( $s_{cnt} \leq lastNegVal$ ) then
32 |           |    $lastNegVal = s_{cnt}$ 
33 |           else
34 |               |    $isPeak = \text{true}$ ; // lokales Minimum wurde erkannt
35 |               end
36 |       else if ( $s_{cnt} > negThresh$  and  $lastNegVal \leq posThresh$ ) then
37 |           |    $isPeak = \text{true}$ ; // lokales Minimum wurde erkannt
38 |       end
39 |   end
40 |   return  $isPeak, lastPosVal, lastNegVal$ ;

```

Abbildung 3.5: Sequentielles, kontinuierliches Suchen von Bewegungsereignissen in multidimensionalen Zeitreihen anhand von Schwellwerten und Extraktion mit statischen Fenstergrößen nach Ebert et al. [49, 50].

die Extraktion von menschlichen Bewegungsereignissen auf Basis von multidimensionalen Zeitreihen ermöglicht, wird in Abbildung 3.5 gegeben. Eingangsparameter des Algorithmus ist dabei in Abbildung 3.5, Zeile 1, die Menge $S = \{S_1, \dots, S_n\}$ der $n = 30$ Einzelsignale S_i des SensX-Sensorsystems. Aus diesen wird anschließend das Signal, welches die größte Dynamik innerhalb seines Signalverlaufs aufweist, das sogenannte *Most-Meaningful-Signal* (MMS), in $\text{getMMS}()$ ausgewählt. Grundannahme hierbei ist, dass das MMS die Identifikation einzelner Bewegungsereignisse durch seine stark ausgeprägten Amplitudendifferenzen besonders gut ermöglicht. Dazu wird für jedes einzelne Signal S_i der Länge m dessen normalisierte Standardabweichung mithilfe der Berechnung 3.1 ermittelt. Als S_{MMS} wird anschließend nach Zuordnung 3.2 das Signal S_i ausgewählt, dessen korrespondierendes σ_i innerhalb der Menge

$M = \{\sigma_1, \dots, \sigma_n\}$ den maximalen Wert repräsentiert.

$$\sigma_{S_i} = \sqrt{\text{Var}(S_i)} = \sqrt{\frac{\sum_{j=1}^m (x_j - \mu)^2}{m}} \quad (3.1)$$

$$S_{MMS} = S_i, \quad \text{wobei gilt } \sigma_{S_i} = \mathbf{arg\,max}_{\sigma \in M} \{\sigma_i\} \quad (3.2)$$

Anschließend wird das S_{MMS} an `processData()` übergeben (vgl. Abbildung 3.5, Zeile 12). In diesem Schritt wird nun sequentiell jeder einzelne Parameter s_{cnt} aus $S_{MMS} = \{s_1, \dots, s_m\}$, zusammen mit dem letzten vorhergehenden positiven $lastPosVal$ und dem letzten vorhergehenden negativen Parameter $lastNegVal$, zur Überprüfung an `detectPeak()` übergeben. Initial gilt dabei $isPeak = false$ sowie $lastPosVal$ und $lastNegVal = 0$.

In `detectPeak()` wird nun überprüft, ob s_{cnt} einen empirisch ermittelten positiven Schwellwert $posThresh$ übersteigt. Wenn dem so ist und s_{cnt} ebenso größer als $lastPosVal$ ist, so befindet sich das Signal im positiven Anstieg und es gilt $s_{cnt} = lastPosVal$ (vgl. Abbildung 3.5, Zeile 23). Gilt hingegen $s_{cnt} \geq posThresh$ aber $s_{cnt} < lastPosVal$, so hat der Signalverlauf ein lokales Maximum passiert, welches als Indikator für ein dort befindliches Bewegungsereignis steht (vgl. Abbildung 3.5, Zeile 25). Zusätzlich liegt auch ein lokales Maximum vor, wenn S_{cnt} zwar kleiner und $lastPosVal$ größer als $posThresh$ ist, da $lastPosVal$ jeweils nur mit Werten von vorhergehenden s_{cnt} , welche größer als $posThresh$ sind, belegt werden kann (vgl. Abbildung 3.5, Zeile 28). Damit muss das Signal also auch hier einen Richtungswechsel vorgenommen und von einem positiven Anstieg in einen negativen Abfall gewechselt haben. Analog zu dem Vorgehen zur Prüfung auf lokale Maxima erfolgt dieses auch für lokale Minima in den Zeilen 30-37 der Abbildung 3.5. Rückgabewerte aus `detectPeak()` sind für jeden überprüften Parameter jeweils der letzte überprüfte positive Parameter $lastPosVal$, der letzte überprüfte negative Parameter $lastNegVal$ und der boolesche Wert $isPeak$. Dieser beinhaltet die Information, ob an der Position $p_{peak} = s_{cnt} - 1$ ein lokales Extrema vorliegt oder nicht.

Gilt für den Rückgabeparameter $isPeak == true$, so wird die Extraktion des Bewegungsereignisses in `createEventFrame()` angestoßen. Dazu wird ein Fenster statischer Länge um die Position p_{peak} gelegt und an den Fenstergrenzen der Beginn des Ereignisses t_s und das Ende des Ereignisses t_e in Form von zu den Sample-Positionen gehörigen Zeitstempeln vermerkt. Anschließend werden t_s und t_e genutzt, um die Bewegungsinformationen des Ereignisses aus allen $n = 30$ Signalen S auszuschneiden. Jedes extrahierte Ereignis e_x in der Menge aller extrahierten Ereignisse $E = \{e_1, \dots, e_j\}$ wird also von einem Set von n Einzelsignalen gleicher Länge beschrieben.

Dieses Vorgehen zur Ereignisextraktion basiert stets auf der Annahme, dass die ungefähre, zeitliche Dauer eines Bewegungsereignisses im Vorfeld bekannt

ist. Im Fall der zur Evaluation herangezogenen Studie (vgl. Abschnitt 3.6.1), in welcher unterschiedliche Körpereigengewichtsübungen erfasst wurden, ergab sich je nach Übungsklasse ein gemittelttes Zeitfenster Δt von 1 s bis 1,5 s. Die individuelle Dauer eines Ereignisses und deren Auswirkungen auf dessen Klassifikation wird anschließend in Kapitel 4 im Rahmen eines Vergleichs mit einem neuen adaptiven Segmentierungsansatz genauer untersucht. Lösungsmöglichkeiten zur Extraktion von Bewegungsereignissen mit zuvor nicht bekannter Dauer werden schließlich in Kapitel 5 vorgestellt.

3.5.2.2 Feature-Engineering

Das Feature-Engineering ist in Abbildung 3.4 als Unterschritt (4) dargestellt und befasst sich mit der Erstellung eines diskreten Merkmalsvektors auf Basis der individuellen Bewegungsereignisse, welche zuvor in Abschnitt 3.5.2.1 extrahiert wurden. Ziel ist es einerseits, die Informationen dieser Ereignisse für Algorithmen des maschinellen Lernens lesbar zu machen, andererseits sollen Ereignisse aber auch bestmöglich beschrieben werden, um ein Vorhersagemodell aussagekräftig und ohne Generalisierungsfehler zu trainieren (vgl. 2.3.2.1). Um die mit SensX erfassten Bewegungsinformationen im Rahmen der folgenden Evaluation auf ihre Brauchbarkeit hinsichtlich der Erkennung von Bewegungsaktivität zu evaluieren, werden sie in diesem Abschnitt in ihrer Dimensionalität reduziert und zu diskreten Merkmalsvektoren zusammengefasst, wobei jeder Merkmalsvektor ein individuelles Bewegungsereignis beschreibt.

Die dazu verwendete Eingangsinformation ist die Menge aus individuellen Ereignissen $E = \{e_1, \dots, e_j\}$, wobei alle e_x jeweils von $n = 30$ Einzelsignalen beschrieben werden. Für die initiale Verstärkung der eingehenden Sensorinformationen soll nun aus jedem Ereignisset e_x ein diskreter Merkmalsvektor der Länge l auf Basis der erfassten Zeitreihen erstellt werden. Eine Herausforderung ergibt sich dabei aus den je nach Sensorplattform unterschiedlichen Sampling-Raten, da l für jede später zu klassifizierende Ereignisdistanz als Vorbedingung des Einsatzes von Algorithmen des maschinellen Lernens gleich sein muss.

Eine Möglichkeit zur Erfüllung dieser Bedingung ist die Anwendung eines Resampling-Verfahrens auf alle Zeitreihen eines Ereignissets e_x auf Basis der Fourier-Transformation, wobei alle Parameter p_i innerhalb jedes der $n = 30$ Signale in e_x auf die gleiche Länge l entlang der X-Achse interpoliert werden [77].

Abbildung 3.6 verdeutlicht das Vorgehen anhand der Beschleunigungsinformation des Beins eines Probanden bei der Durchführung der Körpereigengewichtsübung *Crunches* in X-Richtung. Das ursprüngliche Signal, welches in Blau dargestellt ist, besitzt dabei über 70 Sampling-Punkte. Durch Anwendung des Resamplings wird das Signal entlang der X-Achse gestaucht; dessen beinhaltete Sampling-Punkte werden neu berechnet. Das interpolierte Signal besteht nun aus einer fixen Anzahl von l Sampling-Punkten, wobei die Startzeit des Signals noch immer t_0 , also der Startzeitpunkt des ursprünglichen

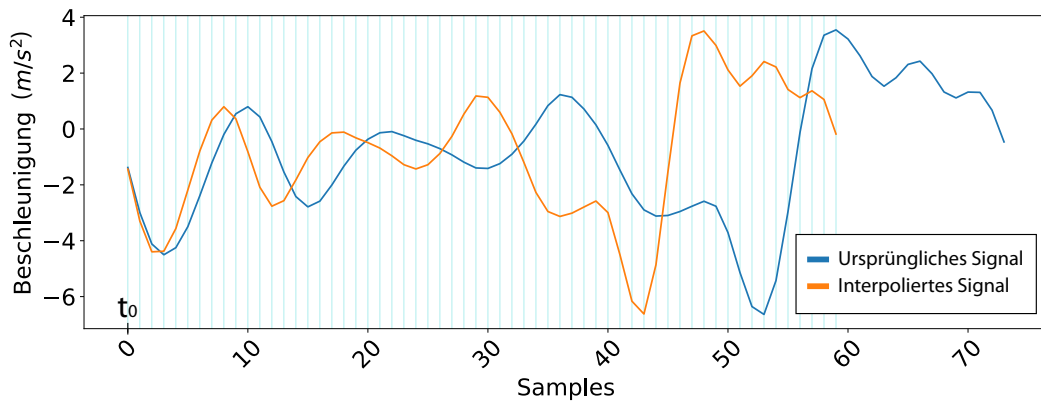


Abbildung 3.6: Resampling des am Bein eines Probanden in X-Richtung erfassten Beschleunigungssignals bei der Durchführung einer Wiederholung der Körpereigengewichtsübung *Crunches* entlang der X-Achse (vgl. 3.6.1).

Signals, ist. Die vorgegebene Sampling-Rate, welche im Rahmen des folgenden Anwendungsszenarios auf $l = 60$ festgelegt wurde, ergibt sich aus dem Zusammenspiel der technischen Kapazitäten des SensX-Sensorsystems, welches mit seinen externen Sensorplattformen die Bewegungen der menschlichen Gliedmaßen erfasst. Im Setup mit einem HTC One als zentrale Recheneinheit wird für die externen Plattformen eine durchschnittliche Sampling-Rate von 40 Hz realisiert, die maximale durchschnittliche Dauer der erfassten Körpereigengewichtsübungen betrug $t = 1,5\text{ s}$. Damit ergibt sich $l = 40\text{ Hz} * 1,5\text{ s} = 60$, wobei zugunsten des nahezu unveränderten Informationsgehalts der externen Sensorik die initial deutlich höhere Sampling-Rate der zentralen Recheneinheit etwa um Faktor 0,6 vermindert wird.

Anschließend wird jedes der n Signale innerhalb aller Ereignisse e_x auf die Länge l gesampelt. Anschließend wird aus diesen interpolierten Signalen für jedes e_x eine Ereignisinstanz I_x in Form eines geordneten Merkmalsvektors der Länge $l = n * 60$ aufgebaut, wobei jeder einzelne Sample-Parameter der interpolierten Signale einem Parameter im Merkmalsvektor entspricht. Die Menge aller Instanzen I_x soll folgend im Rahmen der Evaluation als Basis zum Training eines Vorhersagemodells mithilfe überwachter maschineller Lernprozesse dienen, wobei jeder Instanz ein Label in Form der ihr zugehörigen Ereignisklasse zugewiesen wird.

3.5.2.3 Klassifikation

Die zuvor extrahierte Menge aller Bewegungsinstanzen I_x wird im Rahmen des letzten Schrittes der in Abbildung 3.4 dargestellten Prozesskette mithilfe von maschinellen Lernprozessen untersucht. Dazu wurde im Rahmen von SensX das *WEKA* Machine-Learning-Framework genutzt, welches sich auch auf mobilen Endgeräten, wie in diesem Fall der *zentralen Recheneinheit* in Form

eines handelsüblichen Smartphones, implementieren lässt [82]. Dieses bietet nicht nur die Möglichkeit zu überwachten und nicht überwachten Lernvorgängen mittels unterschiedlicher Lernalgorithmen in Echtzeit, sondern auch zur Feature-Analyse, zum Feature-Ranking und zur automatisierten Optimierung von Hyper-Parametern [199, 73]. Dieses Setup wird im Folgenden genutzt, um die mit SensX erfassten Bewegungsinformationen näher zu untersuchen.

3.6 Evaluation

Im Rahmen der folgenden Evaluation wird zunächst die zur Vertestung von SensX durchgeführte Studie zur Erfassung von Bewegungsinformationen im Detail vorgestellt. Anschließend wird eine Einsicht in die Fähigkeit zur Erkennung von Bewegungsaktivität von SensX unter Berücksichtigung unterschiedlicher Sensorkonfigurationen gegeben. Zusätzlich werden auch Experimente hinsichtlich der Laufzeit während des Trainings von Vorhersagemodellen und der Klassifikation von einzelnen Bewegungsinstanzen durchgeführt. Anschließend werden diese Ergebnisse zusammen mit den an SensX gestellten Anforderungen aus Abschnitt 3.4.1 in Bezug zu den Eigenschaften verwandter Arbeiten gesetzt.

3.6.1 Studiendesign

Um die mithilfe von SensX erfassten Bewegungsinformationen auf ihren Informationsgehalt und deren Eignung zur Aktivitätserkennung hin zu überprüfen, wurde zunächst eine umfangreiche Studie zur Datenerhebung definiert und umgesetzt [53]. Diese umfasst in ihren Rohdaten die Beschleunigungs- und Rotationsinformationen von insgesamt 11.087 Wiederholungen von acht Körpereigengewichtsübungen und wurde mit 26 unterschiedlichen Studienteilnehmern durchgeführt.

Während der Studiendurchführung trugen die Probanden das SensX-Sensorsystem, analog zu dem in Abschnitt 3.5.1.4 beschriebenen Vorgehen. Bei den einzelnen Übungen handelte es sich in der Reihenfolge ihrer Durchführung geordnet um (1) *Crunches* (CR), (2) *Lunges* (LU), (3) *Hampelmann* (HA), (4) *Bicycle Crunches* (BC), (5) *Kniebeugen* (KN), (6) *Mountain Climber* (MC), (7) *Russian Twist* (RT) und (8) *Liegestützen* (LI). Einen Überblick über die einzelnen Übungen sowie deren Durchführung wird in Abbildung 3.7 gegeben. Für jede einzelne Übung mussten die Probanden drei Durchgänge mit 20 Einzelwiederholungen absolvieren, wobei nach jedem Durchgang eine Pflichtpause von 30s vorgegeben war. Die Reihenfolge der einzelnen Übungen wurde dabei so gewählt, dass immer unterschiedliche Muskelgruppen nacheinander beansprucht werden. So sollten zum einen Überlastungen verhindert werden, zum anderen sollte die Chance zum Abschluss des gesamten Trainingsprogramms durch möglichst viele einzelne Teilnehmer erhöht werden. Vor Durchführung

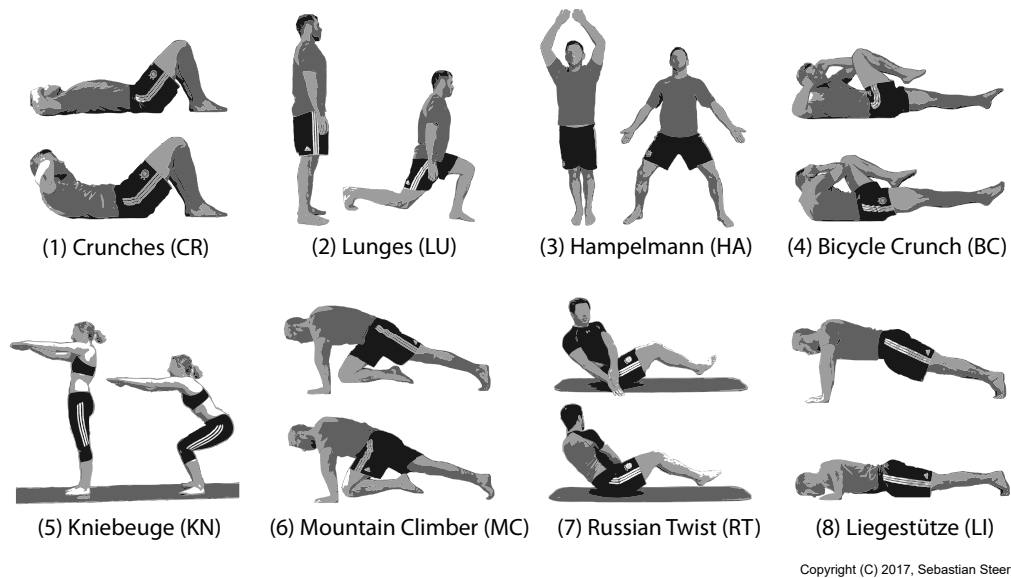


Abbildung 3.7: Übersicht über die acht im Rahmen der Studie durchgeführten und erfassten Körperereigengewichtsübungen, in Reihenfolge ihrer Durchführung und mit zugehörigen Abkürzungen.

jeder einzelnen Übungsart wurde den Teilnehmern ein professionelles Instruktionsvideo gezeigt, in welchem die genaue Übungsdurchführung erklärt wird. Weiterhin wurde die Durchführung aller Übungen von allen Probanden zum späteren qualitativen Labeling auf Video aufgezeichnet.

Neben den Bewegungsinformationen zu den einzelnen Übungen und den korrespondierenden Videoaufzeichnungen wurden auch Informationen zu Alter, Beruf, Geschlecht, Gewicht und grundlegender Fitness der einzelnen Probanden erhoben. Insgesamt nahmen 7 weibliche und 19 männliche Probanden an der Studie teil. Davon wurden in einem ersten Testlauf 2 Probanden (1 weiblich, 1 männlich) mit dem LG Nexus Setup aufgenommen und 6 weibliche und 18 männliche Teilnehmer im Rahmen der Hauptstudie mit der HTC One Konfiguration erfasst (vgl. 3.5.1.1). Das durchschnittliche Alter der Probanden lag bei 27,15 Jahren, das durchschnittliche Gewicht liegt bei 67,04 kg. Die älteste Studienteilnehmerin war zur Zeit der Studiendurchführung 53 Jahre alt, der jüngste Teilnehmer war zu diesem Zeitpunkt 20 Jahre alt. Zusätzlich sollten alle Teilnehmer im Vorfeld eine Einschätzung ihres persönlichen Fitness-Levels sowie zur Häufigkeit ihrer sportlichen Aktivität pro Woche vornehmen. Die Angaben zur sportlichen Aktivität pro Woche sind in Tabelle 3.1 nachzuvollziehen – der Großteil der Teilnehmer trieb zur Zeit der Studiendurchführung 2-3 Mal die Woche Sport. Ihr persönliches Fitness-Level sollten die Teilnehmer auf einer Skala von 1-5 bewerten (*keine Fitness* (1), *geringe Fitness* (2), *normale Fitness* (3), *gute Fitness* (4), *sehr gute Fitness* (5)) (vgl. Tabelle 3.2), wobei der Großteil der Probanden ihren Fitness-Zustand als normal bis gut einschätzte.

Sportliche Aktivitäten pro Woche	1	1-2	2	2-3	3	3-4	4	4-5	5
Anzahl Probanden	6	-	12	3	2	-	2	-	1

Tabelle 3.1: Anzahl der sportlichen Aktivität pro Woche nach Teilnehmern der durchgeführten Studie zur Datenerhebung.

Level der Sportlichkeit	1	1-2	2	2-3	3	3-4	4	4-5	5
Anzahl Probanden	2	-	1	-	9	2	9	1	2

Tabelle 3.2: Selbsteinschätzung des Fitness-Levels der Studienteilnehmer auf einer Skala von 1-5 (*sehr unsportlich* (1), *unsportlich* (2), *normal* (3), *sportlich* (4), *sehr sportlich* (5)).

3.6.2 Untersuchung unterschiedlicher Sensorkonfigurationen

In diesem Abschnitt wird nun die Eignung der, mithilfe des SensX-Sensorsystems erfassten, Bewegungsinformationen zur Erkennung von menschlicher Bewegungsaktivität auf feingranularem Level und für unterschiedliche Sensorkonfigurationen untersucht. Dabei gilt es herauszufinden, auf Basis welcher, der individuellen Sensorplattformen innerhalb des verteilten Systems, besonders gute Ergebnisse in der Aktivitätserkennung erzielt werden können und wie sich das System für unterschiedliche Sensorkombinationen verhält. Bei den unterschiedlichen, zu erkennenden Bewegungsaktivitäten handelt es sich um die einzelnen Klassen der Körpereigengewichtsübungen. Dazu wurden zunächst verschiedene Untergruppen von Sensorkombinationen definiert, welche in Tabelle 3.3 dargestellt sind. Anschließend wurden alle Untergruppen einmal zusammen mit dem Brust-Sensor und einmal ohne Brustsensor im Rahmen eines Prozesses des überwachten maschinellen Lernens mit dem NB-Klassifikator klassifiziert, welcher in verwandten Arbeiten zur Erkennung menschlicher Aktivität bereits mit guten Ergebnissen eingesetzt wurde [178, 212, 188].

Grundlage der folgenden Untersuchungen sind die Bewegungsinformationen der in Abschnitt 3.6.1 beschriebenen Hauptstudie mit 24 Probanden, von denen analog zur Reihenfolge der Übungsdurchführung die ersten 6 Übungen nach den in Abschnitt 3.5.2 definierten Verarbeitungsschritten zur Segmentierung und Bildung individueller Merkmalsvektoren für einen überwachten Lernprozess vorbereitet wurden. Die Daten der zwei im Rahmen von Vortests mit einer anderen Hardware-Konfiguration erfassten Probanden waren nicht teil dieser Evaluation (vgl. Abschnitt 3.6.1). Insgesamt wurden von 7.534 erfassten Wiederholungen aller Bewegungsereignisse 7.473 Instanzen I_x extrahiert, denen jeweils ein Label l_x zugewiesen wurde. Dieses korrespondiert mit der Klasse der jeweiligen Körpereigengewichtsübung. Die Instanzen teilen sich in 1.317 Wiederholungen von *Crunches*, 1.372 *Lunges*, 1.396 Wiederholungen

Abkürzung	Linke Hand	Rechte Hand	Linker Fuß	Rechter Fuß
Alle	✓	✓	✓	✓
Oben Links (OL)	✓	×	×	×
Oben Rechts (OR)	×	✓	×	×
Oben	✓	✓	×	×
Unten Links (UL)	×	×	✓	×
Unten Rechts (UR)	×	×	×	✓
Unten	×	×	✓	✓
Links (L)	✓	×	✓	×
Rechts (R)	×	✓	×	✓

Tabelle 3.3: Aufschlüsselung der einzelnen untersuchten Sensorkonfigurationen nach beinhalteten Sensorplattformen und zugehörigen Abkürzungen.

der Übung *Hampelmann*, 692 *Bicycle Crunches*, 1.379 *Kniebeugen* und 1.317 Wiederholungen der Übung *Mountain-Climber* auf. Die Menge der extrahierten Ereignisse für die Übung *Bicycle Crunch* beinhaltet dabei nur etwa halb so viele Wiederholungen, da auf Grund der gespiegelten Durchführung des gleichen Bewegungsablaufs eine linksseitige und eine rechtsseitige Ausführung zusammen als eine Wiederholung gezählt werden. Die Länge der einzelnen I_x ergibt sich aus der Anzahl der innerhalb der untersuchten Sensorkonfiguration enthaltenen Sensorplattformen (jeweils Accelerometer und Gyroskop) für jeweils X-, Y-, und Z-Richtung und multipliziert sich mit Faktor 60, welcher der interpolierten Sampling-Rate jedes einzelnen Bewegungsereignisses nach dem Vorgehen in Abschnitt 3.5.2.2 entspricht. Zusätzlich erhält jede Instanz I_x noch eine Stelle für die Zuordnung zu seiner Ereignisklasse l_x . Alle Untersuchungen innerhalb dieses Abschnittes wurden im Rahmen einer *Training-Set-Test-Set-Validierung* (TS/TS) mit einer 80:20 Teilung der zu untersuchenden Instanzen nach Ng durchgeführt [174].

Tabelle 3.4 und Abbildung 3.8 zeigen die prozentualen Werte zur gemittelten erfolgreichen Erkennung der unterschiedlichen Bewegungsereignisse, aufgeschlüsselt nach unterschiedlichen Sensorkonfigurationen und jeweils mit sowie ohne zusätzliche Information des Brustsensors. Diese zeigen, dass alle Konfi-

Brust-Sensor	OL (%)	OR (%)	Oben (%)	UL (%)	UR (%)	Unten (%)	Links (%)	Rechts (%)	Alle (%)	Brust (%)
mit	97,86	98,19	97,93	97,99	99,4	98,19	98,81	99,16	99,13	94,31
ohne	94,72	93,78	94,98	97,59	98,06	97,66	98,60	97,66	98,92	-

Tabelle 3.4: Durchschnittliche korrekte Erkennungsrate (in Prozent) der einzelnen Körpereigengewichtsübungen mit SensX nach unterschiedlichen Sensorkonfigurationen (vgl. Tabelle 3.3).

gurationen bis auf OR ohne Brustinformation eine durchschnittliche Aktivi-

tätserkennungsrate von mehr als 94% erreichen. Die Informationen des Brustsensors alleine führen zu einer vergleichsweise schlechten Performanz, jedoch steigert deren Kombination mit der Information von externen Sensorplattformen die Erkennungsrate in allen Fällen erheblich. Kombinationen, welche nur die Information eines Armes (OR, OL) oder des Oberkörpers (Oben) enthalten, erzielen sichtbar schlechtere Ergebnisse als Kombinationen, welche nur die Informationen der Beine (UR, UL) oder des Unterkörpers (Unten) enthalten. Generell sehr gute Ergebnisse liefern Kombinationen einer gesamten Körperhälfte zusammen mit den Informationen des Brustsensors (Links, mit 98,81% und Rechts mit 99,16 %). Die Verwendung der Informationen aller externen

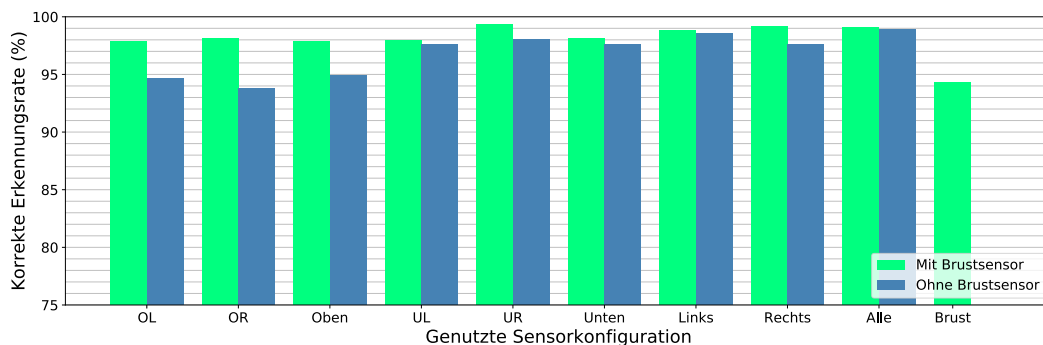


Abbildung 3.8: Visualisierung der durchschnittlichen korrekten Erkennungsrate (in Prozent) der einzelnen Körpereigengewichtsübungen mit SensX nach unterschiedlichen Sensorkonfigurationen (vgl. Tabelle 3.3)

Sensorplattformen zusammen (Alle) führt mit 98% gemittelter Erkennungsrate ohne Informationen des Brustsensors und mit 99.13% zusammen mit diesen, im Durchschnitt zu den besten Ergebnissen. Einen positiven Ausreißer bildet die gemittelte Erkennungsrate von UR im Zusammenspiel mit den Informationen des Brustsensors, welche mit 99.4% das beste Ergebnis innerhalb dieser Testreihe war. Im folgenden Abschnitt werden nun die Konfigurationen mit den drei besten Ergebnissen, welche alle eine durchschnittliche Erkennungsrate von mehr als 99% aufweisen, näher untersucht.

3.6.3 Erkennungsperformanz nach Ereignisklassen

Folgend werden die unterschiedlichen Sensorkonfigurationen mit den drei besten gemittelten Aktivitätserkennungsraten aus Abschnitt 3.6.2 im Detail beleuchtet. Hierbei wird auch ein Einblick in die Performanz bei der Klassifizierung einzelner Klassen von Bewegungsereignissen je ausgewählter Sensorkonfiguration und je ausgewähltem Validierungsverfahren gegeben. Zusätzlich zum vorherigen TS/Ts Experiment mit einer 80:20 Teilung der zu untersuchenden Instanzen wurde auch eine *10-fach gefaltete Kreuzvalidierung* (10fCV) durchgeführt. Die Ergebnisse hierzu sind in Tabelle 3.5 nachzuvollziehen.

Konfiguration	CR (%)	LU (%)	HA (%)	BC (%)	KN (%)	MC (%)	Ø (%)
Alle & Brust (10fCV)	97,0	100	100	100	100	97,8	99,09
Alle & Brust (TS/TS)	96,6	100	100	100	100	98,5	99,13
UR & Brust (10fCV)	97,0	100	100	99,0	99,3	100	99,25
UR & Brust (TS/TS)	96,6	100	100	100	100	100	99,4
Rechts & Brust (10fCV)	98,4	98,6	100	100	100	97,7	99,05
Rechts & Brust (TS/TS)	98,9	99,6	100	100	100	96,7	99,16

Tabelle 3.5: Detaillierte Übersicht über die Erkennungsraten nach Ereignisklassen für die drei performantesten Sensorkonfigurationen aus Abschnitt 3.6.2 im Rahmen einer 10fCV und einer TS/TS.

Deutlich erkennbar ist die sehr gute Erkennungsrate für Bewegungsereignisse, bei welchen entweder alle Extremitäten des menschlichen Körpers oder zumindest ein Großteil derer, klar unterscheidbare Bewegungen mit hohem Energieaufwand durchführen. Ein Beispiel hierfür ist die Übung *Hampelmann*, bei welcher Arme und Beine abwechselnd weit ausholende Bewegungen durchführen, zusätzlich erfasst auch der Brustsensor durch das stetige Hüpfen eines Probanden einen eindeutigen Signalausschlag. Die Übung *Crunches* hingegen wurde für alle Konfigurationen mit einer durchschnittlichen Erfolgsrate von 97,42% verhältnismäßig schlecht erkannt. Dies könnte daran liegen, dass hier nur der Brustsensor die Hauptinformationen zum Bewegungsablauf liefert und die anderen Sensorplattformen während der Bewegungsdurchführung verhältnismäßig wenig Dynamik im Signal verzeichnen. Die Performanz ist im Rahmen der 10fCV durchschnittlich immer etwas niedriger als die der korrespondierenden TS/TS Experimente, wobei allerdings keine Ausreißer entstehen.

3.6.4 Laufzeituntersuchung

Wie in Abschnitt 3.4.1 erwähnt, ist die Fähigkeit zur Untersuchung von Bewegungsereignissen sowie das Erlernen von neuen Vorhersage-Modellen in Echtzeit ein zentrales Anliegen innerhalb der Anforderungen an SensX. Dabei ist eine kurze Laufzeit zum Training von Vorhersagemodellen sowie zur Klassifikation von Ereignissen aber nicht nur für die generelle Anwendbarkeit des Systems relevant, sondern auch für die Akzeptanz des Konzepts seitens potentieller Nutzer sowie für die Systemlaufzeit. Arbeiten zum Energieverbrauch auf mobilen Endgeräten zeigen, dass gerade eine hohe Nutzung von Speicher- und Rechenleistung sich signifikant auf deren Energiehaushalt und damit ebenso auf die potentielle Systemlaufzeit auswirken [47, 48]. Aus diesen Gründen empfiehlt es sich, die benötigten Lauf- und Rechenzeiten für oben genannte Aufgaben so gering wie möglich zu halten.

Um diese Aspekte zu untersuchen, wurden die Laufzeiten zur Klassifikation einzelner Instanzen und zum Training des Vorhersagemodells sowohl im mobilen als auch im stationären Szenario untersucht. Alle Experimente zur Laufzeit wurden unter Einbezug der Informationen von allen verfügbaren Sensoren des SensX-Sensorsystems durchgeführt. Das stationäre Szenario entspricht dabei

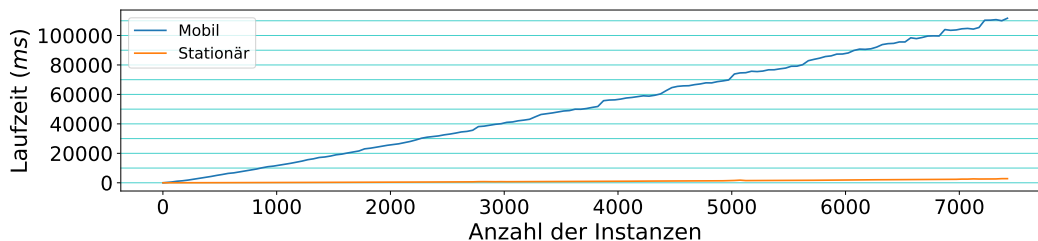


Abbildung 3.9: Trainingsdauer von Vorhersagemodellen mithilfe des NB-Klassifikators im mobilen und im stationären Szenario.

dem Offline-Fall, welcher während der Entwicklungsphase bei dem Vortrainieren von Vorhersagemodellen gegeben ist. Im mobilen Szenario wurde der Online-Fall simuliert, bei welchem ein Nutzer Bewegungsereignisse in Echtzeit klassifiziert und bei Bedarf Vorhersagemodelle für neue Bewegungsereignisse nachtrainieren kann. Das stationäre Szenario wurde auf einem Notebook (Lenovo Think-Pad mit Intel Core i7-6700HQ, 2,6 GHz Octa-Core Prozessor und 16 GB Arbeitsspeicher) getestet, das mobile Szenario mit einem Smartphone (Motorola Moto G5 mit einem 1,4 GHz Snapdragon 430 Octa-Core Prozessor und 3 GB Arbeitsspeicher) als *zentrale Recheneinheit* des SensX-Sensorsystems.

Abbildung 3.9 gibt einen Überblick über die Laufzeit während des Modell-Trainings mithilfe des NB-Klassifikators für beide Szenarien. Hier ist zu sehen, dass die Laufzeit für beide Szenarien nahezu linear mit steigender Größe des Trainingsdatensets ansteigt, wobei der Anstieg im mobilen Szenario deutlich signifikanter ist als im stationären. Für das Training mit allen 7.473 Instanzen werden im stationären Szenario 1,36 s (durchschnittlich 0,18 ms pro Instanz) benötigt, während im mobilen Szenario insgesamt 114,477 s und durchschnittlich 15,319 ms Trainingszeit pro Instanz benötigt werden. Im Anwendungsfall ist es allerdings unrealistisch, mehrere tausend Wiederholungen von einzelnen Ereignisklassen zu verwenden, um diese neu zu lernen. Realistische Größen sind hier im zweistelligen bis unteren dreistelligen Bereich anzusiedeln, in welchem das mobile Setup nach Abbildung 3.9 maximal 8 s benötigt. Wie weitere Untersuchungen mit Verwendung von komprimierten Merkmalsvektoren im Folgekapitel unter 4.5.4 zeigen, genügen schon etwa 25 Instanzen je Ereignisklasse, um ein Vorhersagemodell mit einer korrekten Erkennungsrate von mehr als 98% zu implementieren. Für die Erkennung einzelner Bewegungsereignisse werden im stationären Szenario im Durchschnitt 2,31 ms benötigt, im mobilen Szenario 70,059 ms. Hier ist noch zu vermerken, dass die Klassifikation im Online-Fall

erst nach vollendeter Durchführung eines segmentierten Bewegungsereignisses beginnt. Daraus ergibt sich im Fall der vorliegenden Evaluation ein zusätzliches $\Delta t = 1,5 \text{ s}$ ab Ereignisstart t_s , womit die Gesamtdauer t einer Klassifikation bei etwa $1,502 \text{ s}$ liegt.

Die präsentierten Ergebnisse zeigen, dass das SensX-Sensorsystem Bewegungsereignisse auf Basis von multidimensionalen Zeitreihen mit einer Laufzeit von wenigen Millisekunden nahezu in Echtzeit und mit großem Erfolg erkennt. Weitergehend ist es innerhalb weniger Sekunden rekonfigurierbar und kann Vorhersagemodelle für neue Ereignisse ohne Zuhilfenahme externer Ressourcen nachtrainieren.

3.6.5 Einordnung der Ergebnisse

In den vorhergehenden Abschnitten der Evaluation wurde die Performanz des SensX-Sensorsystems hinsichtlich seiner Fähigkeiten zur Aktivitätserkennung komplexer Bewegungsereignisse, welche von der ganzheitlichen Bewegung des menschlichen Körpers und aller seiner Extremitäten beschrieben werden, demonstriert. Zusätzlich wurden auch Untersuchungen zu den Laufzeiten während des Trainings von Vorhersagemodellen zum überwachten maschinellen Lernen und der Klassifikation einzelner Bewegungsereignisse angestellt, um die Echtzeitfähigkeit des Systems zu validieren. Nun wird noch ein kurzer Vergleich zu den Leistungsmerkmalen verwandter Arbeiten zur Untersuchung menschlicher Aktivität und komplexer Bewegungsereignisse vorgenommen. Dieser Vergleich wird mithilfe der zuvor in Abschnitt 3.4.1 definierten, für SensX relevanten Anforderungen und dem diesem Kapitel zugrundeliegenden Anwendungszweck, durchgeführt. Dazu wurden alle verwandten Arbeiten auf Basis des beschriebenen Ist-Zustands in ihrer jeweiligen Publikation eingeordnet. Aufgrund unterschiedlicher Datengrundlagen und Validierungsmethoden zur Evaluation sowie der verschiedenen Anwendungszwecke der einzelnen verwandten Arbeiten sind endgültige Aussagen in dieser Hinsicht nur bedingt möglich. Hinsichtlich der durchschnittlichen korrekten Erkennungsrate einzelner Bewegungsereignisse wurde für alle verwandten Arbeiten jeweils das beste Ergebnis ausgewählt, welches hinsichtlich seiner zugrundeliegenden Validierungsmethodik der in dieser Arbeit genutzten Methodik am ähnlichsten ist. Tabelle 3.6 gibt hierzu einen Überblick.

Recofit ist in dieser Hinsicht das einzige direkt vergleichbare System zu SensX [139]. Dieses arbeitet ebenfalls auf Basis von Bewegungsinformationen und Untersuchungsgegenstand in der Evaluation sind ebenso Körpereigengewichtsübungen. FEMO beschäftigt sich mit der Erkennung von Gewichthebeübungen auf Basis von RFID-Tags. TDD implementiert einen Ansatz zur Analyse menschlicher Aktivität auf Basis visueller Informationen mithilfe von Deep-Learning Algorithmen [42, 204]. TDD umfasst hierbei keine eigene Hardware, sondern ist lediglich ein logischer Ansatz, welcher auf frei verfügbaren Videodatensätzen als Benchmarks evaluiert wurde.

	SensX	Recofit [139]	DCNN+ [91]	FEMO [42]	TDD [204]
Sensor-Typ	Bewegung	Bewegung	Bewegung	RFID	Visuell
Unabhängigkeit und Mobilität	✓	×	×	×	×
Ganzheitlichkeit	✓	×	×	×	(✓)
Generalisierbarkeit und Erweiterbarkeit	✓	✓	(✓)	✓	×
Einfachheit und Robustheit	(✓)	×	×	×	×
Echtzeitfähigkeit	✓	✓	(✓)	✓	(✓)
Konventionalität	✓	✓	✓	✓	✓
Erkennungsrate (\emptyset in %)	99,4 (komplex)	98,2 (komplex)	99,93 (einfach)	90,0 (einfach)	91.5 (komplex)

Tabelle 3.6: Vergleich der Spezifikationen und Leistungsmerkmale verschiedener Systeme zur Erkennung menschlicher Aktivität.

Durch seine integrierte Recheneinheit und die Fähigkeit zur Klassifizierung und zum Modell-Training in Echtzeit ist SensX das einzige System, welches unabhängig von stationärer Infrastruktur und Laborbedingungen eingesetzt werden kann und somit die Anforderungen bezüglich *Unabhängigkeit und Mobilität* erfüllt. Alle anderen Ansätze benötigen zumindest eine stationäre Recheneinheit.

Auch in puncto *Ganzheitlichkeit* ist SensX das einzige System, welches alle Gliedmaßen des menschlichen Körpers in den Erfassungsprozess von Bewegung miteinbezieht. Lediglich Ansätze wie TDD, welche auf visuellen Informationen arbeiten, können hier eine ähnliche Abdeckung liefern, allerdings leiden diese bei Einsatz von nur einer Kamera an Verdeckungseffekten.

Generalisierbarkeit und Erweiterbarkeit bezieht sich in diesem Kontext auf die Fähigkeit, bei Bedarf auch neue Bewegungsereignisse mit nutzerverträglichem Aufwand zu erlernen und neue Analysemethoden in das System zu integrieren. Diese Anforderung wird nur von TDD nicht erfüllt, da die Dauer zum Modell-Training mithilfe des verwendeten *Deep Convolutional Neural Networks* aktuell mit einer Zeitspanne von bis zu mehreren Tagen zu zeitaufwändig ist [115]. DCNN+ basiert auf einer ähnlichen Analysestruktur, wobei die Eingangsinformationen in Form von einzelnen, vorverarbeiteten Aktivitätsabbildern deutlich weniger komplex sind als die in TDD genutzten Videosequenzen. Deshalb wird die Anforderung für DCNN+ als eingeschränkt erfüllt angesehen.

Die Forderung nach *Einfachheit und Robustheit* wird mit Einschränkungen nur von SensX erfüllt, da es autark von einem Nutzer alleine mithilfe der *zentralen Recheneinheit* während des Einsatzes gesteuert werden kann. Während der Studiendurchführung zeigte sich, dass SensX das Bewegungspotential des Trägers nicht einschränkt und es auch in potentiell widrigen Szenarien, wie beispielsweise zur Untersuchung von Kletterbewegung (vgl. Kapitel 5), eingesetzt werden kann. Genannte Einschränkungen ergeben sich aus dem aktuell

ungeschützten Display der *zentralen Recheneinheit*, welches ohne zusätzliche Schutzmaßnahmen für starke, direkte Stöße anfällig ist.

Die Forderung nach *Echtzeitfähigkeit* wird hinsichtlich der Erkennung von Bewegungsereignissen in der Theorie von allen Ansätzen erfüllt, wobei TDD und DCNN+, wie zuvor angemerkt, aufgrund des rechenintensiven Vorgangs zur Modellbildung bei Bedarf keine oder nur eingeschränkte Rekonfiguration zum Erlernen neuer Bewegungsereignisse mit annehmbarer Laufzeit unterstützen.

Hinsichtlich der *Konventionalität* erfüllen alle Systeme die gestellten Anforderungen, da soweit nachvollziehbar nur handelsübliche und frei verfügbare Komponenten zur Umsetzung eingesetzt wurden.

Bezüglich der Performanz in der Erkennung von Bewegungsaktivitäten sind SensX und Recofit in etwa miteinander vergleichbar, wobei SensX für die gewählten Evaluations-Szenarien um 1,2% bessere Ergebnisse liefert. DCNN+ liefert die nominell besten Ergebnisse hinsichtlich der Erkennung menschlicher Aktivität, allerdings wurden im Rahmen der Evaluation auch nur einfache Aktivitäten wie Gehen, Laufen und Stehen untersucht, ohne dabei auf einzelne Bewegungsereignisse innerhalb einer Bewegungsaktivität einzugehen. FEMO erreicht eine korrekte Erkennungsrate von 90,0% für einfache Bewegungsereignisse, bei welchen nur ein oder zwei Extremitäten derselben Klasse (in diesem Fall die Arme) jeweils zur Durchführung notwendig sind. TDD erkennt 91,5% der Aktivitäten innerhalb des UCF101-Datensatzes, welcher gelabelte Videos von sowohl komplexen als auch von einfachen Bewegungen enthält [194]. Allerdings sind diese Ergebnisse aufgrund stark unterschiedlicher Validierungsverfahren und der unterschiedlich gelagerten Anwendungsfälle nicht direkt mit denen von SensX zu vergleichen.

3.7 Zusammenfassung

Ziel dieses Kapitels war die Konzeption, Umsetzung und Evaluation eines verteilten Sensorsystems, welches auf Basis von Bewegungsinformationen in Form von multidimensionalen Zeitreihen menschliche Bewegungsereignisse erfassen und erkennen kann. Mit SensX wurde dazu ein Konzept entwickelt und präsentiert, welches allen in Abschnitt 3.4.1 definierten Anforderungen, namentlich *Unabhängigkeit und Mobilität, Ganzheitlichkeit, Generalisierbarkeit und Erweiterbarkeit, Einfachheit und Robustheit, Echtzeitfähigkeit, und Konventionalität*, entspricht.

Dieses umfasst neben der Erfassung von Bewegungsinformationen auf der Hardware-Ebene auch Konzepte zur Segmentierung von multidimensionalen Zeitreihen, zum Feature-Engineering und zur Klassifikation von Bewegungsereignissen auf der logischen Ebene. Zur Evaluation des Systems wurde eine umfangreiche Studie durchgeführt, in deren Rahmen über 11.000 Wiederholungen von 8 Körpereigengewichtsübungen mit 26 Studienteilnehmern erfasst wurden. Anschließend wurden die Fähigkeiten des Systems zur Erkennung komplexer

menschlicher Bewegungsabläufe, feingranular und nach unterschiedlichen Sensorkonfiguration aufgeteilt, evaluiert. Die besten Ergebnisse belaufen sich auf eine durchschnittliche korrekte Erkennungsrate von komplexer menschlicher Aktivität von 99,4% im TS/TS-Validierungsszenario und 99,25% im Rahmen einer 10fCV-Validierung. Diese sind damit um 1,2% besser als die Ergebnisse direkt vergleichbarer Arbeiten und als kompetitiv zum aktuellen State-of-the-Art zu bezeichnen. Zusätzlich hat SensX die Fähigkeit zur Klassifizierung von Bewegungsereignissen in nahezu Echtzeit mit einer Klassifikationslaufzeit von durchschnittlich 70,059 *ms* im mobilen Online-Szenario. Die Rekonfiguration und das Erlernen neuer Bewegungen ist, je nach Anzahl der zum Training genutzten Bewegungsinstanzen, innerhalb weniger Sekunden und damit ebenso nahezu in Echtzeit möglich. Nach aktuellem Stand ist SensX damit das bislang erste System, welches die ganzheitliche Erfassung und Erkennung menschlicher Bewegung außerhalb von Laborbedingungen und ohne Zuhilfenahme sonstiger externer Infrastruktur auch für widrige Anwendungsszenarien ermöglicht.

Generell liegt der Anwendungsfokus von SensX auf der Erfassung und Untersuchung der Bewegungen einzelner menschlicher Personen. In anderen Anwendungsszenarien, etwa der Überwachung von großen Menschenmengen oder im Sicherheitsbereich, sind nach wie vor visuelle Analyse-Konzepte, wie beispielsweise TDD, generell besser zur Aktivitätserkennung geeignet [204, 189]. Zwar fällt hier oft eine signifikant längere Zeit zum Training eines Vorhersagemodells an, dafür können diese aber generischer für unterschiedlichste Aktivitäten und ohne auf spezifische Ereignisse abgestimmtes Feature-Engineering angewendet werden. Zusätzlich müssen einzelne Individuen für visuell arbeitende Systeme im Normalfall auch nicht mehr mit einzelnen Sensoren oder Markern bestückt werden, sondern können extern erfasst werden.

In den folgenden Kapiteln werden nun Konzepte zur qualitativen Untersuchung menschlicher Bewegungsaktivität entwickelt und vorgestellt, welche über die reine quantitative Untersuchung und die daraus resultierende Aktivitätserkennung dieses Kapitels hinausgehen. Diese Konzepte arbeiten auf Basis der Sensorinformationen, welche mithilfe des in diesem Kapitel vorgestellten SensX-Sensorsystems erfasst werden.

4 Qualitative Analyse von wiederkehrender Bewegungsaktivität

Im vorhergehenden Kapitel stand die ganzheitliche Erfassung und Erkennung von menschlicher Aktivität mithilfe von Bewegungsinformationen im Vordergrund. In diesem Kapitel sollen diese Konzepte nun mit dem Fokus auf der qualitativer Untersuchung von individuellen, wiederkehrenden Bewegungsereignissen erweitert werden. Diese Ereignisse sind im Rahmen dieser Arbeit analog zu Kapitel 2.1.1 als solche definiert, welche innerhalb einer Klasse von Bewegungsaktivität in vorhersehbarer zeitlicher Abfolge wiederkehrend auftreten. Die qualitative Untersuchung soll in diesem Kontext über die reine Unterscheidung von Ereignissen unterschiedlicher Bewegungsaktivitäten hinausgehen und automatisiert Aussagen hinsichtlich der Ausführungsqualität einzelner Bewegungsereignisse in Echtzeit ermöglichen, um so zum Beispiel Fehlstellungen in Bewegungsabläufen zu identifizieren und die Möglichkeit zur Korrektur zu bieten.

Dazu wird nach einem kurzen Überblick über zugrundeliegende Vorarbeiten in Abschnitt 4.1 anschließend in Abschnitt 4.2 auf die Motivation und die Herausforderungen dieses Kapitels eingegangen. Abschnitt 4.3 stellt verwandte Arbeiten aus den Bereichen der qualitativen Untersuchung von menschlicher Bewegung und zur Extraktion von Ereignissen aus kontinuierlichen Sensorinformationen vor. Auf dieser Basis werden in Abschnitt 4.4 Konzepte zur adaptiven Segmentierung von wiederkehrenden Bewegungsereignissen, zur Formalisierung von qualitativen, subjektiven Beobachtungen in diesen sowie zur qualitativen Untersuchung unter Zuhilfenahme von Ansätzen des überwachten maschinellen Lernens entwickelt. Diese werden nachfolgend im Rahmen der Evaluation in Abschnitt 4.5 unter verschiedenen Gesichtspunkten betrachtet. Wichtigstes Augenmerk liegt dabei auf den Fähigkeiten des vorgestellten Konzepts hinsichtlich der qualitativen Bewegungsanalyse. Im Vergleich zum Vorkapitel werden darüber hinaus auch signifikante Verbesserungen bezüglich der Laufzeit des Systems während der Rekonfiguration und Klassifizierung von einzelnen Bewegungsinstanzen erreicht. Diese Verbesserungen werden in der Evaluation in Abschnitt 4.5.3 mit den Ergebnissen zur reinen Aktivitätserkennung aus dem Vorkapitel 3.6.3 verglichen und erneut auf den Prüfstand gestellt. Weitergehend wird ein Einblick in die Auswirkungen von adaptiven Segmentierungsverfahren auf den Prozess der qualitativen Analyse von Bewe-

gung gegeben. In Abschnitt 4.6 werden die Ergebnisse dieses Kapitels zusammengefasst und in den Kontext der gesamten Arbeit gestellt. Hierbei werden auch Limitierungen und offene Fragestellungen diskutiert, welche dann im Anschluss in Kapitel 5 näher betrachtet werden.

4.1 Vorveröffentlichungen

Die Kerninhalte dieses Kapitels wurden vom Autor bereits in [46] publiziert. Der im Rahmen der Evaluation in Abschnitt 4.5 verwendete Datensatz sowie das Vorgehen zur Formalisierung von qualitativen Beobachtungen in Abschnitt 4.4.3 wurden analog zum Vorkapitel bereits im Rahmen von [53] veröffentlicht. Wie in Kapitel 1.2 dargestellt, stammen alle in den Vorveröffentlichungen und im Folgenden präsentierten Inhalte bezüglich der Idee, des Konzepts und der Evaluation vom Autor der vorliegenden Arbeit. Die Untersuchungen zur System-Laufzeit in Abschnitt 4.5.4 unter Verwendung des in diesem Kapitel vorgestellten Konzepts sowie der Performanzvergleich verschiedener Segmentierungsansätze in Abschnitt 4.5.2 wurden zuvor noch nicht veröffentlicht. Die Tabellen 4.1 und 4.2 und die Abbildungen 4.3 und 4.5 sind bereits in [46] enthalten. Abbildung 4.1 ist bereits in [51] enthalten.

4.2 Motivation und Herausforderungen

Die Popularität von sensorgestützten Anwendungen zur Erfassung von Bewegungsaktivität und zur unabhängigen Planung von Trainingseinheiten im Anwendungsszenario Sport ist in den letzten Jahren stark angestiegen. Neben rein organisatorischen Funktionen stellen diese Applikationen oft auch Instruktionen für anspruchsvolle Übungen und Bewegungsabläufe bereit, welche durch fehlerhafte Ausführung von Bewegungen, durch Überschätzung oder durch Überlastung besonders bei Ungeübten und Laien leicht zu ernsthaften Verletzungen führen können [93, 120, 109, 2]. Um diese Problematik zu adressieren und auch ohne aktive Überwachung von Bewegungsabläufen durch menschliches Personal automatisiert qualitatives Feedback zu individuellen Bewegungsereignissen geben zu können, ist ein Konzept zur qualitativen Analyse notwendig. In verwandten Arbeiten existieren hierzu bereits vereinzelt Ansätze, welche unterschiedliche Teilbereiche der Thematik betrachten. Ein ganzheitliches Konzept, welches die qualitative Analyse wiederkehrender Bewegung generisch und unabhängig von stationärer Infrastruktur betrachtet, fehlt jedoch bislang. Die Entwicklung und Evaluation eines solchen Konzepts ist der Schwerpunkt dieses Kapitels.

Aus dieser Aufgabenstellung ergeben sich verschiedene Herausforderungen. Die Erste ist hierbei die Entwicklung eines adaptiven Verfahrens zur Extraktion von Einzelereignissen innerhalb einer Sequenz von wiederkehrenden Ereignissen der gleichen Aktivitätsklasse. Die Notwendigkeit dazu ergibt sich

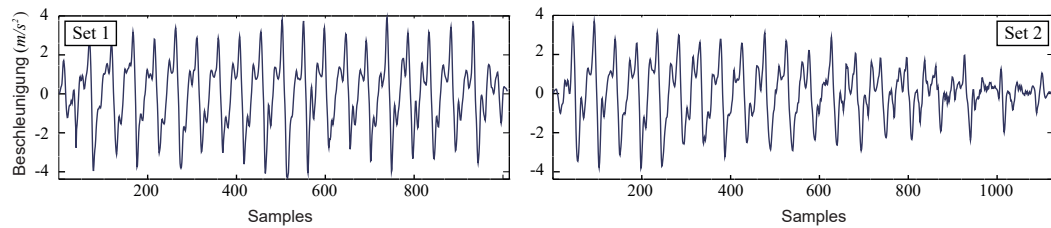


Abbildung 4.1: Beschleunigung des Brustsensors eines Probanden in Z -Richtung für zwei nacheinander durchgeführte Sets der Körpereigengewichtsübung *Crunches* mit jeweils 20 Wiederholungen.

daraus, dass bei der qualitativen Analyse, im Gegensatz zur reinen Aktivitätserkennung, isolierte Einzelereignisse betrachtet und bewertet werden sollen, um anschließend spezifisch über jedes einzelne Ereignis unabhängig von vorherigen oder nachfolgenden Instanzen eine qualitative Aussage treffen zu können. Abbildung 4.1 illustriert in diesem Kontext die Forderung nach adaptiver Segmentierung. Zu sehen ist hier das leicht geglättete Beschleunigungssignal in Z -Richtung (Auf- und Abbewegung der menschlichen Brust) zweier nacheinander durchgeführter Sets der Körpereigengewichtsübung *Crunches* mit jeweils 20 Wiederholungen. Dabei sind nicht nur die einzelnen Wiederholungen deutlich zu erkennen. Zusätzlich lässt sich vor allem in Set 2 anhand der immer länger werdenden Dauer eines periodischen Durchlaufs eine Ermüdung des Probanden erkennen (siehe Verlängerung des Periodenumlaufs des Signals entlang der X -Achse und dessen verminderte Beschleunigung entlang der Y -Achse). Grundannahme ist nun, dass eine qualitative Analyse jeweils für die exakte Dauer jedes individuellen Ereignisses erfolgen muss. So sollen Verzerrungen der Untersuchungsergebnisse durch unvollständige Ereignisbetrachtung oder durch zusätzliche Betrachtung von nicht einem Ereignis zugehörigen Informationen ausgeschlossen werden. Daraus ergibt sich die Notwendigkeit zu einem adaptiven Segmentierungsalgorithmus, welcher fähig ist wiederkehrende Ereignisse vollständig isoliert von benachbarten Ereignissen zu extrahieren und für weitere Untersuchungen bereitzustellen. Eine weitere Herausforderung stellt sich aus der Forderung nach qualitativer Analyse. Um diese mithilfe der Werkzeuge des überwachten maschinellen Lernens zu ermöglichen, müssen subjektive, qualitative Beobachtungen für spätere Untersuchungen formalisiert, diskretisiert und gelabelt werden. Weitergehend sollen im Rahmen des Feature-Engineerings auch die in Kapitel 3 genutzten, interpolierten Merkmalsvektoren komprimiert werden, um die Laufzeiten des Systems während des Erlernens neuer Ereignisse zu verkürzen und den Rechenaufwand auch für mobile Geräte mit beschränkten Ressourcen zu optimieren. Diese Herausforderungen werden folgend mit den in diesem Kapitel vorgestellten Konzepten zur qualitativen Untersuchung von wiederkehrender, menschlicher Bewegungsaktivität adressiert.

4.3 Verwandte Arbeiten

In diesem Abschnitt werden zum einen Arbeiten, welche sich mit der Segmentierung wiederkehrender Ereignisse in kontinuierlichen Zeitreihen befassen, zum anderen Arbeiten zur qualitativen Untersuchung menschlicher Bewegungsereignisse, vorgestellt. Diese dienen nachfolgend als Grundlage zur Entwicklung eines neuen, generischen Konzepts zur qualitativen Untersuchung menschlicher Bewegung.

4.3.1 Segmentierung von Bewegungsaktivität

Die Segmentierung von einzelnen Ereignissen innerhalb kontinuierlicher und multidimensionaler Zeitreihen ist im Kontext des Data-Minings und des maschinellen Lernens auf Basis von Sensorinformationen ein zentraler Aspekt. Generell kann dabei zwischen statischen und adaptiven Segmentierungsverfahren unterschieden werden. Statische Verfahren sind als solche definiert, welche mit einer festen Fenstergröße arbeiten und einzelne Ereignisse innerhalb von deren Grenzen extrahieren oder unterschiedliche Aktivitäten ohne Beachtung von zugrundeliegenden Einzelereignissen trennen. Adaptive Verfahren hingegen versuchen ihre Fenstergröße der individuellen Länge eines unterliegenden Ereignisses anzupassen, um alle dem Ereignis zugehörigen Informationen innerhalb eines Segments zu halten. Informationen vorhergehender oder nachfolgender Ereignisse hingegen sollen nicht darin enthalten sein.

Koskimaki et al. stellen ein statisches Konzept zur Segmentierung von Bewegungsinformationen, welche mithilfe eines an der Hand getragenen Accelerometers aufgenommen wurden, vor [110]. Dieses basiert auf einem überlappenden Sliding-Window-Ansatz mit einer statischen Fensterlänge, welches sequentiell über kontinuierliche Zeitreihen bewegt wird. Dabei werden für jeden Fensterinhalt 84 Features in Form von Maxima, Minima und weiteren berechnet, wobei die 20 signifikantesten zur anschließenden Erkennung von vier unterschiedlichen Bewegungsaktivitäten genutzt werden. Durch die einfache sequentielle Zuordnung von Fensterinhalten zu Aktivitäten werden keine Ereignisgrenzen identifiziert. Dadurch können einzelne Bewegungsereignisse nicht individuell untersucht werden.

Ravi et al. stellen einen ähnlichen Sliding-Window-basierten Ansatz vor, welcher mit einer festen Fensterlänge und 50% Überlappung arbeitet [178]. Aus den jeweiligen Fensterinhalten extrahieren die Autoren den durchschnittlichen Signalwert, die Standardabweichung, den durchschnittlichen Energiegehalt, und die Signalkorrelation als Features für die drei Dimensionen eines triaxialen Accelerometers. Anschließend erkennen die Autoren für alle Fensterinhalte einfache Aktivitäten wie Gehen, Rennen und weitere. Auch hier können unterliegende, sich wiederholende Einzelereignisse aufgrund fehlender Ereignisgrenzen nicht individuell untersucht werden.

Ein weiteres, ähnliches Konzept zur sequentiellen Verarbeitung stammt von Mannini et al. [126]. Dieses variiert ebenso wie die vorherigen Ansätze nicht

in seiner Fensterlänge und ist fähig zur Erkennung von Aktivitätskategorien, ohne dabei einzelne Teilereignisse genauer zu betrachten.

Im Rahmen dieser Arbeit wurde in Kapitel 3 ein Peak-basierter Segmentierungsalgorithmus mit statischer Fensterlänge präsentiert [49, 51]. Im Vergleich zu den zuvor vorgestellten Ansätzen ermöglicht dieser die Untersuchung individueller Bewegungsereignisse. Allerdings adressiert er nicht die in Abschnitt 4.2 geforderte Notwendigkeit nach Adaptivität bezüglich der Festlegung von Ereignisgrenzen. Dadurch kann es besonders bei variierenden Ereignislängen (vgl. Abbildung 4.1) zu unvollständigen Ereignissegmenten oder zu Segmenten kommen, welche ihrem Ereignis nicht zugehörigen Inhalt umfassen. Weitergehend werden feste Schwellwerte verwendet, welche im Fall von hohen Signaldynamiken innerhalb individueller Bewegungsereignisse der gleichen Bewegungsaktivität zu Fehlern in der Ereignisidentifikation führen können.

Um solche Limitierungen der statischen Verarbeitung von Bewegungsaktivität in kontinuierlichen Zeitreihen zu adressieren, können adaptive Konzepte mit dynamischen Fensterlängen verwendet werden. In diesem Kontext stellen Kozina et al. einen Ansatz zur dynamischen Segmentierung von Beschleunigungssignalen vor [113]. Dazu untersuchen die Autoren sogenannte signifikante Änderungen innerhalb einer eingehenden Zeitreihe. Diese sind definiert als Änderungen, bei welchen aufeinanderfolgende Samples eines Signals in absteigender Ordnung einen fest definierten Schwellwert unterschreiten. An diesen Stellen teilen die Autoren anschließend das Signal und erhalten somit einzelne Segmente. Aufgrund der Nutzung eines Schwellwerts und der Notwendigkeit zur Unterschreitung des selbigen ist das Vorgehen allerdings nicht robust gegen Änderungen in der Sensororientierung. So können zwar Ereignisse mit unterschiedlicher Dauer voneinander getrennt werden, es wird aber nicht beachtet, dass ein wiederkehrendes Bewegungsereignis einen festgelegten Schwellwert innerhalb einer Periodendauer gegebenenfalls mehrfach unterschreitet und somit fälschlicherweise in einzelne Fragmente zerteilt wird.

Keogh et al. präsentieren im Kontext adaptiver Ansätze den *SWAB*-Algorithmus (*Sliding Window Bottom-up*) [98, 99]. Hierbei soll die Echtzeitfähigkeit herkömmlicher Sliding-Window-Ansätze mit der Möglichkeit zu adaptiven Fensterlängen, welche das *Bottom-up*-Verfahren bietet, kombiniert werden. Verfahren dieser Art finden auch in der Computer-Grafik unter dem Begriff *Decimation*-Methoden und im Bereich medizinischer Applikationen häufigen Einsatz [80, 84]. Dazu werden Informationen kontinuierlicher Zeitreihen sequentiell mit einer statischen Fensterlänge in einen Buffer geladen. Im Buffer findet nun mithilfe des *Bottom-up*-Verfahrens die eigentliche Segmentierung statt. Dazu wird über eine Kostenfunktion der Aufwand für das Zusammenführen jedes Paares von benachbarten Segmenten in Form von einzelnen Fensterinhalten berechnet (vgl. Kapitel 2.2.6). Anschließend wird jeweils das Paar mit den niedrigsten Kosten iterativ zusammengeführt und bildet ein neues Segment. Dieser Vorgang wird so lange wiederholt, bis ein Abbruchkriterium erfüllt ist. Nach dessen Erfüllung wird das zusammengeführte Segment aus dem Buffer

entfernt und neue Fensterinhalte werden hinzugefügt. Somit können je nach Granularität und Wahl des Abbruchkriteriums auch Einzelereignisse einer Aktivitätsklasse extrahiert werden. Allerdings ist *SWAB* durch die Nutzung von statischen Fensterlängen zur Befüllung des Buffers anfällig für Rauschen und Fragmentierung an den Segmentgrenzen, da es unwahrscheinlicher ist, dass der Beginn und das Ende eines realen Ereignisses genau von den Begrenzungen des Sliding-Windows geschnitten werden.

Zusammenfassend lässt sich feststellen, dass in verwandten Arbeiten verschiedene Ansätze zur Segmentierung kontinuierlicher Zeitreihen mit jeweils unterschiedlichen Anwendungsgebieten präsentiert werden. Diese segmentieren zumeist übergeordnete Klassen von Bewegungsaktivität ohne Betrachtung von Einzelereignissen. Unabhängig von der Granularität der Segmentierung sind alle Ansätze limitiert durch die Nutzung statischer Fensterlängen, welche für jedes Segment zu fehlerhaftem Inhalt in Form von Fragmenten vorheriger oder nachfolgender Ereignisse führen kann.

4.3.2 Qualitative Untersuchung und Bewertung von Bewegung

Die qualitative Untersuchung und Bewertung von Bewegung hat in den letzten Jahren als logischer Folgeschritt der ihr zugrundeliegenden, reinen Erkennung menschlicher Aktivität in verwandten Arbeiten zunehmende Aufmerksamkeit erfahren.

Velloso et al. präsentieren in diesem Kontext eine Arbeit zur qualitativen Untersuchung von fünf unterschiedlichen Gewichthebeübungen [201]. Dazu erfassen die Autoren Gewichthebeübungen in ihrer korrekten Ausführung, anschließend zeichnen sie gezielt Fehl Ausführungen auf. Dazu nutzen sie vier Bewegungssensoren, welche an der Hüfte, am Handgelenk, am Arm sowie am zu hebenden Gewicht selbst angebracht sind und welche über ein Accelerometer, ein Gyroskop und ein Magnetometer verfügen. Mithilfe eines Sliding-Windows von statischer Länge werden die erfassten Informationen sequentiell verarbeitet und insgesamt 17 Features extrahiert. Anschließend wird jeweils der individuelle Inhalt der Sliding-Windows analysiert, ohne jedoch Rücksicht auf die tatsächlichen Grenzen einer individuellen Wiederholung eines Bewegungsereignisses zu nehmen. Zur eigentlichen Untersuchung kommt ein *Random Forest* (RF) Klassifikator zum Einsatz. Im Rahmen der folgenden *Leave-One-Out*-Validierung können Fehler in der Ausführung mit einer Erfolgsrate von 78,2% von der korrekten Ausführung unterschieden werden. Aufgrund der geringen Größe des Datensets und die durch gezielte Aufnahme von Fehlern hervorgerufenen Laborbedingungen kann auf Basis dieser Ergebnisse allerdings keine allgemeingültige Aussage getroffen werden. Weiterhin ist das Konzept durch dieses Vorgehen auch nur bedingt skalierbar.

Parisi et al. stellen ebenfalls eine Arbeit zur Bewertung von Gewichthebeübungen vor [148]. Diese werden mithilfe des Tiefensensors einer *Microsoft*

Kinect erfasst und anschließend auf Basis von Bildinformationen, welche das rekonstruierte Skelett des Athleten nachbildet, analysiert. Dazu nutzen sie ein zweischichtiges Konzept aus neuronalen *Growing when Required* (GWR) Netzwerken, welche aus den gegebenen Eingangsinformation selbstorganisiert lernen können und dabei ihre topologischen Eigenschaften erhalten. Hierzu wird in jeder Lerniteration die Aktivierung als die Differenz zwischen der Eingangsinformation und dem Wert des ihr am ähnlichsten prototypischen Neurons berechnet. Im Rahmen der Studie zur Validierung des Vorgehens trainierten die Autoren ihre Architektur für jedes Ereignis in 100 Epochen und untersuchen anschließend drei unterschiedliche Übungsklassen auf Fehlstellungen, wie etwa einen gekrümmten Rücken oder zu niedrig gehaltene Ellbogen. Dabei gelingt es ihnen, im Mittel für drei Fehlerklassen Fehlstellungen mit einer Genauigkeit von 66,6% Prozent wiederzuerkennen.

Möller et al. stellen mit *Gymskill* ein System zur qualitativen Analyse von Bewegung in Kombination mit einem Balance-Board vor [137]. Dabei untersuchen sie unterschiedliche Balance-Übungen anhand von via Smartphone erfasster Bewegungsinformation und einer erweiterten Hauptkomponentenanalyse (PCA). Zusätzlich zur qualitativen Bewertung einzelner Übungen wurden diese auch automatisiert gezählt und dem Nutzer als Feedback zur Verfügung gestellt. Im Rahmen der abschließenden Bewertung des Vorgehens durch die Studienteilnehmer wird das System als sehr hilfreich bei der Steigerung des Trainingserfolgs empfunden. Eine Einschränkung ist allerdings die Fokussierung des Systems auf seinen sehr spezifischen Aufbau, welcher nur den vorliegenden Anwendungszweck erlaubt. Bewegungsereignisse generischer Art können damit nicht untersucht werden.

Strohmann et al. entwickelten ein Konzept zur qualitativen Analyse von Laufbewegung mithilfe von Bewegungssensorik, welche am Bein eines Probanden getragen wird [197]. Ziel ist die Untersuchung des Auftritts- und Abstoßverhaltens und der Ermüdung von Läufern während des Laufens sowie die Bewertung von deren Fitnesslevel, um so automatisierte Unterstützung während des Lauftrainings zu bieten. Die im Rahmen ihrer Studie erfassten Informationen verarbeiten die Autoren manuell mithilfe verschiedener Schwellwertverfahren und visualisieren verschiedene Muster des Laufverhaltens über Clustering-Algorithmen. Auch hier ist das Vorgehen stark anwendungsbezogen und nicht generisch auf andere Bewegungsaktivitäten übertragbar. Ein Echtzeit-Feedback an den Nutzer ist aufgrund des stationären Verarbeitungsverfahrens nicht möglich.

Zusammenfassend ist zu vermerken, dass für verschiedene Anwendungsfälle bereits unterschiedliche Ansätze zur qualitativen Analyse und Bewertung von menschlicher Bewegung existieren. Diese sind bislang jedoch durchgehend an spezifische, stationäre Systemkonfigurationen gebunden oder nicht zur ganzheitlichen Analyse aller menschlichen Gliedmaßen fähig. Dies macht eine generische Anwendung, welche nicht auf spezifisch definierte Klassen von Bewegungsaktivität ausgelegt sind, unmöglich. Weitergehend können auch meist

einzelne Bewegungsereignisse innerhalb einer Kette von Bewegungsaktivitäten nicht isoliert betrachtet werden. Diese Einschränkungen werden in den folgenden Abschnitten durch die Entwicklung neuer Konzepte zur qualitativen Analyse wiederkehrender, menschlicher Bewegungsaktivität adressiert und gelöst.

4.4 Konzept zur qualitativen Untersuchung und Bewertung

Folgend wird ein Konzept zur generischen, qualitativen Analyse von wiederkehrender Bewegungsaktivität entwickelt. Die grundlegenden Herausforderungen zu diesem Vorhaben wurden dabei schon in Abschnitt 4.2 motiviert.

Zum einen soll eine Möglichkeit zur exakten Extraktion von Bewegungsereignissen entwickelt werden, um so zu verhindern, dass Bestandteile vorhergehender oder nachfolgender Ereignisse fälschlicherweise in einem Ereignissegment enthalten sind. Grundannahme ist hierbei, dass die Bildung eines Vorhersagemodells zur Erkennung von qualitativen Eigenschaften in ihrer Komplexität weit über die reine quantitative Erkennung einzelner Aktivitäten aus Kapitel 3 hinausgeht und von solchen Störinformationen negativ beeinflusst wird.

Zum anderen muss eine Metrik definiert werden, nach welcher qualitative Eigenschaften eines Bewegungsereignisses in einer diskreten Repräsentation formalisiert werden können, um diese in Form eines kompakten Merkmalsvektors für Prozesse des maschinellen Lernens lesbar zu machen. Auch hier dient die im Rahmen der Vorkapitel genutzte und in Abbildung 3.4 gezeigte Prozesskette des maschinellen Lernens als Untersuchungsgrundlage.

4.4.1 Adaptive Segmentierung von Bewegungsereignissen

Ziel dieses Kapitels ist zunächst die Entwicklung eines adaptiven, Online-fähigen Segmentierungsalgorithmus für periodische, multidimensionale Zeitreihen, wie sie von SensX als Eingangsinformationen bereitgestellt werden. Diese sind schematisch in Abbildung 4.2 dargestellt. Das Schaubild verdeutlicht, dass die $n = 30$ Eingangssignale S_i jeweils eine sequentielle, geordnete Menge $M = \{E_1, \dots, E_m\}$ zusammenhängender Ereignisse E_x liefern, wobei jedes Einzelereignis aus einem Set an Unterbewegungen der gleichen Bewegungsaktivität besteht. Alle Einzelereignisse weisen in ihrem Signalverlauf zueinander ein gewisses Maß an Ähnlichkeit auf, alle können eine variable Laufzeit Δt besitzen. Jedes dieser Einzelereignisse soll nun ohne Überlappung oder Lücken in individuelle Ereignissegmente extrahiert werden.

Dazu wird analog zum Vorgehen in Kapitel 3.5.2.1 zunächst das Most-Meaningful-Signal (MMS) S_{MMS} aus der Menge aller Eingangssignale S_i identifiziert. Grundannahme hierbei ist, dass dieses, auf Grund der im Vergleich zu anderen S_i größten beinhalteten Dynamik und Varianz im Signalverlauf

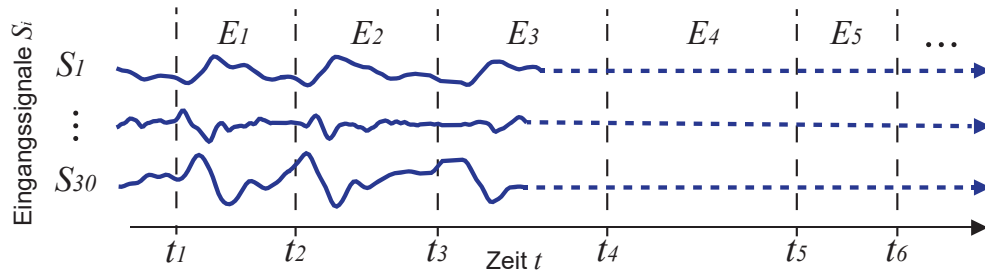


Abbildung 4.2: Schematische Darstellung der erfassten Signale des SensX-Systems und der darin beinhalteten sequentiellen Ereignisse mit ihrer individuellen und variablen Laufzeit.

besonders stark ausgeprägte Charakteristika einzelner Bewegungsereignisse innerhalb seiner Amplitude nachzeichnet.

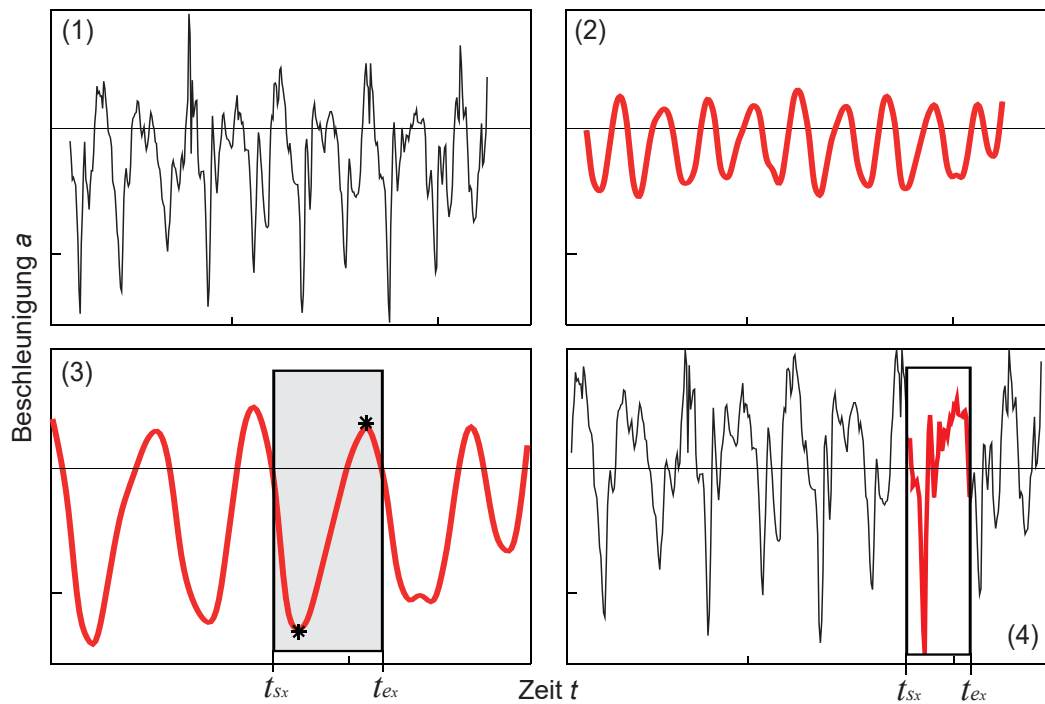


Abbildung 4.3: Einzelschritte des Vorgehens zur adaptiven Segmentierung multidimensionaler, wiederkehrender Zeitreihen anhand von Extrema-Fingerprints.

Der nun folgenden Arbeitsschritt zur Identifikation einzelner Bewegungssegmente innerhalb einer kontinuierlichen Ereignissequenz macht sich eine Beobachtung zunutze, welche in Abbildung 4.3, (1) und 4.3, (2) deutlich wird. Abbildung 4.3 gibt dabei einen Überblick über die gesamte Prozesskette zur adaptiven Segmentierung. In Schritt (1) (vgl. Abbildung 4.3, (1)) ist das nur schwach mit einem *Butterworth*-Filter geglättete S_{MMS} einer Ereignissequenz

zu sehen. In jeder individuellen Periodendauer weist das Signal lokale Minima und Maxima auf, welche in unterschiedlich starker Ausprägung für die unterschiedlichen Umläufe erkennbar sind. Deren Existenz gibt charakteristische und wertvolle Informationen auf individuelle Eigenheiten eines Bewegungsereignisses wieder, ist aber aus dem Blickwinkel der Segmentierung aufgrund oft wechselnder Dynamiken und somit einer fehlenden Generalisierbarkeit zunächst hinderlich.

In Schritt (2) (vgl. Abbildung 4.3, (2)) ist nun die gleiche Sequenz von Bewegungsereignissen zu sehen, welche einer erneuten, sehr starken Tiefpassfilterung unterzogen wurde. Das Signal hat nun nahezu die Form einer Sinusfunktion angenommen und Wiederholungen einzelner Bewegungsereignisse sind deutlich erkennbar. Es wird also deutlich, dass einzelne Bewegungsereignisse innerhalb einer kontinuierlichen Sequenz bei starker Glättung anhand eines individuellen Extrema-Fingerprints, im vorliegenden Beispiel durch einen lokalen Hochpunkt und einen lokalen Tiefpunkt, identifizierbar sind. Der folgende Abschnitt 4.4.2 betrachtet Extrema-Fingerprints und deren unterschiedliche Ausprägungen anschließend noch im Detail.

Die Beobachtung von Extrema-Fingerprints wird in Schritt (3) der Abbildung 4.3 nun dazu genutzt, um sequentiell den Startzeitpunkt t_s und den Endzeitpunkt t_e einzelner Bewegungsereignisse zu ermitteln. Dazu wird auch hier ein Fenster sequentiell über das stark geglättete S_{MMS} bewegt, wobei dieses im Vergleich zu zuvor vorgestellten verwandten Arbeiten aber keine statische Fenstergröße benutzt und in seiner Länge variieren kann. Die initiale Fenstergröße Δt wird dabei anhand der Autokorrelation des Signals ermittelt. Abbildung 4.4 verdeutlicht dabei das Vorgehen anhand des autokorrelierten Signals des SensX-Brustsensors in Y -Richtung während einer wiederkehrenden Bewegungsaktivität. Es ist zu erkennen, dass jedes sich wiederholende Einze-

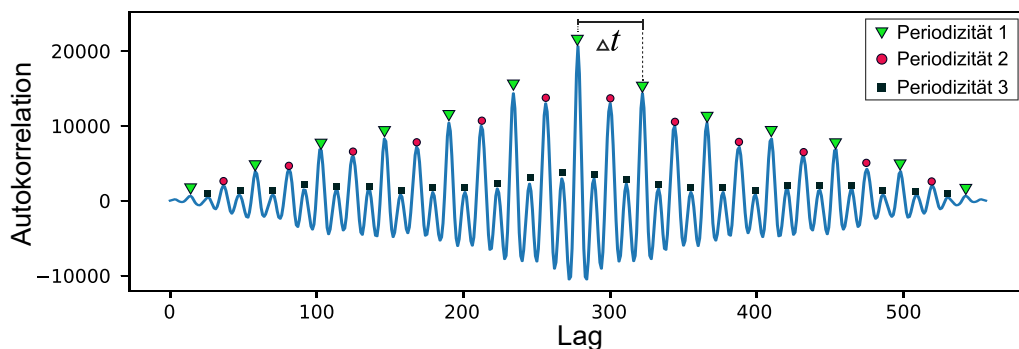


Abbildung 4.4: Autokorreliertes Signal des SensX-Brustsensors in Y -Richtung, erfasst während der Durchführung einer Abfolge der Körpereigengewichtsübung *Hampelmann*.

lereignis innerhalb der Ereignissequenz aus drei unterschiedlichen Periodizitäten zusammengesetzt ist, welche sich wiederum ebenfalls mit einer regelmäßigen Periodendauer wiederholen. Die Dauer Δt eines Einzelereignisses wird

im Schaubild durch die Anzahl der von ihr beinhalteten Samples definiert und dient als Initialisierungsparameter für die Fenstergröße des adaptiven Segmentierungsalgorithmus. Sie ergibt sich aus dem Wertebereich zwischen dem globalen Maximum als Startpunkt und dem Ende der am stärksten ausgeprägten Periodizität. Das Ereignisende ist durch den Beginn des darauffolgenden Periodenumlaufs der *Periodizität 1* gekennzeichnet (vgl. Abbildung 4.4). Anschließend wird zu Δt noch ein Sicherheitsabstand in Höhe der halben Fensterlänge hinzugefügt, um zu erkennende Extrema-Fingerprints in jedem Fall damit zu umschließen. Nach Ermittlung der initialen Fensterlänge wird das Signal nun so vertikal verschoben, dass seine Nulldurchgänge auf dem durchschnittlichen Amplitudenniveau der vor Beginn einer Bewegungsaktivität erfassten Ruhephase liegen. Wenn anschließend während der sequentiellen Verarbeitung innerhalb des Fensters F mit der initialen Länge Δt , wie in Abbildung 4.3, (3) dargestellt, einer der bekannten Extrema-Fingerprints erkannt wird, stoppt die sequentielle Verarbeitung und eine adaptive Suche nach den Segmentgrenzen wird angestoßen.

Dazu wird zunächst überprüft, ob neben dem gesuchten Extrema-Fingerprint noch andere relevante lokale Maxima oder Minima im Signal vorhanden sind. Ist dies der Fall, so werden die Segmentgrenzen zunächst so verschoben, dass diese außerhalb des Segments liegen. Grundannahme dieses adaptiven Prozesses zur iterativen Justierung der Grenzen ist, dass ein individuelles Bewegungsereignis durch die Nulldurchgänge vor und nach dem Auftreten des Extrema-Fingerprints begrenzt wird. Ist das letzte Sample innerhalb des Segments von einem positiven Wert bestimmt, so werden solange darauffolgende Samples zum Segment hinzugefügt bis der nächste Nulldurchgang erreicht ist. Ist der letzte Wert hingegen negativer Natur, so werden bis zum Erreichen des Nulldurchgangs Samples aus dem Segment entfernt. Dieser Nulldurchgang definiert den Endpunkt des Segments und der dazugehörige Zeitstempel t_{e_x} wird zur späteren Ereignisextraktion vermerkt. Analog zu diesem Verfahren wird auch der Startzeitpunkt t_{s_x} des Segments identifiziert. Ausgehend vom Wert des ersten Samples werden vorhergehende Samples zum Segment hinzugefügt, wenn dieser positiv ist und es werden Samples entfernt, wenn dieser negativ ist. Somit sind die Segmentgrenzen t_{s_x} und t_{e_x} eines individuellen Bewegungsereignisses innerhalb einer Bewegungsaktivität adaptiv identifizierbar.

Diese Segmentgrenzen werden anschließend, wie in Abbildung 4.3, (4) dargestellt, verwendet um aus allen $n = 30$ Eingangssignalen S_i ein Bewegungsereignis innerhalb der Grenzen t_{s_x} und t_{e_x} in ein Segment von individueller Länge zu extrahieren. Da es sich bei S_i jetzt um die anfangs nur leicht geglätteten Signale handelt, sind somit auch alle zur späteren Untersuchung relevanten Informationen wieder darin enthalten. Ausgangsinformation aus dem Segmentierungsprozess in Schritt (4), Abbildung 4.3, ist also ein Bewegungsereignis E_x . Dieses wird von einem Set aus 30 Einzelsignalen, welche alle die dem unterliegenden Bewegungsereignis eigene, individuelle Laufzeit $\Delta t_{E_x} = t_{e_x} - t_{s_x}$ besitzen, beschrieben. Nach Extraktion eines Ereignisses E_x wird erneut der

sequentielle Suchvorgang mit t_{e_x} als Startpunkt angestoßen.

In diesem Kontext zeigt Abbildung 4.5 die Anzahl extrahierter Segmente aus einer wiederkehrenden Bewegungsaktivität in Abhängigkeit der Wahl des Cutoff-Faktors des zur Glättung genutzten Tiefpass-Filters (vgl. Kapitel 3.5.2.1). Wie diese zeigt, ist die Stärke einer Glättung in Schritt (2) der Ab-

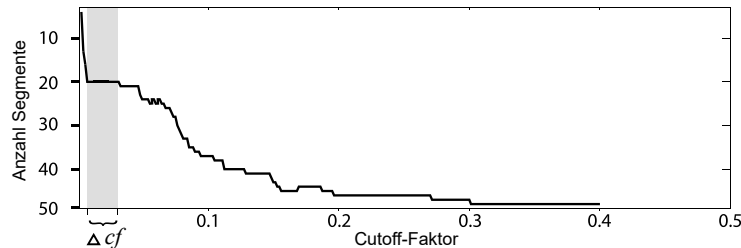


Abbildung 4.5: Bestimmung des optimalen Bereichs des Cutoff-Faktors Δcf innerhalb einer Sequenz mit 20 bekannten Bewegungsereignissen. Der optimale Bereich Δcf ist grau hervorgehoben.

bildung 4.3 essentiell, um für das adaptive Segmentierungsverfahren benötigte Extrema-Fingerprints optimal sichtbar zu machen. Dazu wird für jedes Sensor-Setup vor dem Online-Einsatz eine einmalige Initialisierungsphase benötigt, in welcher der Bereich des optimalen Cutoff-Faktors Δcf bestimmt wird. Hierbei wird analog zu Abbildung 4.5 eine Sequenz mit bekannter Anzahl von einzelnen Bewegungsereignissen iterativ mit verschiedenen cf schrittweise verarbeitet. Der optimale Bereich Δcf ergibt sich, wenn die bekannte Anzahl an Ereignissen mithilfe des Verfahrens gefunden wird.

4.4.2 Extrema-Fingerprints zur Ereignisdetektion

Wie in Abbildung 4.6 verdeutlicht, können Ereignisse unterschiedlicher Bewegungsaktivitäten, je nach Anzahl von Unterbewegungen, Dauer und Konfiguration der angewendeten Tiefpassfilterung Fingerprints mit unterschiedlichen Zusammensetzungen lokaler Extrema aufweisen. In 4.6, (1) und 4.6, (4) ist jeweils ein ähnlicher Extrema-Fingerprint, bestehend aus zwei Maxima und zwei Minima zu sehen. 4.6, (2) entspricht dem im Vergleich einfachsten Fingerprint, welcher auf einer sinusförmigen Schwingung beruht und aus einem Maximum sowie einem Minimum besteht. 4.6, (r) bildet mit jeweils vier Maxima und vier Minima den komplexesten Fingerprint.

Grundsätzlich können auch die S_{MMS} in 4.6, (1), 4.6, (3) und 4.6, (4) durch eine stärkere Filterkonfiguration auf einen sinusförmigen Signalverlauf hin geglättet werden. Dabei ist allerdings zu beachten, dass je stärker das zugrundeliegende S_{MMS} mithilfe des Tiefpassfilters geglättet wird, desto weiter verschieben sich durch die Glättung auch die Segmentgrenzen im Vergleich zu ihrer Position im ursprünglichen, ungefilterten Signal. Diese Abweichung ist in Abbildung 4.6 bei dem Vergleich aller Segmentgrenzen der gefilterten und ungefilterten Signale zu sehen. Solange die Abweichung dabei nicht zu groß ist

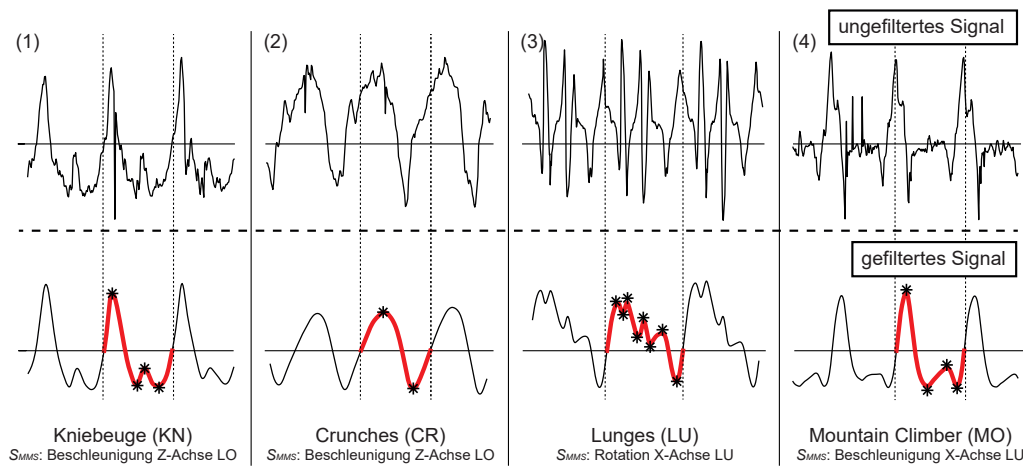


Abbildung 4.6: Beispiele von Extrema-Fingerprints verschiedener Bewegungsereignisse mit unterschiedlich starker Konfiguration eines Butterworth-Tiefpassfilters. Oben ist jeweils das schwach gefilterte Eingangssignal zu sehen, unten das stark gefilterte Signal zur Identifikation der Extrema-Fingerprints.

und in jedem Ereignissegment einer Aktivitätsklasse gleichermaßen enthalten ist, ergibt sich für die spätere Klassifizierung kein Problem. Es lässt aber eine Wechselwirkung zwischen der Notwendigkeit zur Vereinfachung des Signals durch Schaffung von Extrema-Fingerprints mit möglichst geringer Komplexität und der mit der Vereinfachung eingehenden Grenzverschiebung erkennen.

4.4.3 Qualitative Bewertung und Formalisierung von Bewegung

Eine weitere Herausforderung bei der qualitativen Untersuchung von wiederkehrender Bewegungsaktivität ist die Formalisierung von qualitativen Beobachtungen und daraus hervorgehend das Bilden von diskreten Merkmalsvektoren mit entsprechendem Labeling. Als Grundlage für einen solchen Prozess sollen im folgenden die Bewegungsinformationen dienen, welche zur quantitativen Evaluation von SensX bereits im vorhergehenden Kapitel genutzt wurden (vgl. Abschnitt 3.6.1). Diese wurden im Abschnitt 4.4.1 bereits in einzelne Bewegungsereignisse E_x von jeweils individueller Länge segmentiert. Nun soll jedem Ereignis ein Label zur qualitativen Beschreibung von dessen zugrundeliegenden Bewegungsinformation zugeordnet werden. Da während der Durchführung der Studie zur Datenerfassung alle Probanden auf Video aufgenommen wurden, können diese Videos nun als Basis zur subjektiven qualitativen Bewertung einzelner Bewegungsereignisse dienen. Dazu wurde folgendes Vorgehen angewandt.

Jedes einzelne Bewegungsereignis wurde anhand der Videoaufzeichnungen von zwei Experten analysiert. Einer der Experten ist professioneller Sport-

Ingenieur mit fundierten Kenntnissen in den Bereichen Ausdauer-Training und Körpereigengewichtsübungen. Die zweite Expertenanalyse wurde von einem ambitionierten Sportler in den Bereichen Körpereigengewichtsübungen und CrossFit durchgeführt. Zu Beginn des Bewertungsprozesses wurden alle Bewegungsereignisse mit dem initialen Startwert $p_s = 1$ bewertet, was zugleich der besten Qualitätsklasse 1 und damit einer fehlerlosen Ausführung des Bewegungsereignisses entspricht. Im Anschluss wird für jede Anomalie innerhalb der Ausführung, zum Beispiel Fehlstellungen von einzelnen Körperextremitäten oder Fehler in der Ausführung, ein Strafwert p_a zum Initialwert p_s hinzugefügt. Als Anomalien gelten hierbei sämtliche Abweichungen in der Bewegungsausführung im Vergleich zu der Ausführung der Bewegungen im vor der Erfassung bereitgestellten Instruktionsvideo. Der maximale Strafwert wird dabei als $p_a = 1$ festgelegt, was einer signifikanten Anomalie entspricht. Bei erkennbarer, aber weniger starker Ausprägung einer Anomalie wird $p_a = 0,5$ gesetzt, im Falle kaum merklicher Auffälligkeiten gilt der Strafwert $p_a = 0,25$. Das finale Qualitätslabel L_x eines individuellen Ereignisses E_x ergibt sich anschließend aus der Summe aller i Strafwerte p_a analog zu Berechnung 4.1.

$$L_x = p_s + \sum_{n=1}^i p_{a_n}, \text{ mit } L_x > 5 : L_x = 5 \quad (4.1)$$

Für den Fall, dass der so ermittelte Wert, welcher der Gesamtbewertung entspricht, keine ganze Zahl ist, wurde dieser auf eine solche gerundet. Daraus ergibt sich eine Skala von 5 möglichen Bewertungsklassen im Bereich von 1-5. Die schlechteste mögliche Einstufung bildet hierbei Qualitätsklasse 5. Wird in der Summe aller p_a für ein Ereignis ein Wert größer 5 erreicht, so gilt $L_x = 5$, was der schlechtesten möglichen Bewertung entspricht. Die Anzahl aller beobachteten Anomalien je Aktivitätsklasse bewegte sich in einem Bereich von 7-13 Vorfällen je Klasse.

Dieses Labeling gibt nun Aufschluss über die Gesamtqualität der zugrundeliegenden Bewegungsinformationen eines Segments hinsichtlich der Korrektheit der Bewegungsausführung im Vergleich zur Vorgabe der Videoinstruktionen. Eine explizite Adressierung konkreter Anomalien ist darin nicht vorgesehen. Die Validität dieses Vorgehens zur Formalisierung qualitativer Bewertungen von Bewegung wird im Anschluss im Rahmen der Evaluation in Abschnitt 4.5 überprüft.

4.4.4 Aufbau komprimierter Merkmalsvektoren

Um die zuvor segmentierten Bewegungsereignisse E_x und die ihnen jeweils zugeordneten Qualitätsklassen in Form ihrer Label L_x für einen Prozess des überwachten maschinellen Lernens vorzubereiten, ist es notwendig, diese in diskreten Merkmalsvektoren zu kodieren. Jede Zuordnung aus Label und Ereignis soll dabei eine Instanz I_x beschreiben. Dazu muss das bislang aus $n = 30$ Si-

gnalfragmenten bestehende Ereignis E_x in einen eindimensionalen Vektor X_{E_x} überführt werden, um eine Zuordnung $I_x = \{X_{E_x}|L_x\}$ abzubilden. Im Rahmen dieser Überführung soll X_{E_x} aber nicht nur Informationen über den Verlauf aller 30 Einzelsignale enthalten, sondern zusätzlich, im Gegensatz zu auf Basis von Interpolation und Resampling gebildeten Merkmalsvektoren (vgl. 3.5.2.2), auch ein hohes Maß an Kompression realisieren. So sollen geringere Laufzeiten für die spätere Online-Untersuchung und das Echtzeittraining von Vorhersagemodellen mit neuen Bewegungsereignissen ermöglicht werden.

Eine Hilfestellung zur Umsetzung dieser Anforderungen gibt eine genauere Betrachtung der Abbildungen 4.7 und 4.8. Diese stellen die Verteilung der Stan-

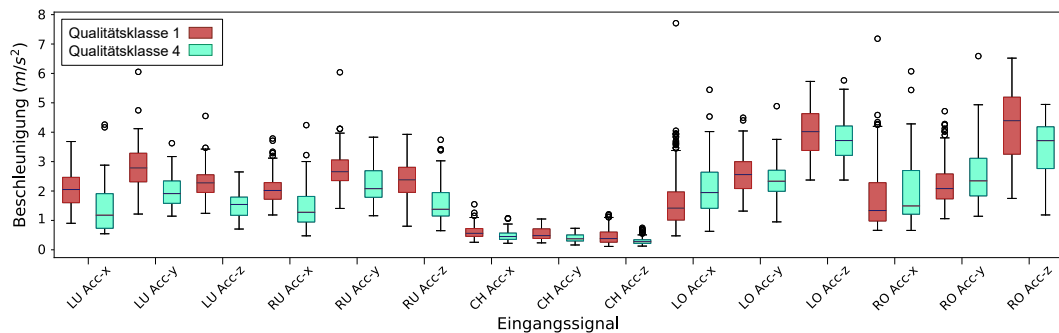


Abbildung 4.7: Standardabweichung der Beschleunigungsinformation von jeweils 100 Bewegungsereignissen der Körpereigengewichtsübung *Bicycle Crunch* der Qualitätsklassen 1 und 4, aufgeteilt nach zur Bewegungserfassung genutztem Sensor.

dardabweichungen von Beschleunigung (vgl. Abbildung 4.7) und Rotation (vgl. Abbildung 4.8) für jeweils 100 Ereignisse der gelabelten Qualitätsklassen 1 und 4, jeweils angeordnet nach erfassendem Sensor, gegenüber. Die übergeordnete Aktivitätsklasse ist dabei die Körpereigengewichtsübung *Bicycle Crunches*. Bei genauerer Betrachtung ist zu erkennen, dass die einzelnen Qualitätsklassen bereits anhand der Ausprägung ihrer Standardabweichung voneinander unterschieden werden können. Besonders signifikant sind hier stark ausgeprägte Rotationswerte für den linken und den rechten Arm in X -Richtung (LO Rot-x und RO Rot-x), welche in Qualitätsklasse 1 eine deutliche Drehung des Ellenbogens eines Athleten zu seinem Knie beschreiben, während diese in Qualitätsklasse 4 wenig bis gar nicht ausgeprägt ist. Analog dazu verhält es sich mit den Beschleunigungswerten beider Arme in Z -Richtung (LO Acc-z und RO Acc-z). Diese Signalausprägung ist auch in den damit verbundenen Rotationsinformationen der Brust in Y -Richtung nachzuvollziehen (CH Rot-y).

Im Bereich der Beine ist deutlich zu sehen, dass deren Beschleunigung in der Qualitätsklasse 1 größer ist als in Qualitätsklasse 4. Die Rotation der Beine ist in Qualitätsklasse 4 hingegen vor allem für Bewegungen in Z -Richtung stark ausgeprägt. Zusätzlich tritt dabei eine deutlich stärkere Streuung der Werte im

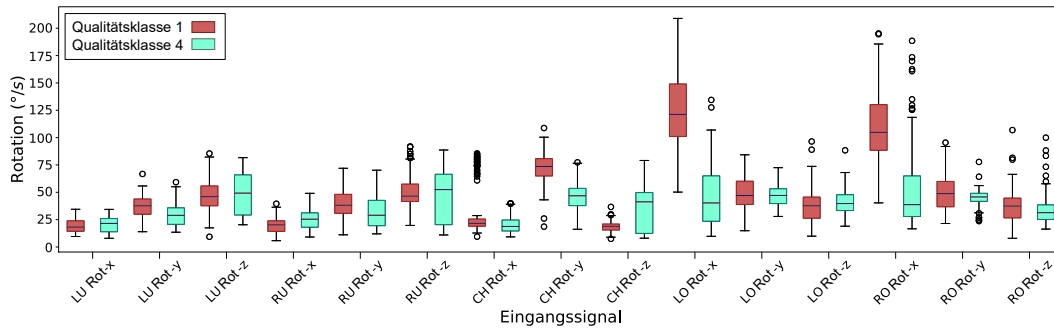


Abbildung 4.8: Standardabweichung der Rotationsinformationen von jeweils 100 Bewegungsereignissen der Körpereigengewichtsübung *Bicycle Crunch* der Qualitätsklassen 1 und 4, aufgeteilt nach zur Bewegungserfassung genutztem Sensor.

Gegensatz zu Qualitätsklasse 1 auf. Dies zeigt, dass die dem *Bicycle Crunch* eigene Beinbewegung für Ereignisse der Qualitätsklasse 1 von den betreffenden Athleten deutlich zielgerichteter und kontrollierter ausgeführt wurde. Unnötige Bewegungen wurden in Ereignissen der Qualitätsklasse 1 größtenteils vermieden oder sind deutlich weniger stark ausgeprägt.

Insgesamt sind relevante Bewegungen in der Qualitätsklasse 1 stärker ausgeprägt. Irrelevante Bewegungen hingegen, welche von Gliedmaßen stammen die zur Durchführung der Bewegungsaktivität nicht benötigt werden, sind deutlich geringer ausgeprägt und weisen eine geringe Streuung in ihrer Werteverteilung auf. Relevante Bewegungen der Qualitätsklasse 4 hingegen sind oft durch deutlich geringere Intensität ausgezeichnet, während irrelevante Bewegungen oft größere Werte, eine höhere Streuung und mehr Ausreißer aufweisen. Dies deutet auf eine ungenaue und weniger kontrollierte Bewegungsausführung von Ereignissen in Qualitätsklasse 4 hin.

Auf dieser Beobachtung aufbauend werden nun diskrete, eindimensionale Merkmalsvektoren X_{E_x} gebildet, welche jeweils ihrem Label L_x zugeordnet sind und die Bewegungsinformationen individueller Einzelereignisse beschreiben. Abbildung 4.9 gibt einen Überblick über das Vorgehen. Für jedes erfass-

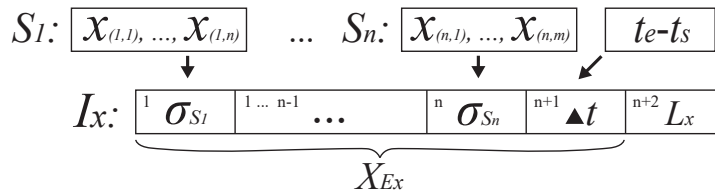


Abbildung 4.9: Schematische Darstellung der Konstruktion der einzelnen Ereignisinstanzen I_x auf Basis des Merkmalsvektors X_{E_x} und dem zugehörigen Qualitätslabel L_x .

te Ereignis $E_x = \{S_1, \dots, S_n\}$, welches aus den Einzelsignalen S_i der Anzahl

$n = 30$ besteht, wird nun eine Instanz $I_x = \{X_{E_x}|L_x\}$ definiert. Wie in Abbildung 4.9 zu sehen ist, definiert sich der Feature-Vektor X_{E_x} dabei für jedes Einzelereignis aus den individuellen Standardabweichungen σ für jedes der 30 eingehenden S_i . Als zusätzlicher Parameter wird noch die Laufzeit $\Delta t = t_e - t_s$ für jedes individuelle Ereignis in X_{E_x} integriert. Damit entsteht für jedes Ereignis ein Merkmalsvektor mit 31 Stellen, welche alle Informationen über das Ereignis in diskreter Form enthalten. Zusammen mit dem korrespondierenden Label L_x , welches an Stelle 32 eingefügt wird, kann nun die dem Ereignis E_x zugehörige Instanz I_x gebildet werden. Die Menge aller Instanzen kann jetzt mit automatisierten Verfahren zum überwachten maschinellen Lernen verarbeitet werden und somit als Grundlage zum Training eines Vorhersagemodells dienen.

Durch die stark komprimierte Vektor-Länge mit nur 31 Stellen zur eigentlichen Beschreibung von Bewegung sind merkliche Verkürzungen der Laufzeit während der Modellbildung und während der Online-Klassifikation einzelner Ereignisse zu erwarten. Vor allem im Vergleich zur interpolationsbasierten Vektorkonstruktion aus Kapitel 3, welche aus 1800 Einzelstellen zur Beschreibung von Bewegung besteht, wird ein Kompressionsfaktor von etwa 1:58 erreicht. Die Auswirkungen der Kompression auf die Laufzeit und die Vorhersageperformanz der in diesem Kontext trainierten Vorhersagemodelle wird im anschließenden Kapitel 4.5 betrachtet.

4.5 Evaluation

Im Rahmen der Evaluation dieses Kapitels sollen mehrere Aspekte untersucht werden. Zuerst werden der zuvor entwickelte, stark komprimierte Merkmalsvektor und dessen korrespondierendes qualitatives Labeling hinsichtlich ihrer Fähigkeiten zur qualitativen Untersuchung von Bewegungsereignissen evaluiert. Im Anschluss arbeiten verschiedene Experimente die Vorteile und Nachteile von adaptiven und statischen Segmentierungsverfahren heraus. Zusätzlich wird der in diesem Kapitel entwickelte, komprimierte Merkmalsvektor hinsichtlich seiner Fähigkeiten zur quantitativen Klassifizierung unterschiedlicher Bewegungsereignisse untersucht. Die Ergebnisse dieser Untersuchungen werden dabei auch mit denen aus Kapitel 3.6.3, in welchem der genutzte Merkmalsvektor mithilfe von Interpolation gebildet wurde, verglichen. Damit verbunden werden auch Ergebnisse zur Trainingslaufzeit von Vorhersagemodellen sowie zu den Auswirkungen variierender Anzahlen von Trainingsinstanzen auf die generelle Ergebnisqualität, präsentiert. Datengrundlage der dazu durchgeführten Experimente bilden die mithilfe des adaptiven Segmentierungsverfahrens (vgl. Abschnitt 4.4.1) extrahierten Bewegungsereignisse, welche im Rahmen der in Kapitel 3.6.1 durchgeführten Studie erfasst wurden. Bei diesen handelt es sich analog zu diesem Kapitel um *Crunches* (CR), *Lunges* (LU), *Hampelmann* (HA), *Bicycle Crunch* (BC), *Kniebeuge* (KN) und *Mountain Climber* (MC). Dabei wurden insgesamt 7,413 Ereignisse aus einer Gesamtmenge von

7.534 erfassten Ereignissen adaptiv segmentiert und gelabelt (vgl. Abschnitt 4.5.2 für nicht extrahierte Ereignisse während der Segmentierung).

4.5.1 Automatisierte, qualitative Bewertung von Bewegung

Schwerpunkt dieses Kapitels ist die qualitative Analyse von wiederkehrender Bewegungsaktivität, weshalb dieser Aspekt zuerst untersucht werden soll. Die Untersuchungen wurden mithilfe von Konzepten des überwachten maschinellen Lernen implementiert, wobei verschiedene Algorithmen, deren Eignung zur Erkennung menschlicher Aktivität bereits in verwandten Arbeiten demonstriert wurde, als Klassifikatoren zum Einsatz kamen. Diese sind namentlich der *C.4.5*, der *Random Forest* (RF), der *Naive Bayes* (NB) und eine Implementierung der *Support Vector Machine* (SVM) [32, 6, 118, 178, 188]. Zusätzlich wurden alle Instanzen auch im Rahmen einer Hyper-Parameter-Optimierung (HPO) mit automatisiert ausgewählten Klassifikatoren untersucht [112]. Alle in diesem Abschnitt beschriebenen Experimente wurden im Rahmen einer 10fCV untersucht. Tabelle 4.1 zeigt die Ergebnisse der nach Qualitätsklassen vor-

Klassifikator	CR (%)	LU (%)	HA (%)	BC (%)	KN (%)	MC (%)	Ø (%)
RF	88,0	90,0	92,1	92,1	93,4	82,5	89,7
C4.5	79,1	80,5	83,6	82,1	84,6	67,9	79,6
SVM	73,2	80,8	85,0	85,7	80,2	60,2	77,5
NB	54,3	70,3	72,5	76,3	58,5	54,6	64,4

Tabelle 4.1: Erfolgreiche Erkennungsrate im Rahmen der automatisierten Klassifizierung nach Qualitätsklassen für die untersuchten Körpereigengewichtsübungen.

genommenen Klassifikation mithilfe der manuell parametrisierten Klassifikatoren, welche im Rahmen von überwachten, maschinellen Lernprozessen und auf Basis der stark komprimierten Merkmalsvektoren aus Abschnitt 4.4.4 ermittelt wurden. Hierbei hat die Anwendung des RF-Klassifikators mit einer gemittelten korrekten Erkennungsrate von 89,7% zu den besten Ergebnissen geführt. Der *C4.5*-Entscheidungsbaum erreicht an zweiter Stelle 79,6%. Der SVM-Klassifikator erreicht vergleichbar gute Ergebnisse und ist für einzelne Übungen, wie den *Lunges*, dem *Hampelmann* und den *Bicycle Crunches* sogar performanter als der *C4.5*. Allerdings sind die Vorhersagen der SVM im Vergleich zu denen des *C4.5* für andere Ereignisse, wie den *Crunch*, die *Kniebeuge* und den *Mountain Climber* schlechter, woraus eine durchschnittliche Erfolgsrate von 77,5% resultiert. Der NB-Klassifikator weist mit durchschnittlich 64,4% die generell schlechteste Performanz auf. Auffällig ist auch, dass die einzelnen Qualitätsklassen des *Mountain Climbers* von allen Klassifikatoren bis auf den NB am schlechtesten erkannt wurden.

Einen feingranularen Einblick hierzu gibt Abbildung 4.10, welche die Konfusionsmatrizen der Klassifikationen nach Qualitätsklassen mit dem RF-

Algorithmus darstellt. Es ist deutlich zu erkennen, dass Fehlklassifizierungen meist durch Zuweisung zu benachbarten Qualitätsklassen zustande kamen. Gründe hierfür können das subjektive Vorgehen zum qualitativen Labeling sowie das abschließende Runden auf ganze Zahlen zur Ermittlung des finalen Qualitätslabels sein. Dadurch werden Ereignisse, welche auf Basis der diskre-

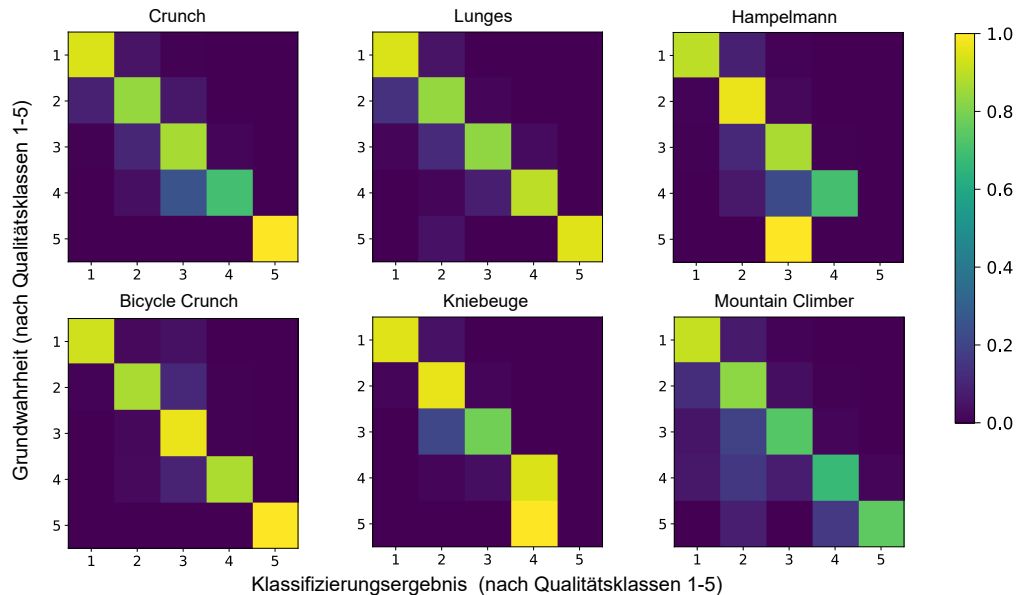


Abbildung 4.10: Zuordnungen nach den Qualitätsklassen 1-5 für die sechs untersuchten Ereignisklassen unter Zuhilfenahme eines RF-Klassifikators mit manuellem Parameter-Setting.

ten Information ihres Merkmalsvektors von einem Klassifikator nach objektiven Kriterien einer bestimmten Klasse zuzuordnen wären, fälschlicherweise benachbarten Qualitätsklassen zugeordnet. Eine weitere Auffälligkeit ergibt sich auch aus den Darstellungen der Konfusionsmatrizen des *Hampelmanns* und der *Kniebeuge*. Hier wird die schlechteste Qualitätsklasse 5 für alle Instanzen falsch zugeordnet. Hintergrund hierbei könnte sein, dass diese Ereignisse in ihrer Ausführung im Vergleich zu anderen Körperereignisübungen in puncto Koordination und Anspruch deutlich weniger komplex sind. Somit wurden im Datenset nur wenige Instanzen der schlechtesten Qualitätsklasse 5 zugeordnet, was deren Repräsentation innerhalb eines Vorhersagemodells, welches mithilfe von manuell konfigurierten Klassifikatoren erlernt wurde, deutlich erschwert. Es liegt also ein *Class-Imbalancement* vor (vgl. Kapitel 2.3.2.2). An der Darstellung des *Mountain Climbers* spiegeln sich die in Tabelle 4.1 präsentierten Ergebnisse wieder. Viele Fehlzuordnungen befinden sich dabei in den benachbarten Qualitätsklassen – sie werden aber auch über deren Grenzen hinaus platziert. Eine weitere generelle Beobachtung ist, dass im Rahmen der automatisierten, qualitativen Untersuchung fehlerhafte Zuordnungen meist zu besseren Qualitätsklassen tendieren. Selten werden Ereignisse schlechter eingeordnet als deren Label besagt.

Nach den zuvor beschriebenen Untersuchungen mit manuellem Parameter-Setting wurden auch Experimente unter Zuhilfenahme automatisierter Prozesse zur Identifikation geeigneter Klassifikatoren mit optimierten Hyperparametern durchgeführt. Dabei bekam die Testmaschine im stationären Offline-Szenario je zu untersuchender Bewegungsaktivität einen Zeitrahmen von 15 Minuten zur automatisierten Ermittlung und Konfiguration eines genau auf die jeweilige Aktivität zugeschnittenen Klassifikators. Einen Überblick

	CR (%)	LU (%)	HA (%)	BC (%)	KN (%)	MC (%)	Ø (%)
Erkennungsrate	100	100	99.9	96.0	100	100	99.3
Klassifikator	RF	RF	K*	kNN	RF	K*	-

Tabelle 4.2: Gemittelte korrekte Erkennungsraten der Qualitätsklassen für die sechs untersuchten Bewegungsereignisse unter Zuhilfenahme eines automatisch ermittelten und parametrisierten Klassifikators.

über die diesbezüglichen Ergebnisse gibt hierbei Tabelle 4.2. Generell konnte die Erkennungsrate der einzelnen Qualitätsklassen für jedes Bewegungsereignis im Rahmen des automatisierten Ansatzes erheblich verbessert werden. Die gemittelte korrekte Erkennungsrate liegt in diesem Kontext bei 99,3%. Auffällig ist, dass auch im automatisierten Szenario oft der RF-Algorithmus als optimaler Klassifikator ermittelt wurde, weitere sind zweimal der *K-Star*-Algorithmus (K*) für den *Hampelmann* und den *Mountain Climber* und der *K-Nearest Neighbours*-Algorithmus (kNN) für *Bicycle Crunches*. Hervorzuheben ist dabei, dass vor allem die Instanzen des *Mountain Climbers*, für welche im manuellen Parameter-Setting durchgehend vergleichsweise schlechte Vorhersagewerte erreicht wurden, im automatisierten Ansatz mit einer Erfolgsrate von 100% ihrer zugehörigen Qualitätsklasse zugeordnet werden.

Zusammenfassend lässt sich zur qualitativen Untersuchung sagen, dass individuelle Bewegungsereignisse verschiedener menschlicher Bewegungsaktivitäten auf Basis eines hoch komprimierten Merkmalsvektors mit einer gemittelten korrekten Erkennungsrate von 99,3% ihrer zugehörigen Qualitätsklasse zugeordnet werden können. Es ist also möglich, mit SensX erfasste Bewegungsereignisse hinsichtlich der Qualität und Korrektheit ihrer Ausführung auf einer Skala von 1-5 generisch und automatisiert zu bewerten.

4.5.2 Adaptive versus statische Segmentierungsverfahren

Eine wichtige Grundannahme dieser Arbeit bezüglich der Anwendung adaptiver Segmentierungsverfahren ist, dass adaptiv segmentierte Ereignisse, deren Segmentgrenzen genau an den tatsächlichen Anfangs- und Endzeitpunkten liegen, mit höherem Erfolg korrekt klassifiziert werden können als statisch segmentierte Ereignisse. Hauptgrund dafür ist, dass sich bei der adaptiven Segmentierung nur die Information eines einzelnen Ereignisses innerhalb des zu untersuchenden Segments befindet. Es befinden sich keine Stör- oder Rauschinformationen von vorhergehenden und darauffolgenden Ereignissen darin,

ebenso werden Ereignisse auch nicht vor ihrem eigentlichen Ende abgeschnitten, sondern sind vollständig enthalten. Abbildung 4.11 verdeutlicht hierbei

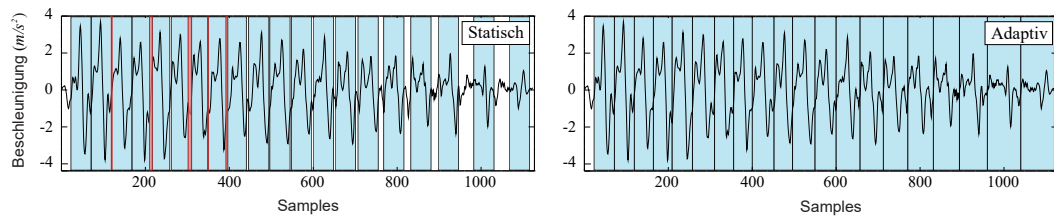


Abbildung 4.11: Vergleich der Segmentierungseigenschaften des statischen und des adaptiven Segmentierungsalgorithmus am Beispiel einer Folge von 20 Bewegungsereignissen mit variierender Dauer.

den Unterschied der beiden Segmentierungsvorgehen anhand des S_{MMS} des bereits in Abbildung 4.1 vorgestellten *Set 2* mit 20 Wiederholungen der Körpereigengewichtsübung *Crunches*. Der statische Ansatz extrahiert einzelne Ereignisse anfangs relativ genau, wenn auch mit einigen Überschneidungen (siehe rote Bereiche). Gegen Ende des Sets wirkt sich die Ermüdung des Studienteilnehmers stärker auf dessen Fähigkeit zur gleichmäßigen Übungsdurchführung aus – die individuelle Ereignisdauer verlängert sich. Der statische Ansatz kann die einzelnen Ereignisse nun nicht mehr nahtlos extrahieren und je länger ein Ereignis andauert, desto mehr Teile von dessen Informationen liegen außerhalb des eigentlichen Extraktionsfensters (siehe weiße Bereiche). Der adaptive Ansatz sorgt im Vergleich dazu für nahtlose und überschneidungsfreie Segmentierung, welche auch eine stark verlängerte Ereignisdauer gegen Ende der Sequenz berücksichtigt. Die Hypothese, dass diese erhöhte Genauigkeit durch Segmentierung die Untersuchung von Bewegungsereignissen positiv beeinflusst, soll nun untersucht werden.

Dazu wurden statisch segmentierte Bewegungsereignisse des vorhergehenden Kapitels der Klassen *Kniebeuge* und *Lunges* (vgl. Kapitel 3.5.2.1) nach dem in Abschnitt 4.4.4 definierten Verfahren in komprimierte Merkmalsvektoren überführt, wobei alle überführten Instanzen ebenfalls mit ihrer korrespondierenden Qualitätsklasse gelabelt wurden. Die untersuchten Ereignisklassen können hierbei auf Grund ihrer mittleren bis guten Platzierung innerhalb der Ergebnisse aus dem vorhergehenden Abschnitt als repräsentativ angesehen werden. Tabelle 4.3 und die Abbildungen 4.12 und 4.13 geben diesbezüglich einen Überblick über die Performanz der einzelnen Klassifikatoren bei der qualitativen Untersuchung in Abhängigkeit vom jeweils zugrundeliegenden Segmentierungsansatz. Es ist ersichtlich, dass der statische Segmentierungsansatz in allen Fällen zu einem schlechteren Klassifikationsergebnis führt. Selbst im automatisierten HPO-Szenario (vgl. Abschnitt 2.3.4) gelingen mit diesem, im Gegensatz zum adaptiven Ansatz, keine optimalen Ergebnisse. Dabei erreicht der adaptive Ansatz um 0,4 - 11,0 Prozentpunkte bessere Werte in der qualitativen Klassifizierung. Im Mittel liegt der Vorteil über alle genutzten Klassifikatoren bei

Klassifikator	KN, adaptiv (%)	KN, statisch (%)	LU, adaptiv (%)	LU, statisch (%)
RF	93,4	90,55	90,0	87,45
C4.5	84,6	77,46	80,5	73,94
NB	58,5	52,54	70,3	65,55
SVM	80,2	69,20	80,8	77,61
HPO	100,0	99,60	100,0	93,76

Tabelle 4.3: Gegenüberstellung der Ergebnisse der qualitativen Untersuchung für die Nutzung eines statischen und eines adaptiven Segmentierungsansatzes anhand der Ereignisklassen *Kniebeuge* und *Lunges*.

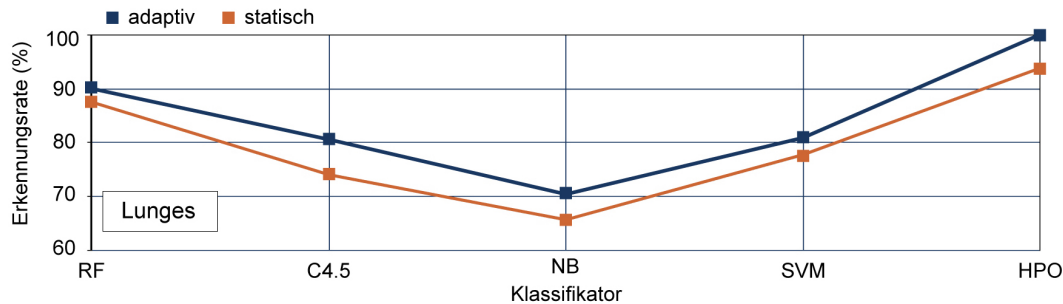


Abbildung 4.12: Ergebnisse der qualitativen Untersuchung der Ereignisklasse *Lunges* für die verschiedenen Klassifikatoren unter Benutzung eines statischen und eines adaptiven Segmentierungsansatzes.

etwa 5,06%. Damit sind die Vorteile des adaptiven Segmentierungsverfahrens gegenüber einem statischen Ansatz im Rahmen der qualitativen Analyse erwiesen. Allerdings fällt das Plus an Performanz kleiner als im Vorfeld erwartet aus.

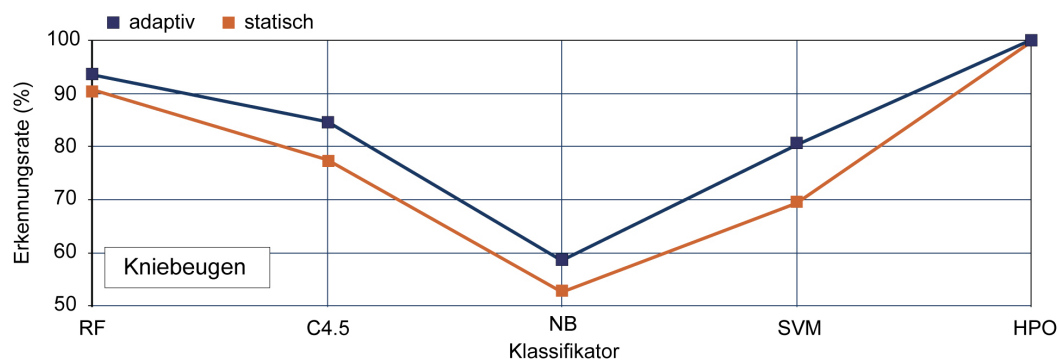


Abbildung 4.13: Ergebnisse der qualitativen Untersuchung der Ereignisklasse *Kniebeuge* für die verschiedenen Klassifikatoren unter Benutzung eines statischen und eines adaptiven Segmentierungsansatzes.

Zusätzlich ist zu erwähnen, dass die gesteigerte Ergebnisqualität in puncto Laufzeit zu merklichen Einbußen führt. So benötigt der adaptive Segmen-

tierungsansatz im Rahmen des zu dieser Arbeit genutzten Test-Setups zur Segmentierung aller 30 Signale eines Einzelereignisses im Mittel $221ms$, wohingegen der statische Ansatz mit $119ms$ in etwa doppelt so schnell ist. Es kann aber davon ausgegangen werden, dass beide Ansätze für eine echtzeitfähige Anwendung geeignet sind und die jeweiligen Laufzeiten im Kontext der Online-Klassifizierung verschmerzbar sind.

Als weiterer Nachteil des adaptiven Algorithmus ist noch die minimal geringere Leistung hinsichtlich der Extraktionsrate einzelner Bewegungsereignisse zu nennen. Während unter Verwendung des statischen Algorithmus 7.473 (99,19%) Bewegungsereignisse aus insgesamt 7.534 extrahiert werden, gelingt dies mit dem adaptiven Algorithmus für 7.413 (98,39%) Ereignisse.

4.5.3 Aktivitätserkennung mit komprimierten Merkmalsvektoren

Nachfolgend sollen die quantitativen Eigenschaften zur Aktivitätserkennung der nach Abschnitt 4.4 segmentierten und komprimierten Instanzen einzelner Bewegungsereignisse untersucht werden. Dazu werden die Ergebnisse dieses Kapitels mit den ermittelten besten Werten im Rahmen der 10fCV und auf Basis der interpolierten Merkmalsvektoren aus Kapitel 3 direkt verglichen. Um dies zu realisieren, werden die nach Kapitel 3.5.2 konstruierten Ereignisinstanzen mithilfe der in Abschnitt 4.5.1 gewählten Klassifikatoren erneut untersucht. Die Ergebnisse zu diesen Experimenten sind in Tabelle 4.4 zusammengefasst. Auffällig ist, dass der neue, adaptiv segmentierte und etwa um Faktor 1:58

Feature-Vektor	NB Ø (%)	SVM Ø (%)	C4.5 Ø (%)	RF Ø (%)	HPO Ø (%)
interpoliert + statisch	99.09	100	100	100	100
komprimiert + adaptiv	97.57	99.66	99.08	99.85	100

Tabelle 4.4: Ergebnisse der quantitativen Untersuchungen der Aktivitätserkennung für sechs unterschiedliche Körpereigengewichtsübungen mit unterschiedlichen Klassifikatoren, jeweils für ein auf Interpolation basierendes und für ein komprimiertes Design der Merkmalsvektoren.

komprimierte Merkmalsvektor im Rahmen der reinen Aktivitätserkennung für die sechs verschiedenen Körpereigengewichtsübungen kaum einen Performanz-Nachteil gegenüber dem auf Interpolation aufbauenden Vektor aufweist. Unter Verwendung einer manuellen Parameterkonfiguration werden im Zusammenspiel mit dem RF-Klassifikator Erkennungsraten von bis zu 99,85% erreicht. Im automatisierten HPO-Szenario (vgl. 2.3.4) können diese sogar auf 100% gesteigert werden. Diese Ergebnisse übersteigen die Performanz aller bekannten, direkt vergleichbaren Ansätze aus Kapitel 3.6. Auch für Instanzen, welche auf mittels Interpolation erstellten Merkmalsvektoren basieren, kann unter

Zuhilfenahme weiterer Klassifikatoren (SVM, $C4.5$ und RF) sowie im automatisierten Szenario eine Erkennungsrate von 100% realisiert werden. Diese Ergebnisse zeigen, dass das im Rahmen von Abschnitt 4.4.4 definierte, stark komprimierte Design der Merkmalsvektoren nicht nur eine qualitative Klassifizierung, sondern auch eine rein quantitative Aktivitätserkennung mit höherer Zuverlässigkeit als in verwandten Arbeiten erlaubt. Zusätzlich ergeben sich aus der starken Kompression signifikante Laufzeitvorteile, wie die folgenden Untersuchungen zeigen.

4.5.4 Laufzeituntersuchung

Im folgenden werden die Laufzeiten zur Modellbildung, jeweils mit auf Interpolation basierendem Design sowie mit auf Kompression basierendem Design, näher beleuchtet. Zusätzlich werden auch Einblicke in die Auswirkungen der Nutzung variierender Anzahlen von Instanzen zum Modelltraining auf die daraus resultierende Fähigkeit zur Klassifikation der trainierten Modelle gegeben. Für alle Experimente wurde aus Gründen der Vergleichbarkeit das gleiche technische Setup wie in Kapitel 3.6.4 verwendet. Abbildung 4.14 zeigt, dass

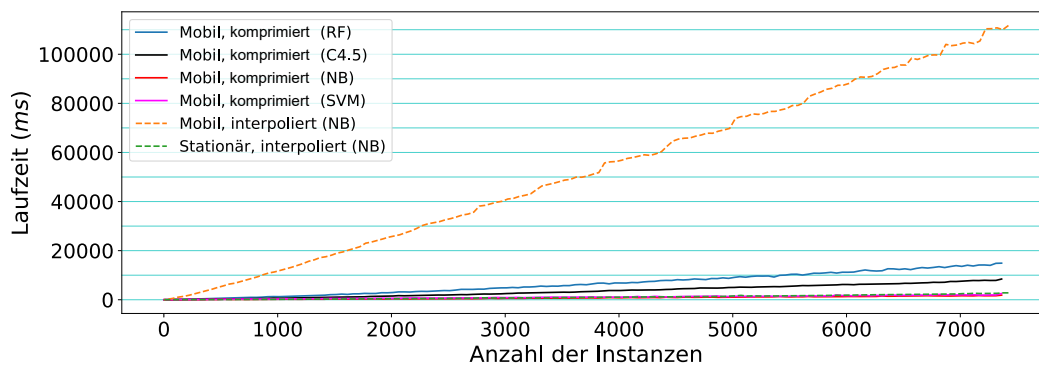


Abbildung 4.14: Laufzeiten zur Bildung von Vorhersagemodellen mit variierenden Klassifikatoren für das mobile und das stationäre Szenario. Dabei wurden unterschiedliche Kombinationen von auf Interpolation basierenden und auf Kompression basierenden Merkmalsvektoren genutzt.

unabhängig vom zur Modellbildung genutzten Klassifikator und dem zugrundeliegenden Design des Merkmalsvektors die Laufzeiten mit der Anzahl der Trainingsinstanzen linear ansteigen. Allerdings weisen alle auf einem komprimierten Design beruhenden Kombinationen im mobilen Szenario, welches im Falle eines Online-Trainings in Echtzeit vorliegt, im Vergleich zu auf Interpolation basierenden Kombinationen signifikante Laufzeitverbesserungen auf. Mit den Kombinationen aus komprimiertem Design und dem NB- sowie dem SVM-Algorithmus werden im mobilen Szenario sogar Laufzeiten erzielt, welche mit denen des stationären Ansatzes aus Kapitel 3.6.4 vergleichbar sind. Die Kom-

binationen aus komprimiertem Design und RF- sowie C4.5-Klassifikator benötigen eine etwas längere Laufzeit zum Modelltraining, was mit der erhöhten Komplexität durch die Aufspannung eines beziehungsweise mehrerer Entscheidungsbäume während der Trainingsphase zu begründen ist. Dafür liefern diese beiden Kombinationen je nach Anwendungsfall deutlich bessere Ergebnisse im Vergleich zu NB-Kombinationen. So können robuste Vorhersagemodelle unter Verwendung von über 3000 Trainingsinstanzen online in etwa 5 s gebildet werden, was die Fähigkeiten des Gesamtsystems zur Rekonfiguration beweist.

Im Rahmen der Laufzeituntersuchungen sind aber auch die Auswirkungen von variierenden Anzahlen von Trainingsinstanzen auf die spätere Vorhersagegenauigkeit der trainierten Modelle von großer Relevanz. Die Untersuchungen wurden dabei wie folgt durchgeführt. Das Testset und das Trainingsset um-

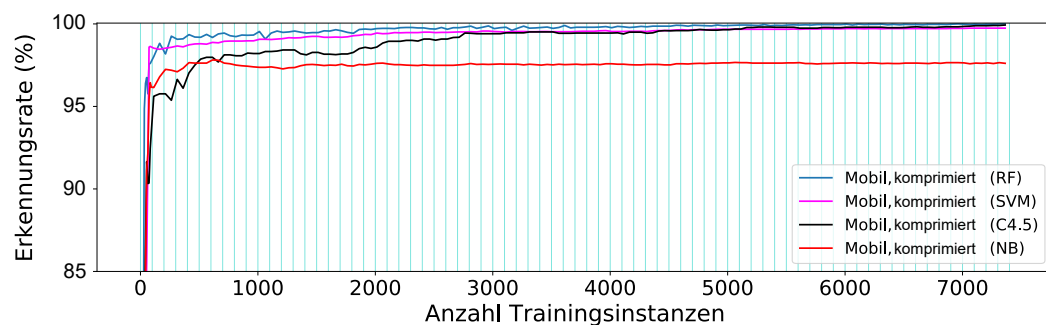


Abbildung 4.15: Korrekte Erkennungsrate für verschiedene Kombinationen aus Klassifikatoren und komprimiertem Design des Merkmalsvektors bei variierender Anzahl von zum Modelltraining genutzter Instanzen im mobilen Online-Szenario.

fassen initial je alle 7.413 extrahierten Instanzen, welche zufällig gemischt und nach Art der Bewegungsaktivität gelabelt sind. Anschließend wird jeweils ein Vorhersagemodell für jeden der untersuchten Klassifikatoren auf Basis des Trainingssets trainiert. Anhand dieser trainierten Modelle wird nun jede einzelne Instanz des Testsets sequentiell klassifiziert. Anschließend wird das Trainingsset um 50 Instanzen verkleinert und der gleiche Vorgang so lange wiederholt, bis das Trainingsset nur noch weniger als 150 Instanzen enthält. Nun wird die Verkleinerung in feingranulareren Schritten von jeweils 10 entfernten Instanzen fortgeführt, um die nun eintretende rapide Verschlechterung der Klassifizierungsergebnisse mit höherer Auflösung sichtbar zu machen. Das Testset hingegen bleibt für alle Iterationen von Training und Klassifizierung durchgehend unangetastet. Auch Vorhersagemodelle, welche auf Basis geringer Anzahlen von Instanzen trainiert wurden, werden mithilfe aller 7.413 Instanzen vertestet, wodurch die Anzahl der nicht im Trainingsset aber dafür im Testset enthaltenen Ereignisse kontinuierlich steigt.

Einblicke in die Ergebnisse zu den betreffenden Experimenten gibt Abbildung 4.15. Alle Klassifikatoren erreichen ab einem Training mit etwa 100 Instanzen bereits eine Vorhersagegenauigkeit von über 95% für alle Klassifika-

ren. Der *RF*- und der *SVM*-Klassifikator konvergieren dabei schnell gegen 100% Genauigkeit, während der *C4.5* hierzu erkennbar mehr Instanzen zum Training benötigt. Nur der *NB*-Klassifikator stagniert auch mit steigender Anzahl von Instanzen im Trainingsset auf Werten von etwa 97%. Damit wurde gezeigt, dass auf Basis von komprimierten Merkmalsvektoren hochperformante Modelle zur Untersuchung von Bewegungsereignissen innerhalb weniger Sekunden sowie online in mobilen Anwendungsszenarien trainiert werden können. Selbst bei einer Anzahl von nur 100 zum Training genutzten Instanzen werden hierbei noch korrekte Erkennungsraten von über 95%, teils sogar über 98%, im Kontext der reinen Aktivitätserkennung für sechs unterschiedliche Ereignisklassen erreicht.

4.6 Zusammenfassung

In diesem Kapitel lag der Schwerpunkt auf der automatisierten, qualitativen Analyse vorhersehbarer und periodischer menschlicher Bewegung, welche über die im Rahmen des vorherigen Kapitels vorgestellte, reine Aktivitätserkennung für komplexe Bewegungsereignisse hinausgeht. Dazu wurde zunächst ein neuer adaptiver Segmentierungsansatz für wiederkehrende Bewegungsereignisse auf Basis von Extrema-Fingerprints entwickelt und vorgestellt. Im Anschluss wurden Möglichkeiten zur Diskretisierung und Formalisierung von qualitativen Beobachtungen innerhalb individueller Bewegungsereignisse definiert, um diese Ereignisse maschinell lesbar zu machen und mithilfe von Werkzeugen des maschinellen Lernens zu untersuchen. Ergebnis dieser Entwicklungen ist ein stark komprimierter Merkmalsvektor, welcher sich nicht nur zur qualitativen Untersuchung von Bewegung eignet, sondern darüber hinaus auch signifikante Verbesserungen hinsichtlich der rein quantitativen Erkennung von menschlicher Aktivität ermöglicht (vgl. Kapitel 3).

Im Rahmen der Evaluation dieses Kapitels wurden diese Konzepte anhand verschiedener Experimente zur qualitativen und quantitativen Untersuchung sowie bezüglich des Laufzeitverhaltens im mobilen Online-Szenario untersucht. Zusätzlich wurde auch die Performanz des adaptiven Segmentierungsverfahrens auf den Prüfstand gestellt und dem statischen Segmentierungsansatz aus Kapitel 3 gegenübergestellt. Dabei konnten im Rahmen der qualitativen Untersuchung 7.413 Instanzen von sechs unterschiedlichen Körpergewichtübungen mit einer gemittelten korrekten Erkennungsrate von 89,7% bei manueller Parameterkonfiguration und mit 99,3% bei automatisierter Hyper-Parameter-Optimierung hinsichtlich ihrer Ausführungsqualität erfolgreich klassifiziert werden. Bislang ist kein vergleichbarer Ansatz bekannt, welcher Bewegungsqualität nur auf Basis von Bewegungsinformationen generisch und ohne gezieltes Erfassen von Fehlern mit entsprechend hoher Sicherheit vorhersagt. Weiterhin hat sich auch die Annahme bestätigt, dass adaptive Segmentierung den Erfolg der qualitativen Analyse positiv beeinflusst. Im direkten Vergleich erzielte der in diesem Kapitel entwickelte adaptive Seg-

mentierungsalgorithmus um bis zu 11,0 Prozent bessere Ergebnisse. Im Mittel wurde über alle getesteten Klassifikatoren ein Performanzvorteil von etwa 5,06 Prozent erreicht. Im Rahmen der quantitativen Erkennung von Aktivität wurde gezeigt, dass der entwickelte komprimierte Merkmalsvektor im Vergleich zum Vorgehen des vorherigen Kapitels trotz einer Kompressionsrate von etwa 1:58 immer noch alle zur Unterscheidung von komplexen Bewegungsvorgängen notwendigen Informationen beinhaltet. Die Vorhersagen konnten im Vergleich zur Aktivitätserkennung in Kapitel 3 je nach gewähltem Klassifikator sogar deutlich verbessert werden und sind auch besser als die bekannter verwandter Arbeiten. Durch Laufzeituntersuchungen konnte nachgewiesen werden, dass mithilfe des komprimierten Merkmalsvektors auch bei großen Mengen von Trainingsdaten das Training von Vorhersagemodellen und damit das Lernen neuer Ereignisse innerhalb weniger Sekunden im mobilen Online-Szenario möglich ist. Die anschließende, eingehende Evaluation der im Rahmen der Laufzeituntersuchungen trainierten Vorhersagemodelle hat deren hohe Leistungsfähigkeit hinsichtlich der Klassifizierung komplexer menschlicher Bewegungsereignisse bewiesen.

Im Rahmen der Untersuchungen zu den vorgestellten Konzepten ergaben sich aber auch neue Herausforderungen und Fragestellungen im Bereich der menschlichen Bewegungsanalyse. Zum einen bietet das komprimierte Design des hier genutzten Merkmalsvektors viele Vorteile hinsichtlich der Systemlaufzeit und eignet sich hervorragend zur generischen, qualitativen Untersuchung von komplexen menschlichen Bewegungsereignissen sowie zu deren Unterscheidung. Allerdings bietet es in seiner jetzigen Form keine Möglichkeit, direkte Rückschlüsse auf konkrete Anomalien innerhalb eines Bewegungsereignisses zu treffen und führt zu einem hohen Informationsverlust. Mit Ausnahme der Integration der absoluten Dauer einzelner Bewegungsereignisse vernachlässigt es weitestgehend die zeitliche Dimension der untersuchten Bewegungsabläufe. Diese kann jedoch signifikante Informationen hinsichtlich der Korrektheit, der Effizienz und des physischen Aufwands einer Bewegung beinhalten. Eine weitere Einschränkung des in diesem Kapitel vorgestellten Konzepts ist der reine Fokus auf wiederkehrende und Periodizitäten beinhaltende Bewegungsaktivität. Viele, aus unterschiedlichen, komplexen Bewegungen bestehende menschliche Aktivitäten beinhalten jedoch keine vorhersehbaren und sich gleichmäßig wiederholenden Bewegungssequenzen und sind somit mit dem bislang vorgeschlagenen Vorgehen nicht zu untersuchen. Diese offenen Fragestellungen und deren nähere Untersuchung sind Thema des folgenden Kapitels, welches ein detailliertes und differenziertes Modell aus Kernkompetenzen zur Untersuchung von nicht vorhersehbarer menschlicher Bewegungsaktivität entwickelt und daraus einen feingranularen Merkmalsvektor zu deren Beschreibung und Untersuchung vorstellt.

5 Qualitative und Quantitative Analyse von nicht vorhersehbarer Bewegungsaktivität

In den vorangegangenen Kapiteln wurde mit SensX die Entwicklung eines Systems zur ganzheitlichen Erfassung und Erkennung von Bewegung mithilfe von am menschlichen Körper befestigten Bewegungssensoren vorgestellt (vgl. Kapitel 3). Anschließend wurde ein Konzept zur Untersuchung von wiederkehrender und periodischer Bewegungsaktivität entwickelt und präsentiert, welches über bislang bekannte, rein quantitative Ansätze zu deren Erkennung hinausgeht und das Treffen von qualitativen Aussagen zu Bewegungsabläufen ermöglicht (vgl. Kapitel 4).

In diesem Kapitel soll nun ein Ansatz zur Untersuchung von nicht vorhersehbarer menschlicher Bewegung, welche keine Periodizitäten oder vorhersehbaren Bewegungsmuster aufweist, vorgestellt werden (vgl. Abschnitt 2.1.1). Daraus ergeben sich nicht nur neue Herausforderungen bezüglich der Segmentierung einzelner Bewegungsabläufe, es muss auch ein neues Modell zur Beschreibung und Formalisierung von menschlicher Bewegung entwickelt werden, welches diese auf diskrete Parameter reduziert und für Algorithmen des maschinellen Lernens lesbar macht. Um die diesbezüglich entwickelten und im Folgenden präsentierten Vorgehensweisen angemessen zu evaluieren, sollen wieder sowohl quantitative als auch qualitative Aspekte von nicht vorhersehbarer Bewegungsaktivität untersucht werden.

Zunächst wird in Abschnitt 5.1 kurz auf diesem Kapitel zugrundeliegende Vorarbeiten eingegangen. Anschließend zeigt Abschnitt 5.2 die Relevanz der angestrebten Konzepte auf und motiviert diese anhand eines Beispiels zur Untersuchung von Kletterbewegung. Danach geben verwandte Arbeiten einen Überblick über Ansätze zur Untersuchung von Kletterbewegung sowie zur Definition von athletischer Leistung im Bezug zur menschlichen Bewegung (vgl. Abschnitt 5.3). Auf dieser Basis wird in Abschnitt 5.4 ein Konzept zur Untersuchung von Kletterbewegung entwickelt, welches anschließend in Abschnitt 5.5 mithilfe einer umfassenden Studie evaluiert wird. Abschließend werden die daraus gewonnenen Erkenntnisse in Abschnitt 5.6 zusammengefasst, diskutiert und in den Kontext zukünftiger Arbeiten gestellt.

Das Vorgehen zur Untersuchung von Bewegungsaktivität allein auf Basis

von Bewegungsinformationen wurde mit vergleichbarer Ganzheitlichkeit, entsprechend feingranularem Maß und dem hier gelegten Fokus auf qualitativen Aspekten bislang noch nicht umgesetzt. Auch der speziell aus dem qualitativen Evaluationsszenario hervorgehende Anwendungsfall, nämlich die automatisierte Bestimmung der Schwierigkeit von Kletterrouten, ist neu.

5.1 Vorveröffentlichungen

Die Kerninhalte und Grundideen bezüglich der automatisierten Erkennung der Schwierigkeit von Kletterrouten und des zur Bewegungserfassung genutzten SensX Sensorsystems wurden bereits im Rahmen von zwei Vorveröffentlichungen publiziert [51, 54]. Wie in Kapitel 1.2 dargestellt, stammen die in den Publikationen und im Folgenden präsentierten Inhalte hinsichtlich Idee, Konzept und Evaluation vom Autor. Die damit verbundenen Abbildungen 5.1, 5.2, 5.3 und 5.4 wurden im Vergleich zur Vorveröffentlichung überarbeitet und in dieser Form zuvor noch nicht publiziert. Das Vorgehen zum Feature-Engineering in Abschnitt 5.4.4 entspricht der ursprünglichen Grundüberlegung, wurde aber in Struktur und Umfang für diese Arbeit noch erweitert. Alle folgend auf diesen Erweiterungen aufbauenden Abbildungen wurden noch nicht veröffentlicht. Die Evaluation in Abschnitt 5.5 wurde auf Basis der Neustrukturierung und Erweiterung des Feature-Engineerings erneut durchgeführt. Die Inhalte zur Unterscheidung einzelner Studienteilnehmer in Abschnitt 5.5 sind ebenfalls durchgehend neu und wurden zuvor nicht veröffentlicht.

5.2 Motivation und Herausforderungen

In den folgenden Abschnitten dieses Kapitels wird ein Konzept zur automatisierten Untersuchung von nicht vorhersehbarer, nicht periodischer Bewegung erarbeitet, umgesetzt und evaluiert. Dabei sollen sowohl qualitative als auch quantitative Aussagen rein auf Basis von Informationen, welche ausschließlich mit Bewegungssensorik zu ermitteln sind, getroffen werden.

Das Anwendungsbeispiel, die Untersuchung von Kletterbewegung, bewegt sich dabei im Kontext der zuvor im Rahmen dieser Arbeit untersuchten Bewegungsaktivitäten aus dem Bereich Sport. Besonders in diesen Szenarien und insbesondere bei der Sportart Klettern ist eine automatisierte und objektive Untersuchung von menschlicher Bewegungsaktivität aufgrund des in dieser Disziplin generell hohen Verletzungsrisikos sinnvoll [146, 209]. Gründe hierfür sind vor allem hohe athletische Anforderungen an die Physis eines Kletterers, aber auch äußere Umstände wie Wetterbedingungen, rauhes Gelände außerhalb von Kletterhallen und nicht zuletzt die Überschätzung der eigenen Fähigkeiten. Zusätzlich wird das Auftreten von Unfällen und Verletzungen durch Überschätzung und Unerfahrenheit noch durch die Entwicklung der letzten Jahre verstärkt, in welchen der Klettersport sich mehr und mehr zu einer

Trendsportart gewandelt hat [168, 176, 169]. Dies spiegelt sich auch im jährlichen Unfallbericht des *Deutschen Alpenvereins* (DAV) für 2016 wider [161]. Zwar ist die Anzahl tödlicher Unfälle in den letzten Jahren nicht unverhältnismäßig stark gestiegen, die Anzahl von Unfällen aufgrund von Selbstüberschätzung und fehlendem Wissen über sachgemäße Materialhandhabung sowie von Rettungsaktionen wegen Blockierens hingegen aber schon. Vom Blockieren spricht man, wenn ein Bergsteiger sich in eine Situation bringt, aus welcher er sich selbst ohne Hilfe nicht mehr befreien kann. Dies kann sowohl durch psychische Belastungen während eines Aufstiegs als auch durch Unterschätzung von dessen physischen Herausforderungen hervorgerufen werden. Zur Unterschätzung von Routenanforderungen kann es dabei durch Unerfahrenheit und falsche Selbsteinschätzung, aber auch durch neu oder falsch bewertete Schwierigkeitsgrade eines Aufstiegs kommen. Dies hängt damit zusammen, dass die Bewertung der Schwierigkeit eines Aufstiegs immer im Kollektiv von vielen Athleten vorgenommen wird. Alle Kletterer, welche einen Aufstieg bewältigt haben, geben eine Bewertung für dessen Schwierigkeit ab und nach einer gewissen Zeit pendelt sich ein finaler Schwierigkeitsgrad aus dem Durchschnitt dieser Einzelbewertungen ein – im Optimalfall entspricht dieser dann der tatsächlichen Routenschwierigkeit. Dieses Problem soll im folgenden durch Entwicklung eines Ansatzes zur automatisierten Erkennung von Schwierigkeitsgraden unterschiedlicher Kletterrouten auf Basis von Bewegungsinformationen adressiert werden.

Dabei ergeben sich verschiedene Herausforderungen, welche im Rahmen dieses Kapitels gelöst werden sollen. Ausgangspunkt soll dabei die Datenbasis sein, welche das in Kapitel 3 vorgestellte SensX Sensorsystem liefert. Dieses System erfasst die Rotations- und Beschleunigungsinformationen des Oberkörpers und der Extremitäten eines Athleten während der Durchführung einer Bewegungsaktivität. Da Kletterbewegung nun im Gegensatz zu zuvor untersuchten wiederkehrenden und periodischen Bewegungen nicht vorhersehbar ist, nicht immer im Vorfeld bekannte Bewegungsmuster enthält und alle Extremitäten völlig frei in verschiedenen Bewegungskombinationen interagieren können, ist ein neues Konzept zur Unterteilung einzelner nicht vorhersehbarer Bewegungsvorgänge in individuelle Sinnabschnitte notwendig. Eine weitere Herausforderung ergibt sich aus der diskreten Beschreibung von nicht vorhersehbarer Bewegung, um auf dieser Basis ein aussagekräftiges Feature-Modell zur Beschreibung von physischer Leistung zu erarbeiten und somit weitere Untersuchungen mithilfe von maschinellen Lernprozessen zu ermöglichen. Dieses Feature-Modell soll so generisch sein, dass es die Unterscheidung der Schwierigkeit unterschiedlicher Kletterrouten auf Basis der Bewegungsinformationen unterschiedlicher Kletterer zulässt. Weiterhin soll es jedoch auch so spezifisch sein, dass es auch Rückschlüsse auf die Eigenheiten der individuellen Bewegungsmuster ermöglicht.

5.3 Grundlagen und verwandte Arbeiten

Folgend wird einerseits ein Überblick über verwandte Arbeiten zur sensor-gestützten Untersuchung von Kletterbewegung gegeben, andererseits zur Beschreibung von Kletterbewegung in Form von athletischer Leistung und daraus resultierenden physischen Anforderungen an einen Kletterer. Diese Beschreibung ist die essentielle Grundlage bei der Entwicklung eines feingranularen Feature-Modells zur Beschreibung von menschlicher Bewegung, welches die Projektion von physischen Kernfähigkeiten eines Athleten auf diskrete Merkmale erlaubt.

5.3.1 Untersuchung von Kletteraktivität

Im Laufe der letzten Jahre wurden einige wenige Arbeiten zur automatisierten und Sensor-basierten Untersuchung von Kletterbewegung und zur Unterstützung von Klettertraining auf Basis von Bewegungssensorik vorgestellt. Dieser Abschnitts gibt einen Überblick über diese.

Pansiot et al. versuchten in diesem Kontext mit einer frühen Arbeit, das Kletterkönnen von Kletterern mithilfe der Informationen eines am Ohr getragenen triaxialen Beschleunigungssensors zu bewerten [147]. Aus den Beschleunigungsinformationen des Accelerometers, welche während dem Aufstieg einer Kletterroute aufgenommen werden, leiten die Autoren spezifische Maße für Kletterkönnen, wie beispielsweise Flüssigkeit von Bewegungen, Kraftaufwand und Ausdauer, ab. Vor der Generierung eines Merkmalsvektors interpolieren und glätten sie zunächst die Rohdaten, anschließend werden diese in sequentiellen Fenstern von $50ms$ Länge verarbeitet. So extrahieren sie 27 Features pro Kletterroute, unter anderem die Aufstiegszeit, den Durchschnitt und die Standardabweichung der Beschleunigung in jeder Dimension, die gesamte Beschleunigung sowie Unterschiede in Mittelwerten und Standardabweichungen in der ersten und der zweiten Hälfte des Aufstiegs. Die Autoren erklären allerdings nicht, auf welche Eigenschaften eines Kletterers und dessen Könnens diese Features genau abzielen. Um nun eine Untersuchung der Kletterbewegung durchführen zu können, wurden im Rahmen einer Studie zur Datenerhebung 19 Routen von 4 Kletterern aufgenommen. Im Rahmen der Evaluation versuchten die Autoren zum einen ihre Merkmale manuell zu interpretieren und Korrelationen auszumachen, zum anderen verwendeten sie mit einem *Expectation Maximization*-Klassifikator Werkzeuge des nicht überwachten maschinellen Lernens, um ihre Daten nach den jeweiligen Kletterern zu clustern. Dabei kam es zu zwei Fehlzuordnungen bei 19 Aufstiegen. Zusammenfassend kann gesagt werden, dass Pansiot et al. eine der ersten Arbeiten zum Thema Analyse von Kletterbewegung auf Basis von Beschleunigungsinformation liefern. Allerdings bleibt die Begründung zur Auswahl der Features weitgehend unklar und die Studie zur Evaluation ist aufgrund ihres geringen Umfangs und teils manuellen Analysevorgehens wenig aussagekräftig. Nichtsdestotrotz bietet die Arbeit einige interessante Aspekte und Ideen zur Untersuchung von

menschlicher Bewegungsaktivität.

Auch Ladha et al. versuchen automatisiert Rückschlüsse auf die Leistung und das Können von Kletterern auf Basis von Bewegungsinformation zu treffen [117]. Hierzu benutzen die Autoren ein Accelerometer, welches am Handgelenk des Kletterers befestigt ist und die Beschleunigungsinformationen während der Begehung einer Route zur späteren Offline-Analyse aufnimmt. Dazu werden die erfassten Bewegungsinformationen zunächst einem Resampling unterzogen und in einem ersten Segmentierungsvorgang von anderen Aktivitäten getrennt. Dazu nutzen sie ein neuronales Netz in Form einer *Restricted Boltzmann Machine* (RBM). Die zum Training benötigten Features extrahieren die Autoren mithilfe eines überlappenden Sliding-Window Algorithmus. Auf die Art der benutzten Features zur Unterscheidung von Aktivitäten gehen sie nicht genauer ein. In einer zweiten Segmentierungsphase werden Kletterbewegungen mithilfe von Schwellwerten für den Energiegehalt eines Signals innerhalb einer erneuten Sliding-Window Prozedur in Übergangs- und Ruhephasen unterteilt, welche zur weiteren Untersuchung dienen. Angelehnt an Pansiot et al. definieren die Autoren anschließend die Berechnung von vier Kernfähigkeiten eines Kletterers, welche dessen Kletterkönnen qualitativ beschreiben sollen [147]. Diese benennen sie mit Kraft P , Kontrolle C , Stabilität S und Geschwindigkeit V . Für diese vier Kernfähigkeiten berechnen sie anschließend über mathematische Operationen je einen Zahlenwert und konstruieren daraus eine vierdimensionale Repräsentation $s = \{P, C, S, V\} \in \mathbb{R}^4$ des Kletterkönnens eines Athleten. Zur Evaluation dieser vierdimensionalen Repräsentation verglichen die Autoren diese mit den offiziellen Ergebnissen eines Kletterwettbewerbs mit 47 Teilnehmern und wiesen dabei vorhandene Korrelationen nach. Somit lässt sich zusammenfassen, dass Ladha et al. das Können von Klettern auf Basis eines komprimierten, simplifizierten Merkmalsvektors, grob einschätzen können. Wie der folgend beschriebene Ansatz von Kosmolla et al. zeigt, ist dabei der erste Segmentierungsschritt zur Extraktion von Kletteraktivität im Vergleich zu seinem Nutzen unnötig komplex strukturiert, differenzierte und präzise Aussagen hinsichtlich qualitativer Aspekte der untersuchten Bewegungen fehlen bislang.

Im Rahmen von *ClimbSense* versuchen Kosmolla et al., unterschiedliche Kletterrouten auf Basis von Bewegungsinformationen voneinander zu unterscheiden [111]. Auch hier wird mit an den Handgelenken getragener Sensorik gearbeitet, welche Beschleunigung, Rotation und Änderungen in der magnetischen Flussdichte an den Handgelenken der Kletterer aufzeichnet. Im Rahmen der Segmentierung findet zuerst eine Trennung von Kletteraktivität und sonstigen Aktivitäten statt, anschließend wird die Kletteraktivität für jeden Sensor in Ruhe- und in Übergangsphasen unterteilt. Im Gegensatz zu Ladha et al. benutzen die Autoren zur Erkennung von Kletteraktivität die Orientierung der Hände; der zweite Segmentierungsprozess wird ebenfalls auf Basis von Schwellwert-basierten Vergleichen des Energiegehalts sequentiell gebildeter Fenster gleicher Länge realisiert. Zur anschließenden Bildung von Featu-

res werden die Orientierungen der Handgelenke der Kletterer mithilfe eines Codebuches in Zeichenfolgen übersetzt. Deren Distanzen zu zuvor gebildeten Routen-Templates werden anschließend unter Zuhilfenahme der Levenshtein-Distanz berechnet. Diese drückt die Distanz zweier Zeichenketten zueinander anhand der Anzahl von Einfüge-, Lösch- und Ersetzoperationen aus, welche benötigt werden, um Zeichenkette 1 in Zeichenkette 2 zu überführen [179]. Zur Evaluation ihres Konzepts zur Unterscheidung von Kletterrouten nahmen Kosmolla et al. insgesamt 50 Aufstiege mit 8 Kletterern an 5 unterschiedlichen Kletterrouten auf und konnten diese unter Betrachtung der Sensorinformationen beider Handgelenke mit einer Rate von 100% korrekt erkennen. Ein großes Problem dieses Ansatzes ist allerdings sein statisches Klassifizierungskonzept. Dadurch, dass nur die Orientierung der Hände während der Ruhephasen untersucht wird, geht jegliche temporale Information bezüglich eines Aufstiegs sowie solche zur eigentlichen physischen Leistung des Kletterers verloren. Der Merkmalsvektor kann genau nur zu einem Zweck, nämlich der Unterscheidung von Routen verwendet werden. Des Weiteren ergeben sich durch die Anwendung des Distanzvergleichs via Templating mit der Levenshtein-Distanz erhebliche Ressourcen- und Laufzeitaufwände, welche durch die $O(mn)$ -Komplexität des Algorithmus auftreten [214]. Dafür zeigen die Autoren eine sehr einfache und effektive Methode zur Segmentierung von Bewegungsaktivität auf.

Feeken et al. präsentieren mit *ClimbingAssist* ein System zur Unterstützung des Klettertrainings von Kletteranfängern [61]. Dazu wird auf Basis vom Druck der menschlichen Füße sowie mithilfe von Beschleunigungsinformationen der Kletterstil der Probanden untersucht. Der Druck wird über eine drucksensitive Folie am Schuh der Athleten gemessen, während die Beschleunigung mithilfe eines im Schuh integrierten Arduinoboards via Accelerometer ermittelt wird. Über Vibrationen, welcher mit einer Vibrationseinheit an die Füße der Athleten übertragen werden, kann Feedback bezüglich des Kletterstils an die Nutzer weitergegeben werden. Feeken et al. konnten im Rahmen ihrer Studie mit 9 Teilnehmern nicht eindeutig nachweisen, dass das System zur Verbesserung des Kletterstils beitrug – auch weil viele Probanden Probleme hatten, ihren Stil direkt anhand des gegebenen Vibrationsfeedbacks anzupassen. Die Autoren halten fest, dass das System von der Ausstattung beider Füße mit Sensorelementen profitieren könnte.

Zusammenfassend lässt sich feststellen, dass bislang entwickelte Konzepte zur Untersuchung von Kletteraktivität meist auf eingeschränkten Feature-Modellen operieren. Andere Arbeiten versuchen Erkenntnisse via manueller Untersuchung zu Erlangen oder über Korrelation von Parametern Zusammenhänge zwischen Leistung und Bewegungsinformation herzustellen. Trotz verschiedener Einschränkungen liefern sie wichtige Erkenntnisse und damit wertvolle Grundlagen hinsichtlich der Untersuchung von Bewegungsaktivität.

5.3.2 Beschreibung von physischer Leistung im Klettersport

Um Kletteraktivität adäquat untersuchen zu können, muss zunächst geklärt werden, von welchen physischen Komponenten physische Leistung in diesem Kontext abhängig ist. Diese Komponenten sollen anschließend in Abschnitt 5.4.4 als Hauptgrundlage zur Entwicklung eines feingranularen Modells zur Diskretisierung und Beschreibung von physischer Leistung im Kontext menschlicher Bewegungsaktivität dienen. Dazu werden zum einen Definitionen zur Beschreibung körperlicher Fitness aus zuvor vorgestellten verwandten Arbeiten verwendet, zum anderen wurde aber auch die Beschreibung physischer Leistungskomponenten von Kletteraktivität in einschlägiger Fachliteratur miteinbezogen.

Goddard und Neumann unterteilen die Komponenten, aus welchen sich die Performanz eines Kletterers zusammensetzt, in verschiedene Kategorien, beispielsweise psychologische Aspekte wie Angst oder Konzentration, äußerliche Bedingungen wie Felstyp und Klima sowie taktische Aspekte wie Erfahrung und Wissen [69]. Die für diese Arbeit relevanteste Kategorie ist die körperliche Fitness, welche sie über Stärke, Ausdauer und Flexibilität definieren. Später sollen mithilfe von SensX Rückschlüsse darauf gezogen werden. Ähnlich definieren auch Watts et al. ein Leistungsprofil von Kletterern anhand von Anthropometrie, Kraft, Ausdauer, Maximalkraft und aerober Kapazität [206, 207, 1]. Die Anthropometrie steht dabei für die Lehre der Ermittlung und Anwendung der Maße des menschlichen Körpers, welche ebenfalls für dessen theoretische physische Leistungsfähigkeit relevant ist.

Ladha et al. definieren Kraft, Kontrolle, Stabilität und Geschwindigkeit als Kernfähigkeiten des Kletterns [117]. Kraft beschreibt für sie den Transfer von isometrischer Stärke in Bewegung. Das Greifen von weit voneinander entfernten Griffen erfordert einen erhöhten Kraftaufwand im Vergleich zu geringeren Entfernungen. Isometrische Stärke meint dabei die größtmögliche Kraft, die ein Nerv-Muskel-System gegenüber einem statischen, unüberwindbaren Hindernis ausüben kann. Kontrolle beschreibt die Fähigkeit zu weichen und gleichmäßigen Bewegungen während Übergänge von Griff zu Griff, wobei ungenügende Kontrolle zu Zittern und stockenden Bewegungen führt. Stabilität definiert sich für Ladha et al. als sicheres Halten ohne Zittern wenn sich Gliedmaßen in Ruheposition, beispielsweise auf einem Griff oder Tritt, befinden. Geschwindigkeit wird beschrieben durch die absolute Zeit, welche für einen Aufstieg benötigt wird.

Pansiot et al. versuchen mit einem Modell, welches die Aufstiegs geschwindigkeit, die Flüssigkeit der Kletterbewegung, das Stärke-zu-Gewicht-Ratio und die Ausdauer eines Kletterers betrachtet, Rückschlüsse auf die Kletterleistung zu ziehen [147].

Alle diese Definitionen zur Beschreibung von Leistung im Klettersport werden nachfolgend im Rahmen von Abschnitt 5.4.4 zu Kernfähigkeiten zusam-

mengefasst und verdichtet. Anschließend dienen diese Kernfähigkeiten als Ausgangspunkt für die Entwicklung eines feingranularen Feature-Modells zur Beschreibung physischer, menschlicher Leistung. Dieses soll eine Projektion der qualitativen Beschreibung von Leistung auf quantitative, diskrete Features ermöglichen und damit auch die anschließende Bildung eines Merkmalsvektors zur ganzheitlichen Beschreibung von Kletterbewegung.

5.4 Konzept zur Untersuchung nicht vorhersehbarer Bewegungsaktivität

Nachfolgend wird ein Konzept zur ganzheitlichen, automatisierten und feingranularen Untersuchung von nicht vorhersehbarer, menschlicher Bewegungsaktivität am Beispiel von Kletterbewegung vorgestellt. Dieses stützt sich auf eine Prozesskette des überwachten maschinellen Lernens und gibt Einblicke in Vorgehen zur Segmentierung von nicht vorhersehbarer menschlicher Bewegung sowie zu deren Projektion auf eine maschinell lesbare Beschreibung in Form von diskreten Merkmalsparametern. Grundgedanke des folgenden Konzepts ist dabei, dass es anhand von Energieaufwand und individuellen Bewegungsmustern möglich ist, Kletterbewegung nachzuvollziehen, zu bewerten und Rückschlüsse auf die Fähigkeiten eines Athleten zu ziehen.

5.4.1 Definition einer Prozesskette

Zur Entwicklung eines Konzepts zur Untersuchung von Kletteraktivität wird zunächst eine Prozesskette definiert, welche sich an der generisch definierten Grundarchitektur aus Kapitel 2.4.2 orientiert. Diese umfasst die (1) *Erfassung von Kletterbewegung* mithilfe des SensX Sensorsystems, die (2) *Vorverarbeitung und Segmentierung* der Sensorinformationen, das anschließende (3) *Feature-Engineering* zur Entwicklung eines expressiven Merkmalsvektors und schlussendlich die (4) *Klassifikation* von Kletterereignissen mithilfe maschineller Lernprozesse. Ein Überblick über diesen Ablauf wird in Abbildung 5.1 gegeben.

Durch die nicht vorhersehbare Natur von Kletterbewegung ergeben sich für die Umsetzung einer solchen Prozesskette nun verschiedene neuer Herausforderungen. Im Gegensatz zu der in den Kapiteln 3 und 4 untersuchten wiederkehrenden, oft symmetrischen und periodischen Bewegungsaktivität handelt es sich nun um Sequenzen bestehend aus nicht zwingend vorhersehbaren Bewegungsereignissen. Deshalb können Dauer und Art von Kletterereignissen nur schwerlich anhand von Korrelation bestimmt werden. Auch steht es jedem Kletterer frei, wie schnell er eine Route meistert und wie viele unterschiedliche Bewegungen er innerhalb eines Aufstiegs durchführt. Flecken et al. beschreiben diesen Umstand damit, dass es beim Klettern im Gegensatz zu anderen Sportarten, wie beispielsweise Turnen, keine Bewegungsvorschriften gibt und

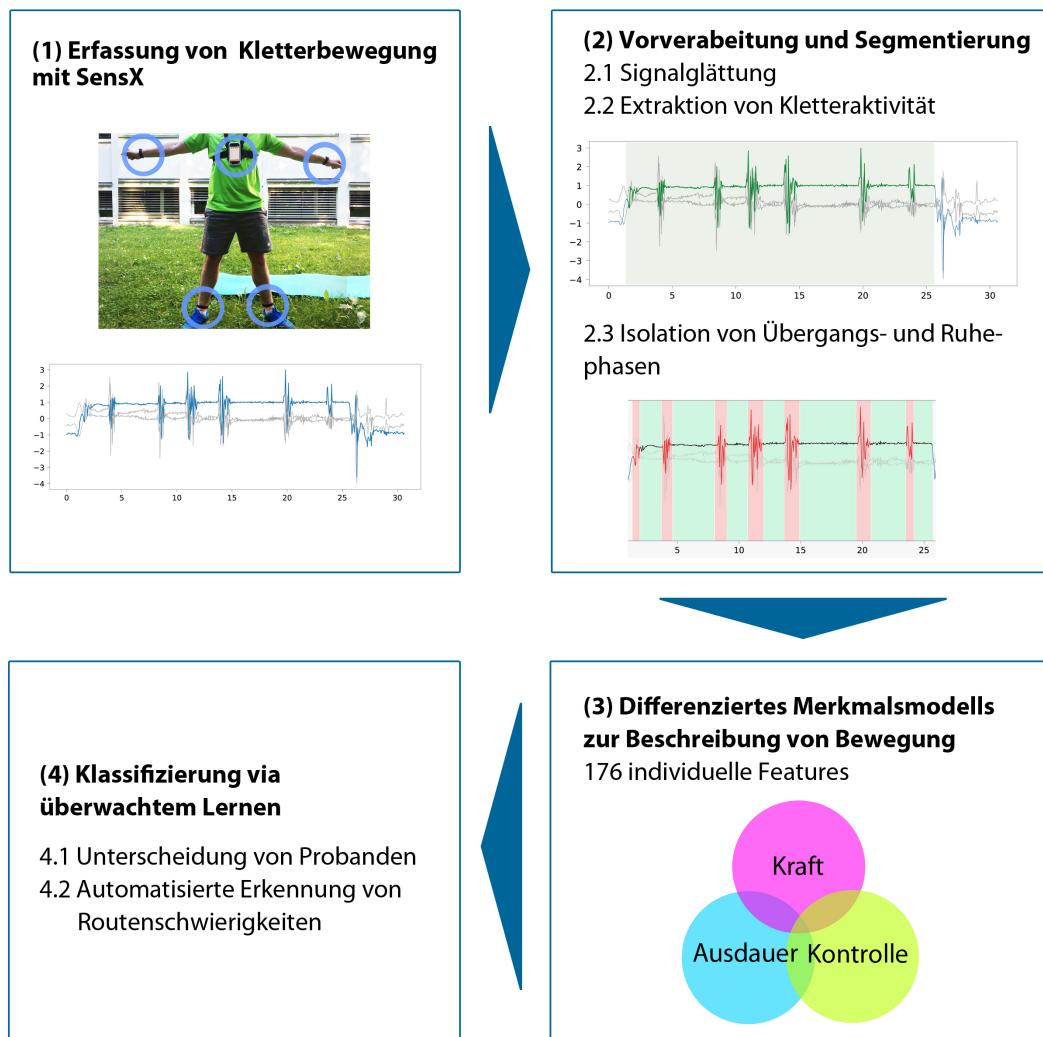


Abbildung 5.1: Übersicht über die adaptierte Prozesskette zur automatisierten, sensorgestützten Untersuchung von Kletteraktivität: (1) Erfassung von Bewegung mit dem SensX Sensorsystem, (2) die Vorverarbeitung und der zweistufige Segmentierungsprozess, (3) das Feature-Engineering und (4) die Klassifizierung mithilfe von überwachten maschinellen Lernalgorithmen.

unzählige Bewegungen zum Lösen ein und desselben Kletterproblems genutzt werden können [64]. Daraus ergibt sich die Notwendigkeit zur Definition eines neuen Segmentierungskonzepts, welches nicht auf Korrelationen und Periodizitäten fußt, sondern individuelle Ereignissegmente anhand neu zu definierender Eigenschaften identifiziert und isoliert.

Im Bereich des Feature-Engineerings besteht die größte Herausforderung in der Identifikation von Merkmalen, welche Kletterbewegung zum einen so generisch betrachten, dass sich Schwierigkeiten routenübergreifend vorhersagen lassen, zum anderen aber Kletterbewegung so spezifisch beschrieben wird, dass

auch individuelle Kletterleistung noch nachvollziehbar ist. Dazu muss ein neues Modell entwickelt werden, mit welchem körperliche Bewegungsaktivität beschrieben und in diskrete Merkmale überführt werden kann. In den folgenden Abschnitten werden zu diesen Herausforderungen individuelle Lösungskonzepte entwickelt und präsentiert.

5.4.2 Erfassung von Kletterbewegung

Die Erfassung von Kletterbewegung ist der erste Schritt der in Abbildung 5.1 vorgestellten Prozesskette und wird mit dem in Kapitel 3 konzipierten SensX Sensorsystem durchgeführt. Dieses bietet durch den von stationärer Infrastruktur unabhängigen Aufbau eine optimale Grundlage zur ganzheitlichen Erfassung von Kletterbewegung außerhalb von Laborbedingungen und in Echtzeit. Auch in diesem Anwendungsfall liefert das System wieder 30 Einzelsignale als Grundlage für weitere Analysen, welche von den 5 individuellen Sensorplattformen für Rotation und Beschleunigung in X -, Y - und Z -Richtung erfasst werden. Wichtig ist hierbei auch die Fähigkeit zur individuellen Erfassung von Bewegungen der einzelnen menschlichen Gliedmaßen, da diese während einer Kletteraktivität unabhängig voneinander agieren.

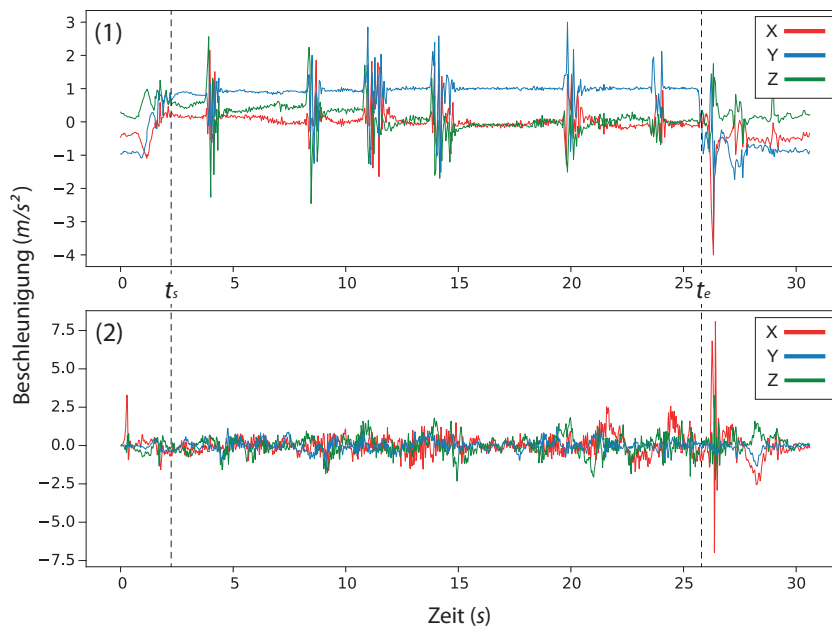


Abbildung 5.2: Beschleunigungsinformationen des SensX Sensorsystems während eines Kletteraufstiegs in X -, Y - und Z -Richtung für die linke Hand (1) und die Brust eines Kletterers (2).

In Abbildung 5.2 ist in (1) die Beschleunigung der linken Hand und in (2) die der Brust eines Kletterers während des Aufstiegs an einer Boulderoute zu sehen. Bei Boulderrouen handelt es sich um kurze, technisch anspruchsvolle

Kletterrouten, welche aufgrund von geringer Gesamthöhe auch ohne Seilsicherung geklettert werden können. In der Beschleunigungsinformation der linken Hand ist der Beginn der Kletteraktivität t_s deutlich auszumachen. Die t_s folgenden Phasen mit gleichmäßig geringem Signalausschlag repräsentieren Ruhephasen, während denen die Hand auf einem Griff liegt. Die Abschnitte mit großen Werten der Signalamplitude hingegen repräsentieren Übergangsphasen, in welchen der Kletterer seine Hand von einem Griff zum nächsten bewegt. Auch das Ende des Aufstiegs ist zum Zeitpunkt t_e ist deutlich zu erkennen, gefolgt von sehr hohen Beschleunigungswerten während des Absprungs des Kletterers vom höchsten Punkt der Boulderoute. Ein besonders differenziertes Bild des Aufstiegs liefert dabei das Signal der Hand in Y -Richtung, welches die höchste Signaldynamik beinhaltet. Die Beschleunigung des Brustsensors, welche in (2) zu sehen ist, liefert hingegen keine eindeutigen Rückschlüsse auf einzelne Bewegungsereignisse während des Aufstiegs. Lediglich zum Absprung nach Beendigung der Route ist wieder ein besonders hoher Amplitudenausschlag zu beobachten. Basierend auf diesen ersten Beobachtungen wird in den folgenden Abschnitten ein zweistufiges Segmentierungsverfahren zur Extraktion einzelner Ereignisse definiert.

5.4.3 Zweistufiger Segmentierungsprozess

Die Segmentierung ist der zweite essentielle Schritt innerhalb der in Abbildung 5.1 dargestellten Prozesskette. Dieser unterscheidet sich in mehreren Punkten maßgeblich von den zuvor in Kapitel 3 und 4 vorgestellten, statischen und adaptiven Segmentierungsansätzen. Hauptgrund hierfür ist vor allem die Art der Bewegung, welche nun nicht mehr wiederkehrender Art ist (vgl. Kapitel 2.1.1). In diesem Szenario können zukünftige Bewegungsereignisse oder Aktivitäten nicht vorhergesehen werden, es existieren keine durchgängigen Periodizitäten und es sind auch keine Muster wie beispielsweise Extrema-Fingerprints bekannt, welche die Identifikation und Extraktion einzelner Bewegungsereignisse ermöglichen. Ferner sind die Ereignisse innerhalb nicht vorhersehbarer Bewegungsaktivität auch in ihrer zeitlichen Dauer weder beschränkt noch im Vorherein definierbar.

Aus diesen Vorbedingungen ergibt sich, dass ein Segmentierungsprozess zur Extraktion von nicht vorhersehbaren Ereignissen nach anderen Regeln arbeiten muss als ein solcher für wiederkehrende Bewegungsereignisse. Keogh et al. beschreiben in diesem Konzept zwei unterschiedliche Verarbeitungsparadigmen zur Segmentierung von menschlicher Bewegungsaktivität innerhalb von kontinuierlichen Zeitreihen: 1) die explizite Segmentierung durch Identifikation von Start- und Endpunkten eines Ereignisses anhand von bekannter Bewegungsdauer und bekanntem Bewegungsmuster (vgl. auch Kapitel 4.3.1) und (2) die implizite Segmentierung durch Anlegen eines sequentiellen *Sliding-Window*, wobei immer der jeweils aktuelle Fensterinhalt untersucht und anschließend einer neuen oder einer gegebenenfalls noch andauernden Aktivität zugeordnet

wird [99]. Eine explizite Segmentierung ist aufgrund der fehlenden initialen Informationen über die Dauer oder den Beginn eines bevorstehenden Kletterereignisses zunächst möglich. Deshalb wird im Folgenden ein zweistufiger, impliziter Segmentierungsprozess definiert, welcher die Anforderungen nicht vorhersehbarer Bewegungsaktivität adressiert. Dieser trennt Kletteraktivität von vor und nach einem Aufstieg aufgetretenen Aktivitäten; anschließend wird die Kletteraktivität dann in einzelne Kletterereignisse, sogenannte Ruhe- und Übergangsphasen, unterteilt. Diese Segmente bilden anschließend die Grundlage für die Extraktion von expressiven Features zur Beschreibung von menschlicher Bewegung in Abschnitt 5.4.4.

5.4.3.1 Segmentierung von Kletteraktivität

Die Trennung von Kletteraktivität und sonstigen Aktivitäten vor und nach einem Aufstieg ist der erste Schritt des zweistufigen Segmentierungsprozesses. Durch die voneinander unabhängigen Bewegungsabläufe der einzelnen Gliedmaßen während einer Kletteraktivität müssen, im Gegensatz zur Segmentierung von vorhersehbarer Bewegung, alle Signale einer Sensorplattform unabhängig von denen anderer Plattformen betrachtet und segmentiert werden. Dies ist dadurch bedingt, dass die Bewegungen der einzelnen Gliedmaßen nicht zwangsläufig einem ganzheitlichen Bewegungsereignis zuzuordnen sind. Somit werden zur Segmentierung zunächst alle Signale einer Sensorplattform isoliert von denen anderer Plattformen betrachtet. Wie bei der Segmentierung vorhersehbarer Bewegung in Kapitel 3.5.2.1 wird dazu zuerst das Most-Meaningful-Signal S_{MMS} anhand der maximalen Standardabweichung identifiziert. Auch hier gilt die Annahme, dass dieses aufgrund seiner hohen Dynamik am besten zur Segmentierung einzelner Bewegungsereignisse geeignet ist. Dazu wird die Standardabweichung (vgl. Berechnung 3.1 in Kapitel 3.5.2.1) σ_i jedes einzelnen der $i = 6$ Signale S_i jeder Plattform (Rotation und Beschleunigung für 3 Dimensionen in X -, Y - und Z -Richtung) berechnet, wobei s_i der x -te Messwert und μ der Erwartungswert ist. Daraus ergibt sich eine Menge M mit i Standardabweichungen $M = \{\sigma_1, \dots, \sigma_i\}$, wobei das S_{MMS} sich nun durch Identifikation von deren Maximum ermitteln lässt:

$$S_{MMS} = \underset{\sigma \in M}{\mathbf{arg\,max}}\{\sigma_x\} \quad (5.1)$$

Zur Verdeutlichung zeigt Abbildung 5.3 in (1) die unsegmentierten Beschleunigungssignale der linken Hand eines Kletterers während eines Aufstiegs. Wie schon zuvor erwähnt, lassen sich der Beginn der Kletteraktivität t_s und deren Ende t_e schon mit bloßem Auge deutlich erkennen. Das S_{MMS} ist in diesem Fall die in Blau dargestellte Beschleunigung in Y -Richtung. Um nun, wie in (2) von Abbildung 5.3 dargestellt, die grün hinterlegte Kletteraktivität von sonstigen Aktivitäten zu trennen, wird im Folgeschritt die Beschleunigung des S_{MMS} an beiden Armen des Athleten ausgewertet. Grundüberlegung ist da-

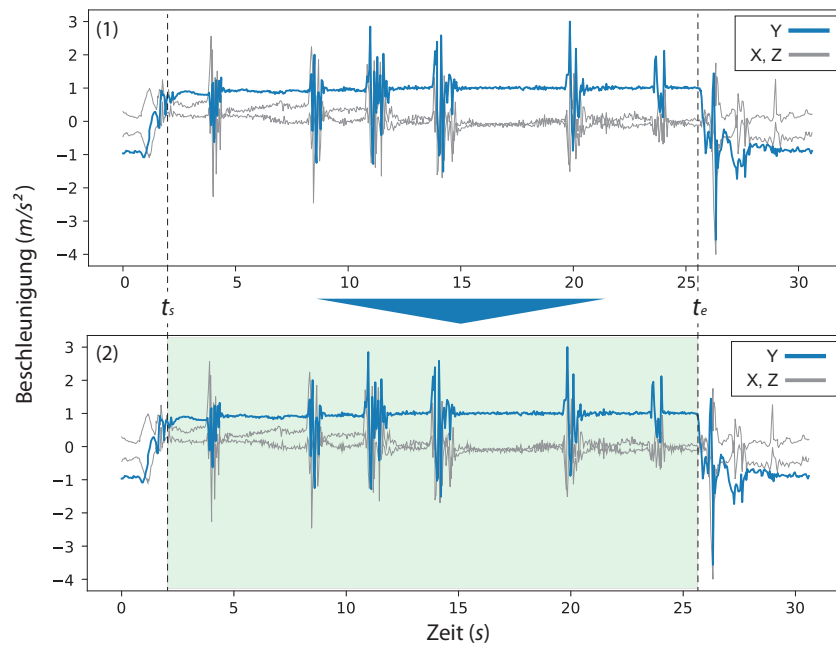


Abbildung 5.3: Beschleunigungsinformationen des SensX-Sensorsystems während eines Kletteraufstiegs in X -, Y - und Z -Richtung für die linke Hand (1) und die daraus extrahierte Kletteraktivität in (2), welche grün hervorgehoben ist.

bei, dass eine Kletteraktivität beginnt, sobald beide Hände eines Kletterers gleichzeitig erhoben sind. Aufgrund der Sensoranordnung von SensX kann dies mithilfe eines Schwellwerts für die Beschleunigung a_{up} ermittelt werden. Ist die Beschleunigung a beider Hände gleichzeitig größer als a_{up} , so hat die Kletteraktivität begonnen. Solange mindestens eine Hand erhoben ist, a größer a_{up} ist und keine Fallbeschleunigung auftritt, dauert die Kletteraktivität an. Sind beide Hände nicht mehr erhoben oder tritt eine starke Fallbeschleunigung auf, so ist der Klettervorgang beendet. Um dies nun sequentiell für die gesamte Zeitreihe zu überprüfen, wird ein ein Sliding-Window der Länge Δt_{sl} über S_{MMS} geführt und die Orientierung der Hand über die auftretende Beschleunigung für jeden einzelnen Fensterinhalt überprüft. t_s ist hier als der Zeitpunkt definiert, an dem a für beide Hände größer a_{up} ist; t_e definiert sich aus dem Zeitpunkt zu dem a für beide Hände kleiner a_{up} ist oder ein Absprung anhand einer starken negativen Beschleunigung erkannt wurde (siehe starker negativer Amplitudenausschlag nach t_e in Abbildung 5.3).

5.4.3.2 Segmentierung von Ruhe- und Aktivitätsphasen

Im zweiten Segmentierungsschritt sollen nun einzelne Bewegungsereignisse aus der in Abschnitt 5.4.3.1 extrahierten Kletteraktivität segmentiert werden. Weitere Grundlagen hierfür sind zum einen das zuvor ermittelte S_{MMS} sowie die in Kapitel 2.3.3 erläuterte Beziehung zwischen Signalamplitude und aufgewende-

ter Bewegungsenergie. In Abbildung 5.4 lassen sich in (1) deutlich sogenannte Übergangsphasen und Ruhephasen erkennen, welche auch in verwandten Arbeiten bereits zur Untersuchung von Kletterbewegung verwendet wurden [117, 111]. Die Übergangsphasen bezeichnen Zeiträume, in welchen ein Klet-

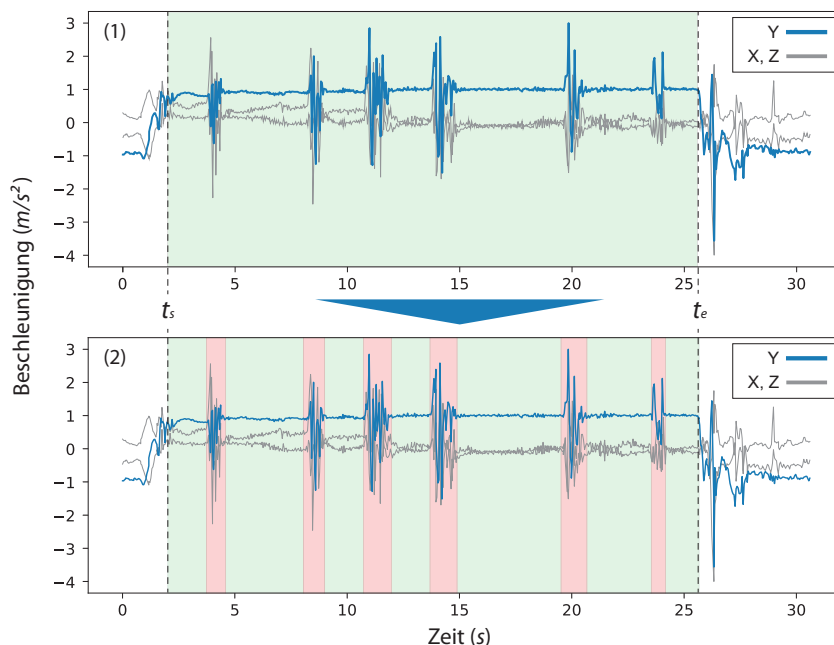


Abbildung 5.4: Darstellung der Überführung des Ergebnisses des ersten Segmentierungsschrittes (1) in den zweiten Schritt (2), bei welchem Ruhephasen (in Grün) und Übergangsphasen (in Rot) voneinander getrennt werden.

terer seine Extremitäten von einem Griff oder Tritt zum nächsten bewegt, wodurch diese durch große Amplitudenausschläge und damit durch einem hohen Energieaufwand gekennzeichnet sind. Ruhephasen hingegen beschreiben die Zeiträume zwischen den Übergangsphasen, in welchen Füße oder Hände statisch auf einem Tritt oder Griff ruhen und vergleichsweise wenig Energie aufgewendet wird. Diese Phasen werden als separate Bewegungsereignisse auf Basis von aufgewendeter Energie extrahiert und anschließend als Grundlage des Feature-Engineerings verwendet.

Um Ruhe- und Übergangsphasen auf Basis von im Signal enthaltener Energie voneinander zu trennen, wird die Energie, wie zuvor im ersten Segmentierungsschritt, innerhalb des sequentiell angelegten *Sliding Window*-Algorithmus untersucht. Der zweite Segmentierungsschritt findet hierbei parallel zum ersten innerhalb der selben *Sliding Window*-Sequenz der Länge i statt, wodurch ein zweiter Signaldurchlauf und damit zusätzlicher Rechenaufwand verhindert wird. Der zweite Segmentierungsschritt setzt also direkt nach einer Zuordnung des Fensterinhalts zu einer Kletteraktivität ein. Ist diese gegeben, so wird angelehnt an die Signalamplitude als Indikator des Energiegehalts eines individuel-

len Fensters F_i die Summe E_f der Standardabweichungen der Beschleunigung in X-, Y-, und Z-Richtung gebildet:

$$E_f = \sigma_x + \sigma_y + \sigma_z \quad (5.2)$$

Liegt E_f über dem zuvor definierten Schwellwert e_{trans} , welcher die Zugehörigkeit zu einer Übergangsphase angibt, so wird der Fensterinhalt als einer Übergangsphase zugehörig, andernfalls als einer Ruhephase zugehörig, deklariert. Ausgangsinformation nach vollständigem Signaldurchlauf sind also zwei geordnete Mengen, Ruhephasen $R = \{R_1, \dots, R_x\}$ und Übergangsphasen $T = \{T_1, \dots, T_y\}$, wobei alle R_x und alle T_y jeweils aus dem Inhalt der ihnen zugeordneten Einzelfenster F_i in zeitlicher geordneter Reihenfolge bestehen und als einzelne Ereignisse behandelt werden. Die einzelnen, segmentierten Ereignisse sind zur Veranschaulichung in Schritt (2) der Abbildung 5.4 dargestellt. Die Übergangsphasen sind dabei rot, die Ruhephasen grün hinterlegt. Ein wichtiger Parameter zur Segmentierung ist in diesem Kontext die Wahl der Fensterlänge l der einzelnen Fenster F_i . Ist diese zu groß gewählt, so kann es sein dass Fensterinhalte bei zu restriktiv gewählten Schwellwerten der falschen Phase zugeordnet werden. Bei zu kleinen Fenstergrößen ist der Prozess anfälliger gegenüber Störinformationen und der Rechenaufwand zur Segmentierung erhöht sich erheblich.

Nach der Identifizierung der beiden Ereignismengen auf Basis der Beschleunigungsinformationen findet nun die Segmentierung der Rotationsinformation anhand der individuellen Zeitstempel jeder Einzelphase statt. Der gesamte Vorgang wird dabei für jede Sensorplattform individuell durchgeführt. Um die Bedingung der bestehenden Kletteraktivität auch für die Sensorsignale an den Füßen eines Athleten zu erfüllen, werden bei deren Segmentierung nur Fenster F_i betrachtet, für deren Zeitraum anhand von a_{up} im ersten Segmentierungsschritt auf Basis der Informationen der Armsensoren bereits eine Kletteraktivität erkannt wurde. Die Informationen des Brustsensors wurden auf Grund des undifferenzierten Signalbildes, welches auch in Abbildung 5.2, (2) zu sehen ist, für diese Arbeit nicht in den Segmentierungsprozess miteinbezogen.

5.4.4 Differenziertes Merkmalsmodell zur Beschreibung menschlicher Bewegung

Wie in den Kapiteln 2.3.3 und 4.4.4 gezeigt, lassen sich anhand der Dynamik und des Amplitudenausschlags eines menschlichen Bewegungssignals Rückschlüsse auf die während eines Bewegungsereignisses aufgewendete Energie ziehen. Auf Basis dieser Vorüberlegungen soll mithilfe der zuvor segmentierten Bewegungsereignisse nun ein neues, expressives Merkmalsmodell zur differenzierten Beschreibung menschlicher Bewegung entwickelt werden. Mit diesem sollen Bewegungsaktivitäten, welche aus einer variierenden Anzahl unterschiedlicher Ereignisse mit jeweils unterschiedlicher Dauer bestehen, beschrie-

ben und untersucht werden können. Dabei soll dieses menschliche Bewegung ganzheitlich beschreiben und sowohl quantitative Rückschlüsse als auch qualitative Aussagen über die zugrundeliegende Bewegungsaktivität ermöglichen. Dafür wurden zunächst die in Abschnitt 5.3.2 aufgeführten Einzelkomponenten zur Beschreibung von physischer, menschlicher Leistung und des Kletterkönnens zu drei im Rahmen dieser Arbeit identifizierten Kernfähigkeiten verdichtet. Diese sind in Abbildung 5.5 zu sehen und gliedern sich in Kraft, Kontrolle und Ausdauer. Zusätzlich sind im Schaubild die Definitionen und Umschreibungen einzelner Komponenten aus verwandten Arbeiten ihren jeweiligen Kernfähigkeiten zugeordnet. Da sich die einzelnen Kernfähigkeiten in der Realität immer bis zu einem gewissen Grad gegenseitig bedingen und somit nahezu niemals isoliert betrachtet werden können, sind sie auch hier teilweise überlappend dargestellt.

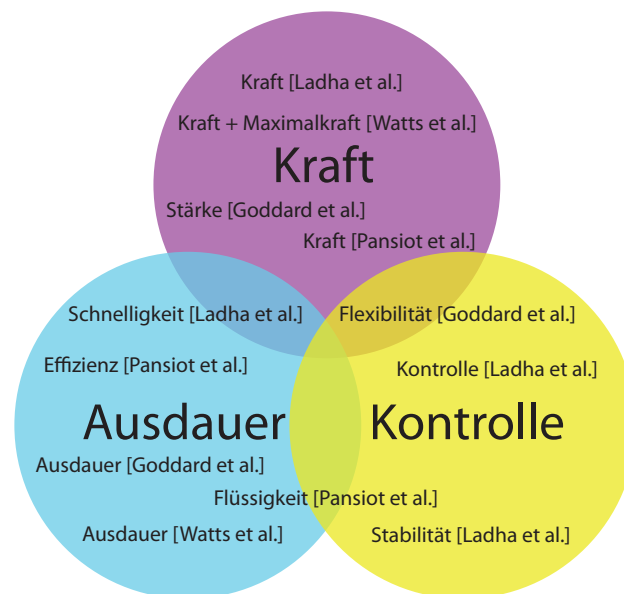


Abbildung 5.5: Verdichtung der verschiedenen Komponenten zur Beschreibung von physischer, menschlicher Leistung und Kletterkönnens aus verwandten Arbeiten auf die drei in dieser Arbeit aggregierten Kernfähigkeiten.

Mithilfe dieser Gliederung soll im Anschluss die strukturierte Extraktion von diskreten Features und deren Zuordnung auf die drei Kernfähigkeiten erleichtert werden. Analog dazu definiert sich Kraft vor allem aus der aufgewendeten Bewegungsenergie in Form von Rotations- und Beschleunigungsbewegungen während der Übergangsphasen. Wie in Abbildung 5.5 zu sehen, wird diese Kernfähigkeit auch in verwandten Arbeiten zumeist mit Kraft [117, 207, 206, 147] oder Stärke [69] beschrieben. Es gibt aber auch Definitionen, welche in den Randgebieten zu anderen Kernfähigkeiten angetragen werden können. So zum Beispiel Schnelligkeit [117], welche zum einen über

Kraft erreicht wird, zum anderen durch ihr konstantes und längerfristiges Vorhandensein aber auch Rückschlüsse auf die Ausdauer eines Athleten zulässt. Ein weiterer Grenzfall ist Flexibilität [69], welche Goddart et al. als die Fähigkeit Kraft kontrolliert, über einen längeren Zeitraum adäquat an ein Problem angepasst anzuwenden, verstehen.

Ausdauer wird nachfolgend vor allem durch die Gleichmäßigkeit des Energieaufwands und der Flüssigkeit von Bewegungen innerhalb beider Phasen sowie durch deren temporale Gleichmäßigkeit definiert. In verwandten Arbeiten wird dazu oft der Begriff Ausdauer verwendet [69, 207, 206], aber auch der Begriff Effizienz [147], welchen die Autoren als einen gleichmäßigen Energieaufwand über einen längeren Zeitraum definieren. Direkt im Grenzbereich angetragen ist für diese Kernfähigkeit die Flüssigkeit von Bewegung [147]. Diese beschreibt die Fähigkeit eines Athleten, Bewegungen langfristig flüssig und gleichförmig durchzuführen und zeugt zum einen von einer hohen Kontrolle, zum anderen aber auch von großer Ausdauer.

Kontrolle wird überwiegend durch niedrige und gleichmäßige Energieaufwände innerhalb der Ruhephasen definiert. Zusätzlich definieren Ladha et al. Stabilität als wichtigen Faktor, um Bewegungen kontrolliert durchführen zu können [117].

In den anschließenden Abschnitten wird nun die Zuordnung und Berechnung von diskreten Features aus den Beschleunigungs- und Rotationsinformationen des SensX Sensorsystems unter Berücksichtigung dieser drei aggregierten Kernfähigkeiten erläutert. Ergebnis des Prozesses ist ein diskreter Merkmalsvektor mit 176 Features in 14 Untergruppen und drei Oberkategorien, welcher anschließend im Rahmen der Evaluation zur Beschreibung und Untersuchung von nicht vorhersehbarer Bewegungsaktivität genutzt wird. Die Oberkategorien der Features unterteilen sich dabei in 1) Features der Übergangsphasen mit drei Untergruppen, 2) Features der Ruhephasen mit ebenfalls 3 Untergruppen und 3) temporale Features mit acht Untergruppen. Die Untergruppen ergeben sich dabei aus den nachfolgenden Überlegungen zur Extraktion der Features.

5.4.4.1 Features der Übergangsphasen

Es wurden drei Untergruppen mit jeweils 24 Features aus den Segmenten der Übergangsphasen extrahiert. Diese ergeben sich aus den Signalen der Rotation und der Beschleunigung in drei Dimensionen der vier externen Sensorplattformen (vgl. Tabelle 5.1). Daraus werden 72 Features extrahiert, welche die erste der drei Oberkategorien von Merkmalen bilden. Im folgenden werden die einzelnen Sensorplattformen mit oben links (OL, für den linken Arm), oben rechts (OR, für den rechten Arm), unten links (UL, für das linke Bein) und unten rechts (UR, für das rechte Bein bezeichnet). Die aus den Übergangsphasen extrahierten Features beschreiben vornehmlich die aufgewendete Kraft eines Athleten, wobei, wie bereits in Abbildung 5.5 zu sehen, immer Grenzbereiche zu anderen Kernfähigkeiten existieren. Abbildung 5.6 und Tabelle 5.1 geben dabei Einsicht in diese Zuordnung: alle Features der Übergangsphasen geben

Aufschluss über den Kraftaufwand P , zwei Gruppen (1 und 2) lassen sich in zweiter Ordnung auch der Ausdauer E zuordnen. Gruppe 2 lässt auch Rückschlüsse auf die Kontrolle C zu. Die $N = 24$ Features der ersten Gruppe X_1

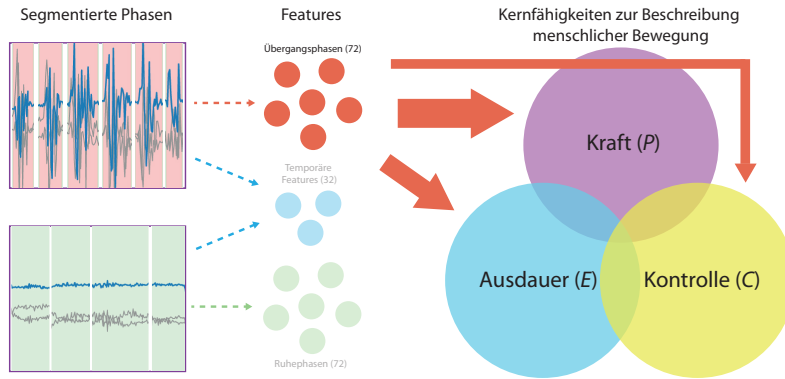


Abbildung 5.6: Übersicht über die Zuordnung der Features aus den Übergangsphasen auf die drei verdichteten Kernfähigkeiten Kraft P , Ausdauer E und Kontrolle C .

Gruppe	Feature-Beschreibung	Zuordnung	Anzahl
X_1	Durchschnittlich aufgewendete Energie (arithm. Mittel) innerhalb aller Übergangsphasen, jeweils individuell für die Rotation und die Beschleunigung jeder Extremität (OR,OL,UR,UL) in X-, Y- und Z-Richtung	P, E	24
X_2	Standardabweichung der aufgewendeten Energie innerhalb aller Übergangsphasen, jeweils für die Rotation und die Beschleunigung aller Extremitäten (OR,OL,UR,UL) in X-, Y- und Z-Richtung	P, E, C	24
X_3	Maximal aufgewendete Energie innerhalb aller Übergangsphasen, jeweils für die Rotation und die Beschleunigung aller Extremitäten (OR,OL,UR,UL) in X-, Y- und Z-Richtung	P	24

Tabelle 5.1: Übersicht und Beschreibung der aus den Übergangsphasen extrahierten Features unter Berücksichtigung ihrer Gruppenzugehörigkeit und der Zuordnung zu spezifischen Kernfähigkeiten.

der Übergangsphasen ergeben sich dabei aus dem arithmetischen Mittel der Summe der Messwerte aller Übergangsphasen $T = \{T_1, \dots, T_n\}$ mit der Anzahl n , jeweils einzeln in X-, Y-, und Z-Richtung für jeden individuellen Sensor:

$$X_1 = \{(x_i)\}_{i=1}^N, \quad \text{wobei} \quad x_i = \frac{1}{n} \sum_{y=1}^n T_y \quad (5.3)$$

Jedes $T_y = (s_1, \dots, s_m)$ ist dabei definiert als diskrete Zeitreihe mit m Messwerten s . Die Features der Gruppe 1 werden vornehmlich der Beschreibung von P zugeordnet, da sie Auskunft über den durchschnittlichen Energieaufwand geben. Zusätzlich lassen sich in Verbindung mit Gruppe 2 und 3 aber auch begrenzte Rückschlüsse auf E und C ziehen.

Die zweite Gruppe von Features X_2 der Übergangsphasen ergibt sich aus der Standardabweichung innerhalb der Summe der Messwerte aller Übergangsphasen $T = \{T_1, \dots, T_n\}$ mit der Anzahl n . Diese wird ebenfalls wieder einzeln in X-, Y-, und Z-Richtung für jeden individuellen Sensor berechnet, was wie in Gruppe 1 auch in Gruppe 2 zu einer Gesamtanzahl von $N = 24$ Features führt. Jedes $T_y = (s_1, \dots, s_m)$ ist dabei definiert als diskrete Zeitreihe mit m Messwerten s :

$$X_2 = \{(x_i)\}_{i=1}^N, \quad \text{wobei} \quad x_i = \sqrt{\text{Var}(T)} \quad (5.4)$$

Dabei lassen die Features der zweiten Gruppe Rückschlüsse auf alle drei Kernfähigkeiten P , E , und C in Form von Fluktuation des Energieaufwandes und dessen Gleichmäßigkeit im Verhältnis zum Durchschnitt und zum Maximum zu.

Als dritte und letzte Feature-Gruppe X_3 mit der Feature-Anzahl $N = 24$ der Übergangsphasen wird der maximale Amplitudenmesswert t_{max} innerhalb aller Messwerte der einzelnen Übergangsphasen $T = \{T_1, \dots, T_n\}$ mit der Anzahl n für jede einzelne Dimension jedes Sensors festgelegt. Jedes $T_y = (s_1, \dots, s_m)$ ist dabei definiert als diskrete Zeitreihe mit m Messwerten s :

$$X_3 = \{(x_i)\}_{i=1}^N, \quad \text{wobei} \quad x_i = t_{max} = \mathbf{arg\,max}\{(T_1, \dots, T_n)\} \quad (5.5)$$

Die Features der dritten Gruppe sind mit ihrer Aussage über den maximalen Energieaufwand der Kernfähigkeit P zugeordnet. Eine Übersicht über alle Features der Übergangsphasen inklusive ihrer Gruppenzugehörigkeit und der Zuordnung zu spezifischen Kernfähigkeiten gibt Tabelle 5.1.

5.4.4.2 Features der Ruhephasen

Auch aus den Ruhephasen werden analog zum Extraktionsprozess der Übergangsphasen 72 Features aus den 24 Grundmengen von Ruhephasen extrahiert. Dabei bilden sie die zweite der drei Oberkategorien von Merkmalen. Wie Abbildung 5.7 und Tabelle 5.2 verdeutlichen, sind diese gleichmäßig in erster Ordnung der Beschreibung der Ausdauer E eines Athleten, in zweiter Ordnung der Kontrolle C über die Bewegung zugeordnet. Generell werden E und C hierbei über möglichst geringe Energieaufwände während der Ruhephasen ausgedrückt, da ein ausdauernder Athlet ruhig und ohne unnötige Bewegung, wie beispielsweise Zittern, auf einem Tritt steht oder sich an einem Griff hält [151, 83, 206].

Die erste Feature-Gruppe der Ruhephasen X_4 mit $N = 24$ Features ergibt

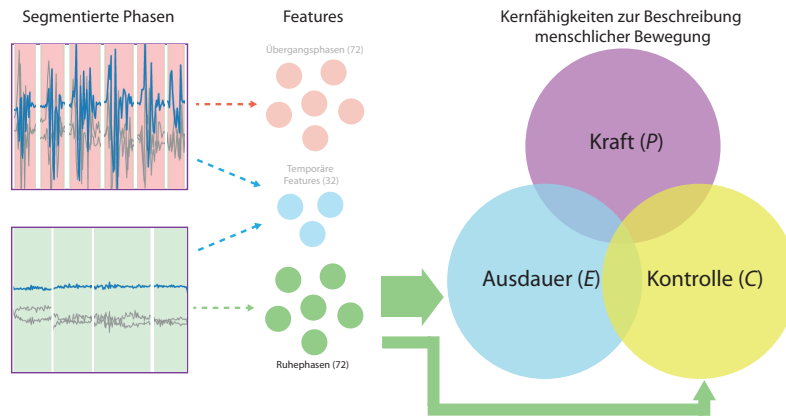


Abbildung 5.7: Übersicht über die Zuordnung der Features aus den Ruhephasen auf die drei verdichteten Kernfähigkeiten Kraft P , Ausdauer E und Kontrolle C .

sich analog zur Berechnung 5.3 von X_1 in Abschnitt 5.4.4.1 aus dem arithmetischen Mittel der Summe der Messwerte aller Ruhephasen $R = \{R_1, \dots, R_n\}$ mit der Anzahl n , jeweils einzeln in X-, Y-, und Z-Richtung für jeden individuellen Sensor. Auch hier ist jedes $R_y = (s_1, \dots, s_m)$ definiert als diskrete Zeitreihe mit m Messwerten s .

Die zweite Gruppe von Features X_5 der Ruhephasen ergibt sich analog zur Berechnung 5.4 der Gruppe X_2 in Abschnitt 5.4.4.1 aus der Standardabweichung innerhalb der Summe der Messwerte aller Ruhephasen $R = \{R_1, \dots, R_n\}$ mit der Anzahl n . Diese wird ebenfalls einzeln in X-, Y-, und Z-Richtung für jeden individuellen Sensor berechnet, was auch für X_5 zu einer Gesamtanzahl von $N = 24$ Features führt.

Für die dritte Feature-Gruppe X_6 mit der Feature-Anzahl $N = 24$ der Ruhephasen wird nun analog zur Berechnung 5.5 der maximale Amplitudenmesswert r_{max} innerhalb aller Messwerte der einzelnen Ruhephasen $R = \{R_1, \dots, R_n\}$ mit der Anzahl n für jede einzelne Dimension jedes Sensors als Feature ermittelt. Jedes $R_y = (s_1, \dots, s_m)$ ist dabei definiert als diskrete Zeitreihe mit m Messwerten s .

5.4.4.3 Temporale Features

Die temporalen Features bilden die dritte Feature-Kategorie und werden sowohl aus den Übergangsphasen als auch den Ruhephasen extrahiert. Grundgedanke bei der Bildung der temporalen Features ist, dass vor allem die menschliche Ausdauer nur in Verbindung mit einer zeitlichen Komponente ausgedrückt werden kann. Da die Größe der Ausdauer auch maßgeblich die Kontrolle über eine Bewegung beeinflusst, sind die 32 temporalen Features in erster Ordnung E und in zweiter Ordnung C zugeordnet. Dies ist auch in Abbildung 5.8 und Tabelle 5.3 nachzuvollziehen. Im Anwendungsfall Klettern wäre es naheliegend, absolute Aufstiegszeiten eines Kletterers in einen Merkmalsvektor zu

Gruppe	Feature-Beschreibung	Zuordnung	Anzahl
X_4	Durchschnittlich aufgewendete Energie (arithm. Mittel) innerhalb aller Ruhephasen zusammen, jeweils für die Rotation und die Beschleunigung aller Extremitäten (OR,OL,UR,UL) in X -, Y - und Z -Richtung	E, C	24
X_5	Standardabweichung der aufgewendeten Energie innerhalb aller Ruhephasen zusammen, jeweils für die Rotation und die Beschleunigung aller Extremitäten (OR,OL,UR,UL) in X -, Y - und Z -Richtung	E, C	24
X_6	Maximal aufgewendete Energie innerhalb aller Ruhephasen zusammen, jeweils für die Rotation und die Beschleunigung aller Extremitäten (OR,OL,UR,UL) in X -, Y - und Z -Richtung	E, C	24

Tabelle 5.2: Übersicht und Beschreibung der aus den Ruhephasen extrahierten Features unter Berücksichtigung ihrer Gruppenzugehörigkeit und der Zuordnung zu spezifischen Kernfähigkeiten.

integrieren [147]. Dies ist allerdings aufgrund von dessen späterer Anwendung zur automatisierten, routenübergreifenden Klassifizierung von unterschiedlichen Schwierigkeitsgraden nicht zielführend, da die absolute Aufstiegsdauer beim Vergleich von unterschiedlichen Kletterrouten kein verallgemeinerbarer Indikator für Schwierigkeit ist. So können kurze Routen durchaus schwieriger sein als längere Routen, bei gleicher Länge muss nicht unbedingt die Schwierigkeit gleich sein, es kann aber auch die längere Route entsprechend schwieriger sein. Deshalb werden nun relative, temporale Features entwickelt und extrahiert, um diesem Fakt ohne die Verwendung einer absoluten Zeitdauer Rechnung zu tragen. Wie bei den zuvor präsentierten Oberkategorien von Merkmalen gilt aber auch bei den temporalen Features, dass jede Untergruppe oft mehreren Kernfähigkeiten zuzuordnen ist. Die erste, der in Tabelle 5.3 dargestellten Merkmalsuntergruppen, X_7 , ergibt sich aus dem arithmetischen Mittel der Summe aller individuellen Zeitfenster Δt_y aller Übergangsphasen $T_{trans} = (\Delta t_1, \dots, \Delta t_n)$ der Menge n für die $N = 4$ externen Sensorplattformen:

$$X_7 = \{(x_i)\}_{i=1}^N, \quad \text{wobei} \quad x_i = \frac{1}{n} \sum_{y=1}^n \Delta t_y \quad (5.6)$$

Die Untergruppe X_8 ergibt sich darauf aufbauend aus der Standardabweichung der individuellen Zeitfenster aller Übergangsphasen je Sensorplattform:

$$X_8 = \{(x_i)\}_{i=1}^N, \quad \text{wobei} \quad x_i = \sqrt{\text{Var}(T_{trans})} \quad (5.7)$$

Analog zu Berechnung 5.6 wird die Untergruppe X_9 ermittelt, das Vorgehen zur Ermittlung der Untergruppe X_{10} geschieht auf Basis von Berechnung 5.7.

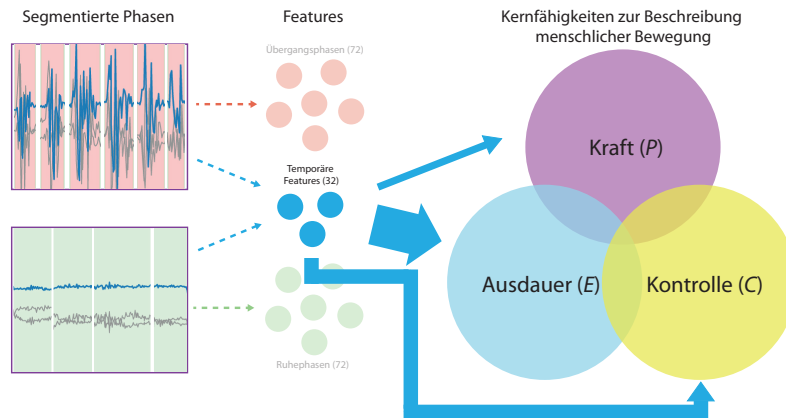


Abbildung 5.8: Übersicht über die Zuordnung der Features aus den Ruhe- und Übergangsphasen auf die drei aggregierten Kernfähigkeiten Kraft P , Ausdauer E und Kontrolle C .

Beide erhalten als Eingangsmenge die individuellen Zeitfenster Δt_y aller Ruhephasen $T_{rest} = (\Delta t_1, \dots, \Delta t_n)$ der Menge n für die $N = 4$ externen Sensorplattformen. Die letzten vier Feature-Gruppen X_{10} bis X_{14} mit insgesamt 16 Features geben zusätzlich Auskunft über die längste Übergangsphase, die kürzeste Übergangsphase, die längste Ruhephase und die kürzeste Ruhephase, jeweils für jede einzelne der 4 externen Sensorplattformen.

5.4.5 Bildung von diskreten Merkmalsvektoren

Um im Rahmen der folgenden Evaluation verschiedene Algorithmen des maschinellen Lernens zur automatisierten, qualitativen und quantitativen Untersuchung von nicht vorhersehbarer Bewegungsaktivität heranzuziehen, werden die im Rahmen der 14 Untergruppen berechneten einzelnen Features nun einem Merkmalsvektor $F = (x_1, \dots, x_i)$ mit der Länge $i = 176$ hinzugefügt, womit er sich über die Gesamtheit aller extrahierten Features definiert. Mit diesem diskreten Vektor lassen sich nun unterschiedliche Abfolgen nicht vorhersehbarer menschlicher Bewegungen beschreiben. Die enthaltenen diskreten Features sind den drei Kernfähigkeiten zur differenzierten Beschreibung von menschlicher Bewegungsaktivität P , E und C zugeordnet und sollen nachfolgend auf ihre Aussagekraft hin überprüft werden. Zusätzlich zur Betrachtung von Energieaufwänden wurden dabei, im Gegensatz zum Vorgehen aus verwandten Arbeiten, auch temporale Features mit relativer Gewichtung extrahiert, welche eine generische, Routen-übergreifende, temporale Beschreibung von Bewegung ohne Bindung an absolute Zeitfenster ermöglichen sollen.

Gruppe	Feature-Beschreibung	Zuordnung	Anzahl
X_7	Durchschnittliche Dauer (arithm. Mittel) der Übergangsphasen aller Extremitäten (OR,OL,UR,UL)	P, E	4
X_8	Standardabweichung der Dauer der Übergangsphasen für alle Extremitäten (OR,OL,UR,UL)	E, C	4
X_9	Durchschnittliche Dauer (arithm. Mittel) der Ruhephasen aller Extremitäten (OR,OL,UR,UL)	E	4
X_{10}	Standardabweichung der Dauer der Ruhephasen für alle Extremitäten (OR,OL,UR,UL)	E	4
X_{11}	Dauer der längsten Übergangsphase je Extremität (OR,OL,UR,UL)	E, C	4
X_{12}	Dauer der kürzesten Übergangsphase je Extremität (OR,OL,UR,UL)	E, C	4
X_{13}	Dauer der längsten Ruhephase je Extremität (OR,OL,UR,UL)	E	4
X_{14}	Dauer der kürzesten Ruhephase je Extremität (OR,OL,UR,UL)	E	4

Tabelle 5.3: Übersicht und Beschreibung aller temporalen Features, extrahiert aus Ruhe- und Übergangsphasen und unter Berücksichtigung ihrer Gruppenzugehörigkeit sowie der Zuordnung zu spezifischen Kernfähigkeiten.

5.5 Evaluation

Im Rahmen der Evaluation soll nun zunächst die Aussagekraft des zuvor konstruierten Merkmalsvektors F überprüft werden. Dazu werden zwei unterschiedliche Experimente durchgeführt. Zum einen soll mithilfe einer quantitativen Klassifizierung untersucht werden, ob die 172 extrahierten Features spezifisch genug sind, um einzelne Kletterer der im Folgenden beschriebenen Studie anhand ihrer Bewegungsinformationen und der temporalen Zusammenhänge innerhalb der Ruhe- und der Übergangsphasen voneinander zu unterscheiden. Zum anderen soll überprüft werden, ob derselbe Merkmalsvektor zugleich generisch genug spezifiziert wurde, um die Schwierigkeit unterschiedlicher Kletterrouten auf Basis von dessen Informationsgehalt automatisiert zu bestimmen.

Weiterhin sollen auch die einzelnen Features auf ihren individuellen Informationsgehalt untersucht werden. Dazu werden unterschiedliche Feature-Subsets mithilfe der Korrelation von Klassifizierungsergebnissen zu den einzelnen Positionen des Merkmalsvektors ermittelt. Anschließend kann auf deren Basis eine erneute Untersuchung mit unterschiedlichen Algorithmen des maschinellen Lernens erfolgen.

5.5.1 Studie zur Erfassung von Kletterbewegung

Um die vollständige Prozesskette, welche in Abbildung 5.1 definiert wurde, zu durchlaufen und eine aussagekräftige Überprüfung zur Leistungsfähigkeit der zuvor entwickelten Konzepte aus den Abschnitten 5.4.3 und 5.4.4 zu realisieren, wurde eine umfangreiche Studie zur Erhebung von Kletteraktivität durchgeführt. Im Folgenden wird deren zugrundeliegendes Studiendesign zusammengefasst.

Insgesamt wurden im Rahmen der Studie 152 Aufstiege von 16 Studienteilnehmern an 13 verschiedenen Boulder Routen erfasst. Bei Boulder Routen handelt es sich um kurze, technisch anspruchsvolle Kletter Routen mit einer Länge von meist 4 – 5 m, welche ohne Sicherung geklettert werden. Im Gegensatz zu Kletter Routen, welche bei professionellen Wettbewerben meist eine Länge von 12 – 18 m aufweisen und mit Sicherung geklettert werden, ist bei Boulder Routen oft eine hohe Schnellkraft und damit eine schnelle Muskelkontraktion sowie eine kontrollierte Stabilisierung des Körpers nach dynamischen Bewegungen erforderlich. Dennoch sind beide Arten des Kletterns miteinander verwandt und bis zu einem gewissen Grad in Bewegung und Technik miteinander vergleichbar [60, 209]. Es wurden Boulder Routen zur Datenerhebung ausgewählt, da durch fehlende Notwendigkeit bezüglich des Wissens über Sicherungstechnik und Materialanwendung ein größerer Kreis an Studienteilnehmern realisiert werden konnte. Außerdem sind Energie- und Kraftaufwand während eines Aufstiegs bei einer Boulder Route tendenziell größer als bei mehr auf Ausdauer ausgelegten Kletter Routen, wodurch ein aussagekräftigeres Feature-Bild zu erwarten ist (vgl. Abbildung 5.6).

Die während der Studie aufgenommenen Routen umfassen drei unterschiedliche Schwierigkeitsgrade, welche mit den Kategorien Blau, Grün und Schweden (blau-gelb) farblich unterschieden werden können. Dabei entspricht jede farbliche Kategorie einem Bereich von *Fontainebleau Climbing Grades*, bei welchen es um ein allgemein anerkanntes Bewertungssystem zur Bewertung der Schwierigkeit von Boulderproblemen handelt. Im Fall der durchgeführten Studie entsprach Blau einem Schwierigkeitsbereich von 1a-2c, Schweden einem Bereich von 3a-4b und Grün einer Schwierigkeit von 4b-5c. Bei Schweden und Grün gab es also eine Überschneidung der Schwierigkeitsbereiche. Generell ist das Schwierigkeitslevel der drei gewählten Farbkategorien als sehr leicht bis mittel zu bezeichnen, wodurch garantiert werden sollte, das möglichst viele Studienteilnehmer alle Routen klettern können. Zusätzlich besitzt jede Route mehr Griffe und Tritte, als zu deren Bewältigung notwendig sind, was unterschiedliche Aufstiegsvarianten zur Lösung des Boulderproblems zulässt.

Von den 16 Probanden waren 12 männlich und 4 weiblich. Durchschnittlich kletterten diese seit 1,69 Jahren, wobei die absolute Klettererfahrung zwischen 0 und 9 Jahren variierte. Zusätzlich wurden auch Alter, Gewicht und Größe der Teilnehmer erfasst. Diese Informationen werden in Tabelle 5.4 zusammengefasst.

Nach Erhebung der persönlichen Informationen erfolgte für jeden Teilneh-

Geschlecht	Ø Alter (Jahre)	Ø Größe (cm)	Ø Gewicht (kg)
m	19,3	178,67	68,3
w	19	169,25	55,25

Tabelle 5.4: Übersicht über Geschlecht sowie durchschnittliches Alter, Größe und Gewicht der Studienteilnehmer.

mer eine kurze Aufwärmphase von 5 Minuten zur Vorbeugung von Verletzungen an Muskeln, Sehnen, Bändern und Gelenken. Anschließend wurde die Aufnahme der Daten mithilfe des in Kapitel 3 vorgestellten SensX Sensor-systems durchgeführt. Die Aufstiegsinformationen wurden dabei zur anschließenden Offline-Untersuchung auf dem als Brustsensor integrierten Smartphone zwischengespeichert.

5.5.2 Feature-Review und Subset-Bildung

Um die Aussagekraft des im vorhergehenden Abschnitt 5.4.4 entwickelten Merkmalsvektors im Vorfeld zu überprüfen und auch Teilmengen davon auf ihre Leistungsfähigkeit hin untersuchen zu können, wird analog zu den in Kapitel 2.3.3.3 beschriebenen Vorgehen eine Feature-Selektion auf Basis der *Pearson*-Korrelation mit anschließendem Ranking implementiert. Dazu wird zunächst für jede der im Rahmen der Studie erfassten $n = 152$ Routen ein Merkmalsvektor $F_n = (x_1^n, \dots, x_i^n)$ der Länge i anhand des in den Abschnitten 5.4.3 und 5.4.4 definierten Vorgehens berechnet. Anschließend wird jedem Merkmalsvektor die dazugehörige beobachtete Klasse in Form eines Labels l_y aus der Menge aller möglichen Beobachtungen $L = \{l_1, \dots, l_z\}$ der Anzahl z zugeordnet, woraus sich für alle F_n eine Instanz I_n ergibt:

$$I_n = \{F_n | l_y\} \quad (5.8)$$

Folgend wird im Kontext der korrelationsbasierten Feature-Analyse nun der Wert jedes Features des gegebenen Merkmalsvektors F_n über dessen Korrelation mit den zugeordneten Ereignislabels innerhalb der Instanzen I bestimmt. Der Feature-Wert resultiert aus dem gewichteten Durchschnitt der Gesamtkorrelation eines Attributes mit der F_n zugeordneten Klasse für alle Instanzen. Auf dieser Basis können alle Features eines Merkmalsvektors nun nach ihrer Wertigkeit sortiert und somit unterschiedliche Feature-Subsets erstellt werden.

Zur qualitativen, automatisierten Bestimmung der Schwierigkeit unterschiedlicher Routen im folgenden Abschnitt 5.5.4 werden die n Merkmalsvektoren F_n mit den Farbkategorien der Routenschwierigkeiten Blau (bl), Schweden (sch) und Grün (gr) gelabelt. Dabei ergibt sich über die korrelationsbasierte Feature-Analyse für diesen Fall ein Wertebereich von $0,00834 - 0,375$. Die gesamte Verteilung der gemittelten Gesamtkorrelation für alle Features und deren Häufigkeiten wird im Histogramm in Abbildung 5.9 dargestellt. Abbil-

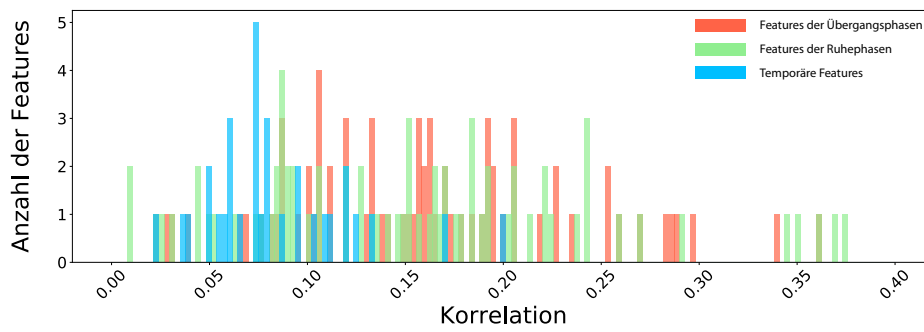


Abbildung 5.9: Histogramm der Korrelation von nach Routenschwierigkeit gelabelten Features mit farblicher Hervorhebung der zugehörigen Oberkategorie der Features.

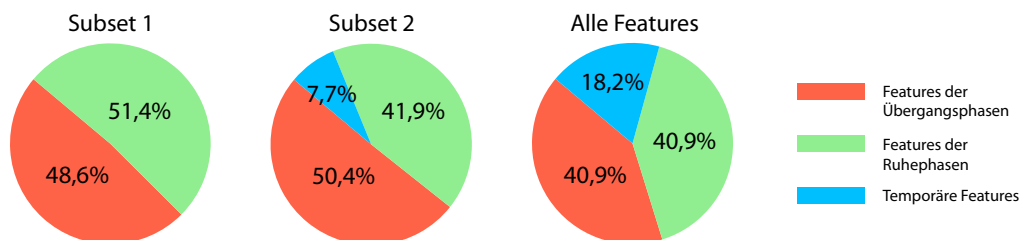


Abbildung 5.10: Prozentuale Verteilung der drei Oberkategorien der Features innerhalb der unterschiedlichen, mithilfe von Feature-Korrelation erstellten Feature-Sets zur Klassifizierung von Routen nach ihrer Schwierigkeit: *Subset 1*, *Subset 2* und *Alle Features*.

Abbildung 5.11 gibt in (1) zusätzlich noch einen differenzierteren Einblick in die Wertigkeit der einzelnen Features bezüglich der automatisierten Schwierigkeitsbestimmung einzelner Routen. Dabei ist zu erkennen, dass die Features der Übergangsphasen mit einer gemittelten Gesamtkorrelation k von meist größer 0,1 generell eine hohe Aussagekraft bei der Unterscheidung von Routenschwierigkeit liefern. Die Features der Ruhephasen hingegen verteilen sich über den gesamten Wertebereich und stellen zusätzlich die global aussagekräftigsten, aber auch weniger aussagekräftigen Kandidaten im Vergleich zu den Features der Übergangsphasen. Die temporalen Features mit häufigen Werten von $k < 0,1$ sind für die Unterscheidung von Schwierigkeit generell weniger relevant, dennoch liegen alle Werte immer noch deutlich über $k = 0$.

Mithilfe dieses Rankings nach Feature-Wertigkeit werden nun zusätzlich zum vorhandenen Merkmalsvektor, welcher alle 176 extrahierten Features enthält, zwei weitere Feature-Subsets für die spätere Klassifizierung erstellt. Diese ergeben sich aus der Definition zweier Korrelations-Schwellwerte. *Subset 1* beinhaltet dabei 37 Features mit jeweils einer gemittelten Gesamtkorrelation von mindestens $k > 0,2$, welche sich in 18 Features der Übergangsphasen und 19

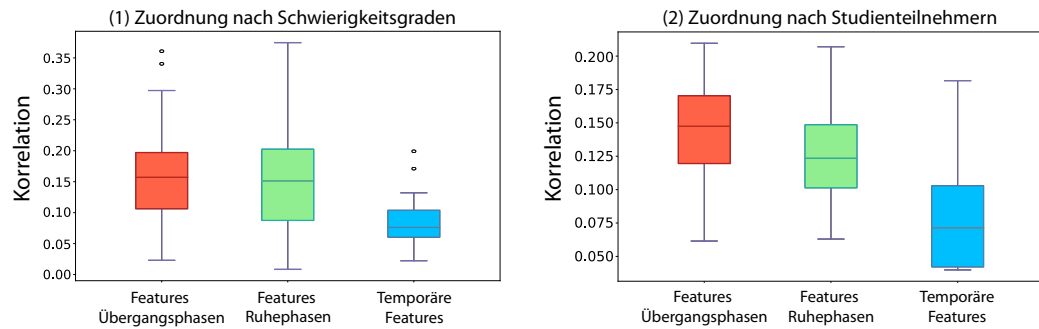


Abbildung 5.11: Verteilung der Features innerhalb ihrer Kategorien nach ihrer auf Korrelation basierten Wertigkeit für die zwei unterschiedlichen Zuordnungen auf Routenschwierigkeit (1) und auf unterschiedliche Studienteilnehmer (2).

Features der Ruhephasen unterteilen. *Subset 2* beinhaltet 117 Features mit jeweils einer gemittelten Gesamtkorrelation von mindestens $k > 0,1$, welche sich in 59 Features der Übergangsphasen, 49 Features der Ruhephasen und 9 temporale Features unterteilen. Abbildung 5.10 gibt hierzu einen Überblick über die Verteilung der Features in den einzelnen Feature-Sets.

Zur quantitativen Unterscheidung der einzelnen Studienteilnehmer, welche in Abschnitt 5.5.4 durchgeführt wird, werden die n Merkmalsvektoren F_n mit den Kürzeln der einzelnen Studienteilnehmer gelabelt. In diesem Fall ergibt sich über die korrelationsbasierte Feature-Analyse ein Wertebereich von $0,0399 - 0,2096$. Damit ist selbst das Feature von geringstem Wert deutlich expressiver für die Unterscheidung der Teilnehmer, als das geringste im Fall der Unterscheidung von Routenschwierigkeit, wobei aber nur 6 Features ein höheres Gewicht als $k > 0,2$ haben und keines $k > 0,3$ erreicht. Die Verteilung ist also wesentlich kompakter.

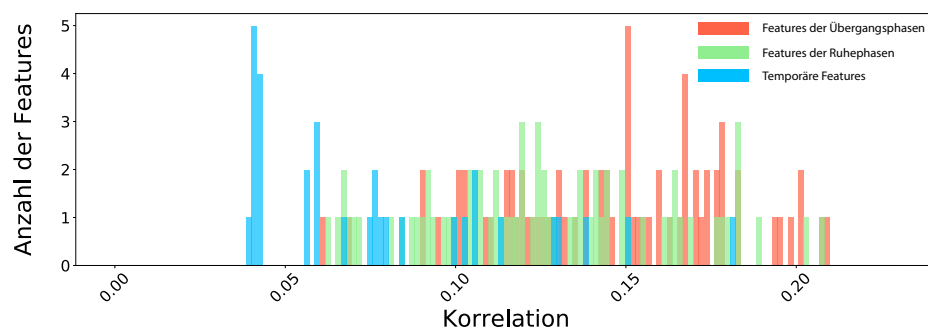


Abbildung 5.12: Histogramm der Korrelation von nach Studienteilnehmern gelabelten Features mit farblicher Hervorhebung der zugehörigen Feature-Kategorie.

Auch hier wird die gesamte Verteilung der gemittelten Gesamtkorrelation für alle Features wieder in einem Histogramm in Abbildung 5.12 dargestellt.

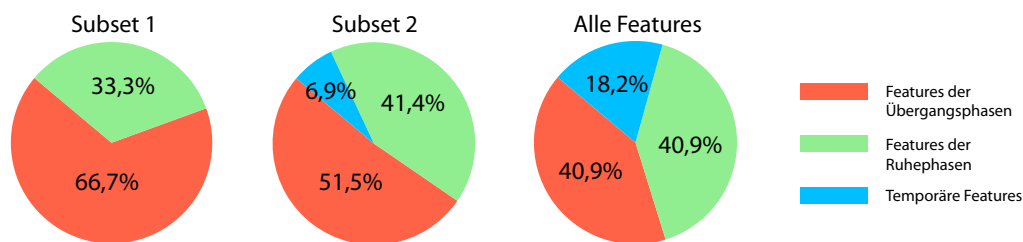


Abbildung 5.13: Prozentuale Verteilung der drei Feature-Kategorien innerhalb der unterschiedlichen, mithilfe von Feature-Korrelation erstellten Feature-Sets zur Klassifizierung nach Studienteilnehmern: *Subset 1*, *Subset 2* und *Alle Features*.

Zusätzlich gibt Abbildung 5.11 in (2) einen Einblick in die Wertigkeit der einzelnen Feature-Kategorien hinsichtlich des Anwendungsfalles Teilnehmerunterscheidung. Die Features der Übergangsphasen sind mit einer gemittelten Gesamtkorrelation von meist $k > 0,1$ und Spitzenwerten von über $0,2$ als sehr aussagekräftig zu bezeichnen. Zusätzlich stellt diese Feature-Kategorie die aussagekräftigsten Kandidaten. Die Features der Ruhephasen verteilen sich über den gesamten Wertebereich, bleiben im Großteil aber hinter den Features der Übergangsphasen zurück. Die temporalen Features stellen zwar erneut gehäuft die schlechtesten Kandidaten, sie sind im Vergleich zu Abbildung 5.10 aber deutlich ausgewogener verteilt und beinhalten sogar einige Kandidaten mit $k > 0,175$.

Anschließend werden auch für die spätere Teilnehmerunterscheidung zwei weitere Feature-Subsets mit den Schwellwerten $k > 0,2$ und $k > 0,1$ erstellt. *Subset 1* beinhaltet dabei 6 Features mit jeweils einer gemittelten Gesamtkorrelation von mindestens $k > 0,2$, welche sich in 2 Features der Übergangsphasen und 4 Features der Ruhephasen unterteilen. *Subset 2* beinhaltet 130 Features mit jeweils einer gemittelten Gesamtkorrelation von mindestens $k > 0,1$, welche sich in 67 Features der Übergangsphasen, 54 Features der Ruhephasen und 9 temporale Features unterteilen. Abbildung 5.10 gibt hierzu einen Überblick über die prozentuale Verteilung der Features innerhalb der einzelnen Feature-Sets für die folgenden quantitativen Untersuchungen.

Ausgangsinformationen des in diesem Abschnitt durchgeführten Schritts der Feature-Korrelation und Selektion sind also jeweils drei Feature-Sets für die quantitative Untersuchung und drei Feature-Sets für die qualitative Untersuchung: 1) jeweils ein Set mit allen Features, 2) jeweils ein Feature-Subset für $k > 0,1$ und 3) jeweils ein Feature-Subset für $k > 0,2$. Diese Feature-Sets werden nachfolgend als Eingangsinformation für die beiden durchgeführten Experimente zur automatisierten Erkennung von Routenschwierigkeit und zur Erkennung von Studienteilnehmern verwendet. Der Selektionsprozess zeigt, dass die in Abschnitt 5.4.4 gewählten Features nahezu durchgehend als aussagekräftig für beide Experimente gewertet werden können.

5.5.3 Automatisierte Bestimmung von Routenschwierigkeit

Nun werden die entwickelten Feature-Sets dazu genutzt, um die Schwierigkeit unterschiedlicher Kletterrouten automatisiert und nur auf Basis von Bewegungsinformation zu bestimmen. Dazu wurden verschiedene Vorhersagemodelle mithilfe der bereits aus Kapitel 4.5 bekannten Klassifikatoren (C4.5, NB, RF und SVM) sowie eines automatisiert via HPO identifizierten Klassifikators trainiert und anschließend validiert. Die während der Validierung verwendeten Feature-Sets entsprechen dabei denen der in Abschnitt 5.5.2 ermittelten zur Bestimmung der Routenschwierigkeit: *Subset 1* mit einer Feature-Anzahl von $n = 37$, *Subset 2* mit einer Feature-Anzahl von $n = 117$ und das komplette Set der Länge $n = 176$, welches alle Features beinhaltet.

In einem ersten Experiment wurde die automatisierte Erkennung der Routenschwierigkeit im Rahmen einer 10-fach gefalteten Kreuzvalidierung (10fCV) durchgeführt, in einem zweiten wurde die Klassifizierungs-Performanz unter Durchführung einer *Leave-One-Out-Cross-Validation* (LOOCV) getestet. Bei der HPO-basierten Wahl eines Klassifizierungsalgorithmus bekam die genutzte Maschine eine Rechenzeit von 15 Minuten, in welcher sie verschiedene Algorithmen mit unterschiedlichen Hyper-Parameter-Konfigurationen testen und einen optimierten Klassifikator ermitteln konnte. Da im vorliegenden Fall keine binären Klassifikatoren trainiert wurden, werden klassische Qualitätsmetriken binärer Entscheidungen wie *Sensitivity*, *Specificity*, *Precision* und *F-Measure* nachfolgend nicht betrachtet.

Feature-Set	C 4.5 (%)	NB (%)	RF (%)	SVM (%)	HPO (%)
<i>Subset 1</i>	71.053	80.921	82.895	80.263	100
<i>Subset 2</i>	66.447	77.631	84.867	82.237	100
<i>Alle Features</i>	63.816	72.368	84.868	86.842	100

Tabelle 5.5: Korrekte Erkennungsraten der unterschiedlichen Routenschwierigkeiten, aufgeschlüsselt nach genutztem Klassifikator im Rahmen der 10fCV.

In Tabelle 5.5 sind die Ergebnisse zur Klassifikation von Routenschwierigkeiten im Rahmen der 10fCV für die verschiedenen Klassifikatoren und die unterschiedlichen Feature-Sets aufgeführt. Auffällig ist, dass selbst unter Verwendung von *Subset 1*, mit nur 37 Features, die Schwierigkeit einer gekletterten Route im schlechtesten Fall zu 71,053% (C4.5), im besten Fall sogar mit bis zu 82,895% (RF) korrekt vorhergesagt werden kann. Für die Teilmengen von *Subset 1* und *Subset 2* erzielt der RF unter den manuellen Klassifikatoren jeweils die besten Ergebnisse, während mit steigender Feature-Anzahl die Ergebnisse der SVM besser werden und diese für die Gesamtheit aller Features schließlich die beste Erkennungsrate unter manueller Konfiguration liefert. Dieser Trend ist auch in Abbildung 5.15 (1) gut zu erkennen. Für die HPO-basierte

Klassifikation konnte in allen drei Fällen eine Erkennungsrate von 100% realisiert werden. In diesem Szenario wurde für *Subset 1* mit dem *Lazy Locally Weighted Learning*-Klassifikator (LWL) ein instanzbasierter Lernalgorithmus identifiziert [66], für *Subset 2* eine optimierte SVM und für ein Set mit allen Features ein optimierter *Multilayer Perceptron* MLP Klassifikator.

Die durchschnittliche Zeit zum Training der Vorhersagemodelle für alle drei Feature-Sets ist mit 0,033s für den C4.5, 0,003s für den NB, 0,08s für den RF und 0,017s für die SVM vergleichsweise gering. Für die im Rahmen der HPO ermittelten Klassifikatoren betrug die Modelltrainingszeit durchschnittlich 0,29s, allerdings muss nominell noch die zur Parameteroptimierung benötigte Zeit hinzugezählt werden.

Feature-Set	C 4.5 (%)	NB (%)	RF (%)	SVM (%)
<i>Subset 1</i>	66.447	79.605	78.947	79.605
<i>Subset 2</i>	62.05	76.316	81.58	82.24
<i>Alle Features</i>	61.184	68.421	83.553	84.868

Tabelle 5.6: Korrekte Erkennungsraten der unterschiedlichen Routenschwierigkeiten, aufgeschlüsselt nach genutztem Klassifikator im Rahmen der LOOCV

Tabelle 5.6 zeigt die Ergebnisse eines zweiten Experiments und gibt Aufschluss über die Performanz der manuellen Klassifikatoren im Rahmen der LOOCV. Zusammen mit der zusätzlichen Betrachtung von Abbildung 5.15 (1) ist auffällig, dass die Ergebnisse mit denen aus 5.5 korrelieren, im Schnitt aber immer um einige Prozentpunkte schlechter sind. Die besten Erkennungsraten mit Parameter-Konfiguration liefert auch hier die SVM mit 84,868%, ebenso steigt die Qualität der Ergebnisse mit zunehmender Anzahl der betrachteten Features.

In Abbildung 5.14 werden die Klassifizierungsergebnisse der 10fCV aufgeschlüsselt nach den Schwierigkeitsbereichen Blau (B), Schweden (S) und Grün (G), genutztem Klassifikator und genutztem Feature-Set mithilfe von Konfusionsmatrizen dargestellt. Vor allem für die weniger performanten Klassifikatoren ist deutlich zu erkennen, dass falsch erkannte Routenschwierigkeiten zumeist immer als in benachbarten Schwierigkeitsbereichen befindliche Grade klassifiziert wurden. Dies lässt sich durch die kontinuierliche Ausprägung einzelner Farben über jeweils einen Bereich von unterschiedlichen Schwierigkeitsgraden sowie die an den Grenzen teils überlappenden Schwierigkeiten erklären. Eine deutliche Ausnahme macht hier nur der NB unter Verwendung aller Features, wobei Grün oft als Blau klassifiziert und Schweden dabei übersprungen wird.

Zusammengefasst lässt sich festhalten, dass eine automatisierte Schwierigkeitsbestimmung von Kletterrouten rein auf Basis von menschlicher Bewegungsinformationen mit sehr guten Erfolgsraten möglich ist. Die besten Ergebnisse in manueller Konfiguration wurden unter Verwendung einer SVM und

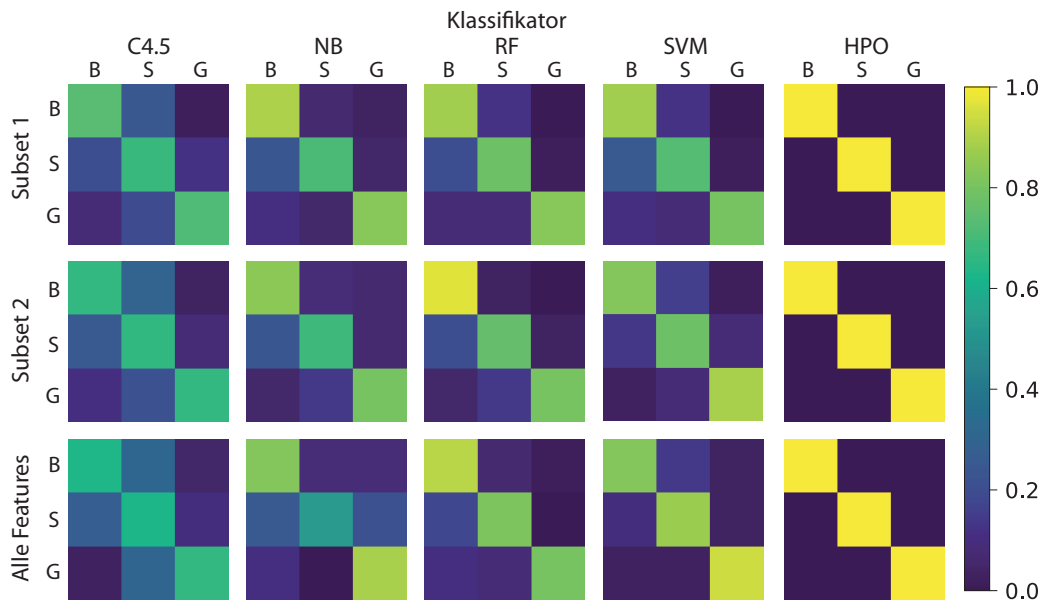


Abbildung 5.14: Klassifizierungsergebnisse des 10fCV-basierten Experiments, aufgeschlüsselt nach den gelabelten Schwierigkeitsbereichen Blau (B), Schweden (S) und Grün (G) und geordnet nach genutztem Klassifikator sowie nach genutztem Feature-Set.

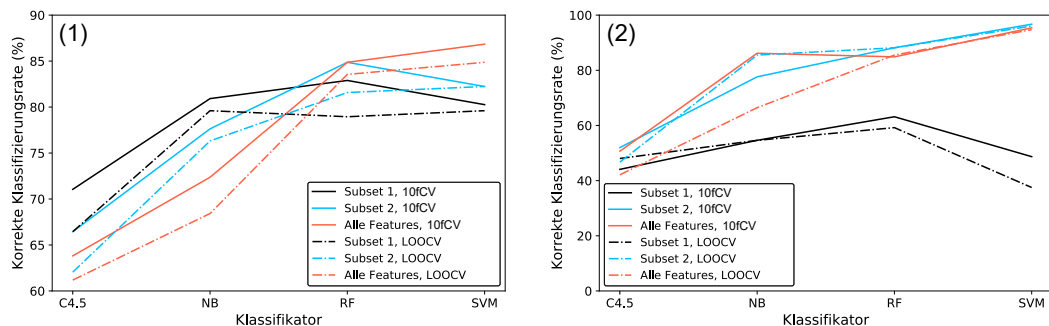


Abbildung 5.15: Durchschnittliche Erkennungsrate für die automatisierte Erkennung von Schwierigkeitsgraden (1) und für die Unterscheidung einzelner Studienteilnehmer (2), aufgeschlüsselt nach Feature-Sets und Validierungsverfahren.

allen 176 Features mit einer korrekten Erkennungsrate 86,842 % im Rahmen einer 10fCV erzielt – via HPO konnte sogar eine Erkennungsrate 100 % erreicht werden. Darüber hinaus zeigen die Ergebnisse auch, dass das in Abschnitt 5.4.4 entwickelte Modell zur differenzierten Beschreibung menschlicher Bewegungsaktivität sehr gut geeignet ist, diese expressiv zu modellieren und für überwachte maschinelle Lernprozesse zu repräsentieren.

5.5.4 Automatisierte Unterscheidung von Studienteilnehmern

In diesem Abschnitt wird eine quantitative Untersuchung durchgeführt, bei welcher die einzelnen Teilnehmer der Studie rein auf Basis von Bewegungsinformation voneinander unterschieden werden sollen. Die dazu genutzten Klassifikatoren sind wieder C4.5, NB, RF und SVM sowie ein automatisiert via HPO identifizierter und konfigurierter Klassifikator. Die drei Feature-Sets umfassen, wie in 5.5.2 festgelegt, $n = 6$ Features in *Subset 1*, $n = 130$ Features in *Subset 2* und im dritten Feature-Set alle der 176 extrahierten Features. Auch zur Erkennung der einzelnen Teilnehmer wurden mit einer 10fCV und einer LOOCV analog zu Abschnitt 5.5.3 zwei Experimente durchgeführt.

Feature-Set	C 4.5 (%)	NB (%)	RF (%)	SVM (%)	HPO (%)
<i>Subset 1</i>	44.08	54.61	63.16	48.684	78.29
<i>Subset 2</i>	51.973	77.632	88.158	96.711	100
<i>Alle Features</i>	50.658	86.184	84.868	95.395	100

Tabelle 5.7: Korrekte Erkennungsraten für die unterschiedlichen Studienteilnehmer, aufgeschlüsselt nach genutztem Klassifikator im 10fCV-Szenario.

Tabelle 5.7 zeigt die ermittelten Klassifizierungsergebnisse zur Teilnehmerunterscheidung im Rahmen der 10fCV. Auffällig ist hier, dass mithilfe eines RF-Entscheidungsbaumes bereits auf Basis von *Subset 1*, welches nur 6 Features enthält, eine Erkennungsrate von über 63 % realisiert werden kann. Dies unterstreicht zusätzlich den zuvor via Korrelation ermittelten hohen Informationsgehalt der Features in *Subset 1*. Im automatisierten HPO-Ansatz können die 16 Probanden auf Basis dieser 6 diskreten Zahlenwerte sogar mit einer Trefferquote von mehr als 78 % korrekt auseinander gehalten werden. Unter Verwendung von *Subset 2* erreichen die SVM und der HPO-Ansatz bereits 96,711 % bzw. 100 % an korrekt erkannten Teilnehmern über alle erfassten Routen. Dies sind gleichzeitig auch die besten erreichten Ergebnisse für diesen Anwendungsfall. Daraus lässt sich schließen, dass für die Unterscheidung von Teilnehmern, im Gegensatz zu der Erkennung von Routenschwierigkeit aus Abschnitt 5.5.3, nicht alle der 176 extrahierten Features notwendig sind und einige die Ergebnisse sogar verschlechtern.

Ähnliche Rückschlüsse lassen sich auch aus den Ergebnissen der LOOCV ziehen. Wie in Kombination mit Abbildung 5.15 (2) zu erkennen ist, orientieren sich diese zumeist wieder an den Ergebnissen der 10fCV. Eine klare Ergebnisverschlechterung mit steigender Feature-Anzahl ist für den NB-Klassifikator zwischen *Subset 2* und dem Set für *Alle Features* erkennbar. Auch hier erreichen wieder fast alle Klassifikatoren (NB, RF und SVM) ihre besten Ergebnisse unter Training und Klassifikation mit *Subset 2*, was in durchschnittlichen Erkennungsraten von 85,526 % bis 96,052 % resultiert. Nur der C4.5 erreicht für

Feature-Set	C 4.5 (%)	NB (%)	RF (%)	SVM (%)
<i>Subset 1</i>	48.026	54.605	59.210	37.5
<i>Subset 2</i>	46.710	85.526	88.158	96.052
<i>Alle Features</i>	42.105	66.447	85.526	94.737

Tabelle 5.8: Korrekte Erkennungsraten für die unterschiedlichen Studienteilnehmer, aufgeschlüsselt nach genutztem Klassifikator im LOOCV-Szenario.

das *Subset 1* die höchste durchschnittliche Erkennungsrate, was aber aufgrund der geringen Streuung der Ergebnisse und der generell schlechten Performanz des C4.5 nicht als aussagekräftig angesehen werden kann. Eine weitere interessante Beobachtung ist, dass die SVM bei geringer Feature-Anzahl in der LOOCV erneut schlechter abschneidet als die anderen Klassifikatoren, bei steigender Feature-Anzahl aber eine deutlich bessere Performanz erreicht.

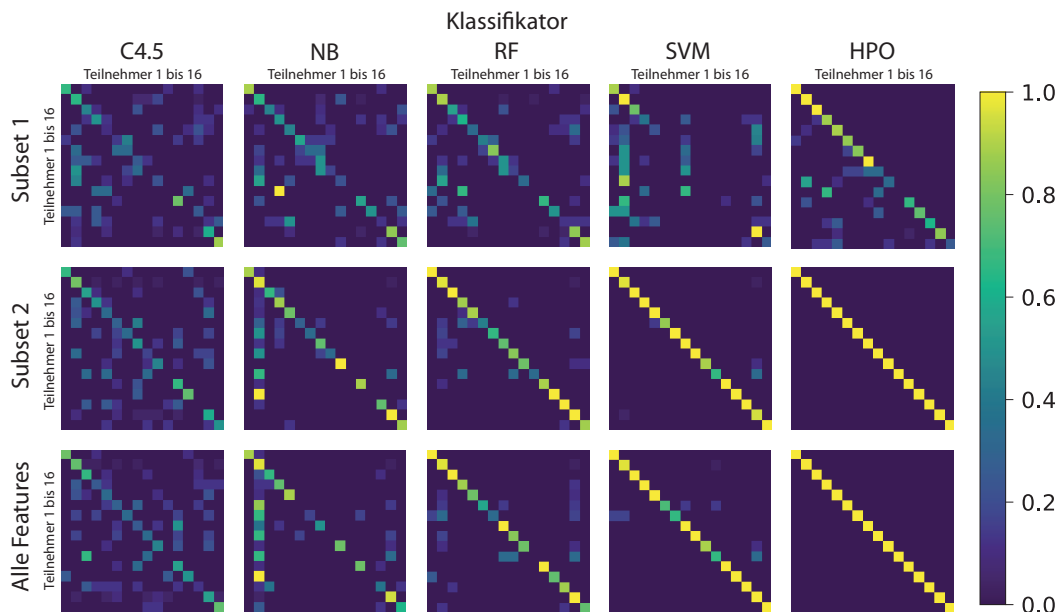


Abbildung 5.16: Konfusionsmatrizen der mithilfe von überwachtem Lernen voneinander unterschiedenen 16 Studienteilnehmer für alle Kombinationen aus Feature-Set und Klassifikatoren im manuellen Setup während der 10fCV.

Abbildung 5.16 zeigt die Konfusionsmatrizen für die Untersuchung der einzelnen Teilnehmer für alle Kombinationen aus Feature-Sets und Klassifikatoren im manuellen Setup während des 10fCV Experiments. In diesem Fall kann aus benachbarten Regionen keine Aussage abgeleitet werden, da zwischen einzelnen Teilnehmern keine direkte Abhängigkeit besteht. Besonders augenscheinlich ist aber, dass für alle Kombinationen, selbst solche mit *Subset 1*, welches nur 6 Features enthält, eine Diagonale von rechts oben nach links unten erkennbar

ist. Diese markiert korrekte Klassifikationen und ist umso deutlicher auszumachen, je besser die Klassifikationsergebnisse für eine Kombination sind. Eine interessante Beobachtung ergibt sich auch hier für den NB Klassifikator. Wie schon in den Tabellen 5.7 und 5.8 zu sehen, werden die Ergebnisse mit steigender Feature-Anzahl wider intuitiver Erwartung schlechter. Zusätzlich zeigt Abbildung 5.16, dass die Vorhersagen mit dem NB-Klassifikator für eine steigende Anzahl von Features immer mehr zu dem Studienteilnehmer an zweiter Position hin tendieren, was als Indiz für mögliches Overfitting während des Modelltrainings gewertet werden könnte.

Zusammenfassend lässt sich festhalten, dass im Rahmen der Experimente dieses Abschnittes bewiesen wurde, dass die zuvor entwickelten Feature-Sets auch einen hohen individuellen Informationsgehalt bezüglich der Unterscheidung einzelner Studienteilnehmer aufweisen. In einem manuellen Parameter-Setting konnten diese mit einer durchschnittlichen Rate von über 96% voneinander unterschieden werden. In automatisiert optimierten Szenarien gelang dies bei 16 Teilnehmern sogar mit 100% Erfolgsrate und nur auf Basis der aus ihren Bewegungsinformationen extrahierten Merkmale.

5.6 Diskussion und Zusammenfassung

In diesem Kapitel wurde gezeigt, dass sowohl quantitative als auch qualitative Untersuchungen von nicht vorhersehbarer Bewegungsaktivität allein auf Basis von Bewegungsinformation mit großem Erfolg zu realisieren sind. Dazu wurde zunächst eine Prozesskette zur Verarbeitung und Untersuchung auf Basis des SensX Sensorsystems definiert. Das anschließend entwickelte Untersuchungskonzept besteht aus einem zweistufigen Segmentierungsprozess und realisiert die Projektion von roher Sensorinformation auf qualitative Eigenschaften physischer menschlicher Leistung. Resultat ist ein diskreter Merkmalsvektor, welcher zur Untersuchung von Bewegungsaktivität mithilfe von Werkzeugen des maschinellen Lernens geeignet ist. Zur Realisierung dieser Projektion wurde ein differenziertes Modell zur Beschreibung menschlicher Bewegungsaktivität, welches auf drei aggregierten Kernfähigkeiten von physischer, menschlicher Leistung basiert, definiert.

So konnten alle 16 Teilnehmer der Evaluationsstudie aus dem Anwendungsszenario Klettern mit einer Erfolgsrate von 100% im automatisiert optimierten Ansatz und mit mehr als 96% im manuellen Ansatz, lediglich auf Basis der 176 Features, voneinander unterschieden werden. Weiterhin konnte auf Basis des gleichen Feature-Sets mit einer Erfolgsrate von 100% (HPO-Szenario) bzw. mehr als 86% (manuelle Konfiguration) die zugrundeliegende Schwierigkeit verschiedener Boulder Routen erkannt werden. Damit ist bewiesen, dass automatisierte, objektive Bewertung und Untersuchung von nicht vorhersehbarer menschlicher Bewegungsaktivität, rein auf Basis von Bewegungsinformation, möglich ist. Aufgrund der Komprimierung der genutzten Feature-Sets, in Verbindung mit den Erkenntnissen aus Kapitel 4.5.4 zur Systemlaufzeit, ist davon

ausgehen, dass das Konzept echtzeitfähig ist und auch neue Vorhersagemodelle innerhalb weniger Sekunden trainiert und genutzt werden können.

Dabei spiegelt sich in den genutzten Feature-Sets quantitativ wieder, dass die automatisierte Bestimmung von Schwierigkeiten ungleich komplexer ist als die Unterscheidung von Teilnehmern. Trotz der geringen Anzahl von drei möglichen Schwierigkeitsbereichen ist zu deren verlässlicher Unterscheidung eine deutliche größere Feature-Anzahl notwendig als zur Unterscheidung einzelner Teilnehmer. Generell hat sich aber gezeigt, dass nahezu alle der in Abschnitt 5.4.4 identifizierten Features für den untersuchten Anwendungsfall eine sehr hohe Wertigkeit besitzen. So ist es möglich, die einzelnen Studienteilnehmer auf Basis von nur 6 diskreten Parametern mit einer korrekten Erkennungsrate von über 78% zu identifizieren. Auch hat sich gezeigt, dass das entwickelte Feature-Set flexibel genug ist, um sowohl zur Teilnehmerunterscheidung als auch zur Schwierigkeitsbewertung von Boulder Routen herangezogen zu werden. Die anfänglich geforderte Balance zwischen Spezifitäts- und Generalisierungseigenschaften ist also gegeben.

Aus der Nutzung von Boulder Routen zur Evaluation des vorgestellten Konzepts ergeben sich einige Folgefragen. Bouldern ist eine Unterart des Kletterns und sowohl Bewegungsabfolgen als auch die Art der Griffe und auftretende Kletterprobleme sind meist übereinstimmend. Es ist denkbar, dass aufgrund der intensiveren Krafternutzung, kürzerer Routenlängen und der generell mehr genutzten Schnellkraft beim Bouldern Features der Übergangsphasen im Rahmen der vorherigen Untersuchungen stärker gewichtet wurden, als dies in einem klassischen Kletterszenario der Fall gewesen wäre [60]. In diesen, eher auf Ausdauer und konstantem Kraftaufwand aufbauenden Kletterproblemen, könnten die temporalen Features gegenüber denen der Übergangs- und Ruhephasen an Relevanz gewinnen, was Gegenstand weiterer Untersuchungen sein kann. Weiterhin ist darauf hinzuweisen, dass es sich bei den erfassten Studienteilnehmern unter Aspekten des Kletterkönnens um eine relativ homogene Gruppe handelte. Die entwickelten Konzepte stützen sich darauf, dass Teilnehmer während des Kletterns vergleichbare Anzeichen von Erschöpfung oder Ausdauerverlust zeigen, um anschließend aus diesen Beobachtungen auf einzelne Teilnehmer oder Schwierigkeitsgrade zu schließen. Daraus folgt, dass für kleine zu untersuchende Gruppen das Level des Kletterkönnens nicht zu stark divergieren darf, da sonst gegebenenfalls keine miteinander in Verbindung stehenden Beobachtungen zwischen unterschiedlichen Teilnehmern gemacht werden können. Ein letzter zu nennender, kletterspezifischer Aspekt ist die Auswahl der aufgenommenen Boulder Routen. Diese hatten alle einen Schwierigkeitsgrad von einfach bis mittel und beinhalteten kaum einschlägige Schikanen wie Leisten oder Überhänge. Um auch diese mithilfe des entwickelten Feature-Set zu beschreiben, müsste dieses um entsprechende Parameter, wie beispielsweise die Orientierung der Körperachse, erweitert werden.

Eine der wichtigsten Fragestellungen ist, ob das entwickelte Modell zur differenzierten Beschreibung von Bewegungsaktivität, welches auf den aggregier-

ten Kernfähigkeiten Kraft, Ausdauer und Kontrolle beruht, auch auf andere menschliche Bewegungsaktivitäten projizierbar ist. Intuitive Grundannahme ist dabei, dass jegliche menschliche Bewegung, ob wiederkehrend, nicht vorhersehbar, periodisch oder nicht periodisch, von diesen drei Kernfähigkeiten maßgeblich bestimmt wird und somit von einem daraus entwickelten Feature-Set auch beschreibbar sein muss. Die Generalisierung und Erweiterung dieses Modells hin zu einem ganzheitlichen Konzept zur Beschreibung menschlicher Bewegungsaktivität ist somit eine vielversprechende, zukünftige Herausforderung.

6 Zusammenfassung und Ausblick

In der vorliegenden Arbeit standen die unabhängige Erfassung, die Erkennung und die qualitative Untersuchung von menschlicher Bewegung im Mittelpunkt. Während vergleichbare Arbeiten zur menschlichen Bewegungsanalyse zumeist durch rein quantitative Analyseaspekte beschränkt werden, wurde im Rahmen dieser Arbeit nicht nur eine Grundlage zur ganzheitlichen Erfassung von menschlicher Bewegung spezifiziert, es wurden auch Arten von menschlicher Bewegung, nämlich die wiederkehrende und nicht vorhersehbare definiert und darauf aufbauend dedizierte Konzepte zu deren fundierter Untersuchung entwickelt.

Um diese Herausforderungen zu adressieren und dazu neue Lösungswege aufzuzeigen, wurden einzelne Aspekte der Thematik in drei aufeinander aufbauenden Teilkapiteln differenziert untersucht.

Zu Beginn stand dabei in Kapitel 3 die Konzeption, Umsetzung und Evaluation eines verteilten Sensorsystems, welches auf Basis von Bewegungsinformation menschliche Aktivität erfassen und erkennen kann im Vordergrund. Ergebnis dieses Kapitels ist das SensX Sensorsystem, welches menschliche Bewegung mobil und unabhängig von komplexer Infrastruktur ganzheitlich erfassen sowie in Echtzeit analysieren kann. Im Vergleich zu verwandten Arbeiten standen dabei die Bewegungen von allen Extremitäten des menschlichen Körpers im Fokus der Untersuchung. Das System ist einfach und robust auf herkömmlichen Hardware-Komponenten aufgebaut und kann autark von einem einzigen Nutzer auch bei widrigen äußeren Bedingungen, wie beispielsweise während fordernder sportlicher Aktivität, benutzt werden. Die Architektur des Systems umfasst die Hardware-Ebene, welche die Erfassung von menschlicher Bewegung realisiert und die logische Ebene, welche Konzepte zur Vorverarbeitung von Rohinformation, zur Segmentierung, zum Feature-Engineering und zur Untersuchung von Bewegung mithilfe von maschinellen Lernalgorithmen implementiert. Die Weiterentwicklung der logischen Untersuchungskonzepte in den Anschlusskapiteln 3 und 4 zeigt, dass die Grundarchitektur modular aufgebaut ist und eine einfache Rekonfiguration für unterschiedliche Anwendungsszenarien erlaubt. Im Rahmen der ersten Evaluation von SensX wurde mit einer umfangreichen Studie gezeigt, dass einzelne Ereignisse komplexer menschlicher Bewegungsaktivitäten bereits auf Basis von statischen Segmentierungsverfahren und im Zusammenspiel mit interpolierten Merkmalsvektoren mit einer Erfolgsrate von über 99% in Echtzeit erkannt werden können.

Weiterhin ist die Rekonfiguration des Systems und das Erlernen neuer Ereignisklassen innerhalb von Sekunden möglich. Damit wurde mit SensX eine potente Grundlage zur Erfassung und Erkennung menschlicher Bewegung geschaffen, welche in den anschließenden Inhaltskapiteln weiter optimiert wurde und anschließend als Basis zur fundierten Untersuchung qualitativer Aspekte von Bewegungsaktivität nutzbar ist.

In Kapitel 4 stand die automatisierte, qualitative Untersuchung von individuellen, wiederkehrenden, menschlichen Bewegungsereignissen im Vordergrund, wobei daraus auch optimierte Konzepte zur rein quantitativen Erkennung von komplexen Bewegungsereignissen hervorgingen. Dazu wurde zunächst ein neuer, adaptiver Ansatz zur Extraktion von individuellen, wiederkehrenden Bewegungsereignissen mit variabler Länge aus menschlicher Bewegungsaktivität vorgestellt. Um auch qualitative Untersuchungen anzustellen, wurden Vorgehen zur Diskretisierung und Formalisierung von qualitativen Beobachtungen innerhalb der extrahierten Segmente definiert, um diese anschließend mithilfe von Werkzeugen des maschinellen Lernens zu untersuchen. Dabei wurde ein stark komprimierter Merkmalsvektor entwickelt, welcher die qualitative Untersuchung von menschlicher Bewegung ermöglicht und darüber hinaus auch signifikante Verbesserungen hinsichtlich deren Erkennung realisiert. In der anschließenden Evaluation wurde die Leistungsfähigkeit dieser Konzepte demonstriert. Unterschiedliche wiederkehrende und komplexe, menschliche Bewegungsereignisse konnten mit einer Erfolgsrate von ebenfalls mehr als 99% in ihrer Ausführungsqualität korrekt eingeordnet werden, ohne dabei durch die spezifische Definition von bestimmten Fehlerklassen in der Untersuchung eingeschränkt zu sein. Die Erkennungsrate von individuellen Ereignissen konnte im direkten Vergleich zum Vorkapitel sogar auf 100% gesteigert werden. Weiterhin wurde auch gezeigt, dass das adaptive Segmentierungsverfahren im Vergleich zum statischen Vorgehen einen spürbaren Genauigkeitsvorteil im Rahmen einer Klassifikator-basierten Qualitätsuntersuchung erzielt. Durch die hohe Kompression des entwickelten Merkmalsvektors konnten so die Laufzeiten während des Online-Klassifikationsprozesses und während des Erlernens von neuen Bewegungsereignissen signifikant verkürzt werden. Zusätzlich wurde gezeigt, dass bereits eine verhältnismäßig geringe Anzahl von Ereignisinstanzen ausreichend ist, um Ereignisse auf Basis dieses Merkmalsvektors mit sehr hoher Genauigkeit zu untersuchen. Dies unterstreicht zusätzlich die Eignung des Systems zum Erlernen und Untersuchen von unbekanntem Bewegungsereignissen in Echtzeit.

Im Fokus des letzten Teilkapitels stand anschließend die automatisierte Untersuchung von menschlicher Bewegungsaktivität, welche sich aus nicht vorhersehbaren Bewegungsereignissen zusammensetzt. Schwerpunkt der Untersuchungen war dabei das Anwendungsszenario Klettern. Dazu wurden die mit SensX erfassten Bewegungsinformationen zunächst mit einem zweistufigen Segmentierungsprozess und individuell für die menschlichen Gliedmaßen in separate Sinnabschnitte unterteilt. Diese Sinnabschnitte, namentlich Ruhe- und Aktivitätsphasen, wurden über ein differenziertes Modell zur Beschreibung

der physischen Komponenten von menschlicher Bewegungsaktivität, welches im Rahmen dieser Arbeit entwickelt wurde, auf drei hier definierte Kernkomponenten zur Beschreibung von körperlicher menschlicher Leistung projiziert. Über diese Projektion ließ sich schließlich ein generischer Merkmalsvektor ableiten, welcher sowohl qualitative als auch quantitative Untersuchungen von nicht vorhersehbarer, menschlicher Bewegung zulässt. Dessen Fähigkeiten wurden anschließend anhand von zwei unterschiedlichen Experimenten im Rahmen der Evaluation demonstriert. So war es möglich, mit einer Erfolgsrate von über 86% (via manueller Parameterkonfiguration) bzw. 100% (via HPO) die Schwierigkeit unterschiedlicher Boulder Routen automatisiert und nur auf Basis der Bewegungsinformationen der Studienteilnehmer zu bestimmen. Zusätzlich konnten im Rahmen des zweiten, quantitativ gelagerten Experiments alle Teilnehmer mit hundertprozentiger Sicherheit voneinander unterschieden werden.

Zusammenfassend ist also zu vermerken, dass mit den Teilaspekten dieser Arbeit ein ganzheitliches Paket zur sowohl qualitativen als auch quantitativen menschlichen Bewegungsanalyse entwickelt, validiert und vorgestellt wurde. Durch die alleinige Nutzung von Bewegungssensorik und den Verzicht auf stationäre Infrastruktur können nun auch Bewegungsaktivitäten, welche zuvor aufgrund von in anderen Systemen bestehenden Einschränkungen nicht untersucht werden konnten, eingehend analysiert werden. Dabei sind die in dieser Arbeit vorgestellten Entwicklungen nicht als Ersatz für bestehende Konzepte zur Aktivitätserkennung, beispielsweise auf Basis von visueller Information oder mithilfe von Tiefensensorik, zu sehen. Stattdessen wirken sie ergänzend und adressieren Anwendungsszenarien, in welchen die menschliche Bewegungsanalyse aufgrund verschiedener Einschränkungen, wie zum Beispiel fehlender Mobilität oder nur partieller Erfassung des menschlichen Körpers, zuvor nicht möglich war. Weiterhin basiert das entwickelte Konzept durchgehend auf handelsüblichen und einfach erhältlichen Grundkomponenten sowie auf offenen Software-Werkzeugen, was dessen Wert bezüglich Verbreitung und Anwendbarkeit sowohl für Endkunden im IoT-Bereich, als auch im professionellen Bereich zur Analyse und zum Training von sportlicher Aktivität oder im medizinischen Bereich zur Bewegungsanalyse erheblich steigert.

Abschließend sollen nun noch einige offene Fragestellungen aufgezeigt werden, welche sich im Laufe dieser Arbeit ergeben haben und welche als Ansatzpunkte für Folgearbeiten dienen können.

Ein offener Punkt hinsichtlich der qualitativen Untersuchung von wiederkehrender Bewegungsaktivität in Kapitel 4 ist die Bereitstellung von direktem Feedback hinsichtlich konkreter Fehlstellungen in Bewegungsabläufen. Im aktuellen Stand kann das System eine generische Aussage für ein Bewegungsergebnis bereitstellen, etwa ob eine Übung in ihrer Durchführung von guter oder von schlechter Qualität ist. Ein denkbarer nächster Schritt wäre nun die dahingehende Erweiterung des Konzepts, dass auch konkrete Fehlstellungen erkannt und benannt werden können. Wichtig ist, dass diese Fehler nicht statisch de-

finiert sein dürfen, wie es bislang in einigen verwandten Arbeiten der Fall ist. Ein denkbarer Lösungsansatz ist, menschliche Bewegung auf Basis von Beschleunigungs- und Rotationsinformation, ähnlich wie bei der Anwendung von Tiefensensorik, auf ein Modell des menschlichen Skeletts zu projizieren und anhand dessen Bewegungsabläufe zu rekonstruieren. Damit ergäbe sich die vielversprechende Möglichkeit, auch geometrische Merkmale in einen Analyseprozess miteinzubeziehen und so Fehlstellungen von Extremitäten zueinander zu berechnen. In diesem Kontext könnte SensX um weitere Sensoren an den menschlichen Extremitäten erweitert werden, um so beispielsweise auch deren Beuge- und Streckbewegungen in den Untersuchungsprozess zu integrieren.

Im Rahmen der Untersuchung von nicht vorhersehbarer Bewegungsaktivität anhand von Kletterbewegung in Kapitel 5 ergeben sich ebenso ungelöste Fragestellungen. Eine davon ist die Überprüfung des Konzepts hinsichtlich seiner Fähigkeit zur Generalisierung innerhalb des Anwendungsszenarios. Wie bereits in Abschnitt 5.6 erwähnt, können sich Kletterrouten und Boulder Routen bezüglich der zum Aufstieg nötigen Kraft und der notwendigen Ausdauer unterscheiden. Es gilt also zu untersuchen, ob das Konzept auch für die Analyse von Routen mit stark unterschiedlichen Anforderungen hinsichtlich Kraft und Ausdauer nutzbar ist. Weiterhin könnte das Konzept auch für generell anspruchsvollere Routen (Überhänge, Leisten, etc.) sowie heterogene Teilnehmergruppen erweitert werden. Intuitiv ist zu erwarten, dass das genutzte Modell zur Beschreibung von menschlicher Leistung zur Lösung dieser Anforderungen einer erneuten Überprüfung unterzogen werden muss, um auch bisher nicht betrachtete Eigenschaften von Routen und Kletterern zu adressieren.

Die wohl wichtigste Fragestellung ist aber, ob das entwickelte Modell zur differenzierten Beschreibung von Bewegungsaktivität, welches auf den aggregierten Kernfähigkeiten Kraft, Ausdauer und Kontrolle beruht, auch generisch auf andere menschliche Bewegungsaktivitäten projizierbar ist. Aktuell beinhaltet es zusätzlich zu rein auf Signaldynamik aufbauenden Merkmalen auch temporale Informationen über zugrundeliegende Bewegungsereignisse. Es ist denkbar, dass es mithilfe entsprechender Erweiterungen weiter generalisiert und für die Untersuchung unterschiedlichster Anwendungsszenarien aus dem Bereich menschlicher Bewegungsaktivität genutzt werden kann.

Diese Fragestellungen sind Anknüpfungspunkte für Folgearbeiten und somit die Grundlage weiterer Forschung auf dem Gebiet der menschlichen Bewegungsanalyse unter Zuhilfenahme von Bewegungssensorik.

Abkürzungsverzeichnis

CNN	<i>Convolutional Neural Network</i> Ein künstliches neuronales Netz mit einem von biologischen Prozessen inspirierten Konzept im Bereich des maschinellen Lernens. CNNs finden Anwendung in zahlreichen modernen Technologien der künstlichen Intelligenz, vornehmlich bei der maschinellen Verarbeitung von Bild- oder Audiodaten.
DAV	<i>Deutscher Alpenverein e.V.</i> , hierbei handelt es sich um die größte nationale Bergsteigervereinigung der Welt sowie den fünftgrößten nationalen Sportfachverband Deutschlands.
DTW	<i>Dynamic Time Warping</i> , auch dynamische Zeitnormierung genannt, bezeichnet einen Algorithmus, der Wertefolgen unterschiedlicher Länge aufeinander abbildet.
BLE	<i>Bluetooth Low Energie</i> ist eine Funktechnik, mit der sich Geräte in einer Umgebung von etwa 10 Metern vernetzen lassen.
FFT	<i>Fast Fourier Transformation</i> Die FFT ist ein Algorithmus zur effizienten Berechnung der diskreten Fourier-Transformation. Mit ihr kann ein digitales Signal in seine Frequenzanteile zerlegt und diese dann analysiert werden.
GPS	Das <i>Global Positioning System</i> ist ein Satelliten-gestütztes System zur Positionsbestimmung und Navigation auf der Erdoberfläche.
HPO	<i>Hyper Parameter Optimization</i> , ein oft automatisiertes Verfahren zur Determination optimaler Eingangsparameter für einen Algorithmus oder einen Klassifikator.
HMM	<i>Hidden Markov Models</i> Das HMM ist ein statistisches Markov-Modell, in welchem angenommen wird, dass das modellierte System ein Markov-Prozess mit unbeobachteten Zuständen ist.
IoT	<i>Internet of Things</i> Hierbei handelt es sich um einen Sammelbegriff für Technologien einer globalen Infrastruktur der Informationsgesellschaften, die es ermöglicht physische und virtuelle Gegenstände miteinander zu vernetzen und sie durch Informations- und Kommunikationstechniken zusammenarbeiten zu lassen.
kNN	<i>k Nearest Neighbour</i> , ein Klassifikator aus dem Bereich des Machine-Learnings.

- K*** Der K^* , auch *K-Star*, ist ein Klassifikator aus dem Bereich des maschinellen Lernens.
- GWR** *Growing when Required*, ein selbstorganisierender, Online-fähiger Clustering-Algorithmus aus dem Umfeld neuronaler Netze.
- LOOCV** *Leave One Out Cross Validation*, ein Validierungsverfahren im Kontext des maschinellen Lernens.
- LWL** *Locally-Weighted-Learning*, ein instanzbasierter Algorithmus zum maschinellen Lernen und zur Regression.
- MAC** Bei der MAC bzw. *Media Access Control*-Adresse handelt es sich um die Hardware-Adresse jedes einzelnen Netzwerkadapters. Sie dient als eindeutiger Identifikator des Geräts in einem Rechnernetz.
- MLP** *Multi Layer Perceptron*, ein Klassifikator aus dem Bereich des Machine-Learnings in Form eines neuronalen Netzes.
- MMS** *Most Meaningful Signal*, im Kontext dieser Arbeit bezeichnet das MMS das Signal innerhalb eines, das gleiche Bewegungsereignis beschreibende Signalsets, welches die höchste Signaldynamik aufweist und somit zur Erkennung einzelner Bewegungsereignisse am besten geeignet ist.
- NB** *Naive Bayes*, ein Klassifikator aus dem Bereich des Machine-Learnings.
- PCA** *Principal Component Analysis*, diese ist ein Verfahren der multivariaten Statistik. Sie dient dazu, umfangreiche Datensätze zu strukturieren, zu vereinfachen und zu veranschaulichen, indem eine Vielzahl statistischer Variablen durch eine geringere Zahl möglichst aussagekräftiger Linearkombinationen genähert wird.
- RF** *Random Forest*, ein Klassifikator aus dem Bereich des Machine-Learnings.
- RFID** *Radio Frequency Identification*, Technologie für Sender-Empfänger-Systeme zum automatischen und berührungslosen Identifizieren und Lokalisieren von Objekten.
- SMA** *Signal Magnitude Area* In der Mathematik ist die SMA, der sog. Signalbetragsbereich, ein statistisches Maß für die Stärke einer variierenden Menge von Werten.
- SoP** *Sum of Pair-Distanz*
Die SoP ermöglicht den Vergleich zweier Zeitreihen gleicher Länge über die Bildung der Summe der euklidischen Distanzen aller sequentiell gegenüberstehenden Punkte-Paare.

-
- SVM** *Support Vector Machine*, die SVM dient im maschinellen als Klassifikator, in der Regressionsanalyse kommt sie auch als Regressor zum Einsatz.
- SURF** *Speeded Up Robust Features*, ein Algorithmus zur schnellen und robusten Erkennung von Bildmerkmalen im maschinellen Sehen.
- TS/TS** *Training Set / Test Set*, ein Validierungsverfahren im Kontext des maschinellen Lernens, wobei ein Datensatz vor der anschließenden Klassifizierung nach einem zuvor festgelegten Schlüssel in Trainings- und Testdaten unterteilt wird.
- 10fCV** *10-fach gefaltete Kreuzvalidierung* (engl. 10-fold Cross-Validation), ein Validierungsverfahren im Kontext des maschinellen Lernens.

Abbildungsverzeichnis

2.1	Beschleunigungsinformationen eines sich wiederholenden Bewegungsereignisses über die Zeit t in Millisekunden. Links sind die Rohdaten dargestellt, rechts sind diese nach Filterung mit einem Butterworth-Tiefpassfilter der 2. Ordnung zu sehen.	11
2.2	Ungefilterte Beschleunigung des Brustsensors eines Athleten in Y -Richtung (auf- und abwärts) (1) während der Durchführung eines Sets mit 20 Wiederholungen der Körperereigengewichtsübung <i>Hampelmann</i> . In (2) lässt die Autokorrelation eines Ausschnitts des Signals aus (1) Rückschlüsse auf enthaltene Periodizitäten und die Periodendauer Δt zu.	14
2.3	Überblick über eine generische Prozesskette zur Aktivitätserkennung mithilfe von Prozessen des maschinellen Lernens. . . .	35
2.4	Schematische Darstellung von Schritt (1) <i>Erfassung von Bewegungsinformation</i> innerhalb der Prozesskette aus Abbildung 2.3. Die Eingangsinformationen setzen sich aus n Eingangssignalen zusammen.	35
2.5	Überführung der im ersten Schritt erfassten Bewegungsinformation in einen im Rahmen von Schritt (2) <i>Vorverarbeitung</i> geglätteten und zur weiteren Untersuchung geeigneten Zustand.	36
2.6	Schritt (3) <i>Segmentierung</i> befasst sich mit der Extraktion von individuellen Ereignissegmenten e_x aus dem im zweiten Schritt vorverarbeiteten Set S' von Zeitreihen.	38
2.7	Schritt (4) <i>Feature-Engineering</i> verdeutlicht die Überführung der im dritten Schritt extrahierten Einzelsegmente e_i in jeweils einen korrespondierenden Merkmalsvektor X_i	39
2.8	Untersuchung der im vierten Schritt kreierten Merkmalsvektoren X_i im Rahmen des letzten Schrittes (5) <i>Klassifizierung</i> mithilfe eines Prozesses des überwachten, maschinellen Lernens. . . .	40
3.1	Übersicht über das Architekturkonzept von SensX, unterteilt in die Hardware-Ebene und die logische Ebene : Die Hardware-Ebene umfasst den technischen Aufbau, die logische Ebene die abstrahierten Schritte 1. Erfassung, 2. Erkennung und 3. qualitative Analyse von menschlicher Bewegung.	52

3.2	SensX getragen von einem Probanden: die vier externen Sensoren sind an den einzelnen Extremitäten befestigt, während das Smartphone als zentrale Recheneinheit und Sensor zugleich an der Brust getragen wird.	55
3.3	Abbildung der externen Sensorplattformen mit Schutzhülle und Gummiband zur Befestigung. Zur autarken Energieversorgung dienen handelsübliche CR 2032 Knopfbatterien.	56
3.4	Zuordnung einzelner Schritte der Prozesskette des maschinellen Lernens aus Abbildung 2.3 zur Hardware-Ebene und zur logischen Ebene von SensX.	59
3.5	Sequentielles, kontinuierliches Suchen von Bewegungsereignissen in multidimensionalen Zeitreihen anhand von Schwellwerten und Extraktion mit statischen Fenstergrößen nach Ebert et al. [49, 50].	60
3.6	Resampling des am Bein eines Probanden in X-Richtung erfassten Beschleunigungssignals bei der Durchführung einer Wiederholung der Körpereigengewichtsübung <i>Crunches</i> entlang der X-Achse(vgl. 3.6.1).	63
3.7	Übersicht über die acht im Rahmen der Studie durchgeführten und erfassten Körpereigengewichtsübungen, in Reihenfolge ihrer Durchführung und mit zugehörigen Abkürzungen.	65
3.8	Visualisierung der durchschnittlichen korrekten Erkennungsrate (in Prozent) der einzelnen Körpereigengewichtsübungen mit SensX nach unterschiedlichen Sensorkonfigurationen (vgl. Tabelle 3.3)	68
3.9	Trainingsdauer von Vorhersagemodellen mithilfe des NB-Klassifikators im mobilen und im stationären Szenario.	70
4.1	Beschleunigung des Brustsensors eines Probanden in Z-Richtung für zwei nacheinander durchgeführte Sets der Körpereigengewichtsübung <i>Crunches</i> mit jeweils 20 Wiederholungen.	77
4.2	Schematische Darstellung der erfassten Signale des SensX-Systems und der darin beinhalteten sequentiellen Ereignisse mit ihrer individuellen und variablen Laufzeit.	83
4.3	Einzelschritte des Vorgehens zur adaptiven Segmentierung multidimensionaler, wiederkehrender Zeitreihen anhand von Extrema-Fingerprints.	83
4.4	Autokorreliertes Signal des SensX-Brustsensors in Y-Richtung, erfasst während der Durchführung einer Abfolge der Körpereigengewichtsübung <i>Hampelmann</i>	84
4.5	Bestimmung des optimalen Bereichs des Cutoff-Faktors Δcf innerhalb einer Sequenz mit 20 bekannten Bewegungsereignissen. Der optimale Bereich Δcf ist grau hervorgehoben.	86

4.6	Beispiele von Extrema-Fingerprints verschiedener Bewegungsereignisse mit unterschiedlich starker Konfiguration eines Butterworth-Tiefpassfilters. Oben ist jeweils das schwach gefilterte Eingangssignal zu sehen, unten das stark gefilterte Signal zur Identifikation der Extrema-Fingerprints.	87
4.7	Standardabweichung der Beschleunigungsinformation von jeweils 100 Bewegungsereignissen der Körpereigengewichtsübung <i>Bicycle Crunch</i> der Qualitätsklassen 1 und 4, aufgeteilt nach zur Bewegungserfassung genutztem Sensor.	89
4.8	Standardabweichung der Rotationsinformationen von jeweils 100 Bewegungsereignissen der Körpereigengewichtsübung <i>Bicycle Crunch</i> der Qualitätsklassen 1 und 4, aufgeteilt nach zur Bewegungserfassung genutztem Sensor.	90
4.9	Schematische Darstellung der Konstruktion der einzelnen Ereignisinstanzen I_x auf Basis des Merkmalsvektors X_{E_x} und dem zugehörigen Qualitätslabel L_x	90
4.10	Zuordnungen nach den Qualitätsklassen 1-5 für die sechs untersuchten Ereignisklassen unter Zuhilfenahme eines RF-Klassifikators mit manuellem Parameter-Setting.	93
4.11	Vergleich der Segmentierungseigenschaften des statischen und des adaptiven Segmentierungsalgorithmus am Beispiel einer Folge von 20 Bewegungsereignissen mit variierender Dauer.	95
4.12	Ergebnisse der qualitativen Untersuchung der Ereignisklasse <i>Lunges</i> für die verschiedenen Klassifikatoren unter Benutzung eines statischen und eines adaptiven Segmentierungsansatzes.	96
4.13	Ergebnisse der qualitativen Untersuchung der Ereignisklasse <i>Kniebeuge</i> für die verschiedenen Klassifikatoren unter Benutzung eines statischen und eines adaptiven Segmentierungsansatzes.	96
4.14	Laufzeiten zur Bildung von Vorhersagemodellen mit variierenden Klassifikatoren für das mobile und das stationäre Szenario. Dabei wurden unterschiedliche Kombinationen von auf Interpolation basierenden und auf Kompression basierenden Merkmalsvektoren genutzt.	98
4.15	Korrekte Erkennungsrate für verschiedene Kombinationen aus Klassifikatoren und komprimiertem Design des Merkmalsvektors bei variierender Anzahl von zum Modelltraining genutzter Instanzen im mobilen Online-Szenario.	99
5.1	Übersicht über die adaptierte Prozesskette zur automatisierten, sensorgestützten Untersuchung von Kletteraktivität: (1) Erfassung von Bewegung mit dem SensX Sensorsystem, (2) die Vorverarbeitung und der zweistufige Segmentierungsprozess, (3) das Feature-Engineering und (4) die Klassifizierung mithilfe von überwachten maschinellen Lernalgorithmen.	111

5.2	Beschleunigungsinformationen des SensX Sensorsystems während eines Kletteraufstiegs in $X-$, $Y-$ und $Z-$ Richtung für die linke Hand (1) und die Brust eines Kletterers (2).	112
5.3	Beschleunigungsinformationen des SensX-Sensorsystems während eines Kletteraufstiegs in $X-$, $Y-$ und $Z-$ Richtung für die linke Hand (1) und die daraus extrahierte Kletteraktivität in (2), welche grün hervorgehoben ist.	115
5.4	Darstellung der Überführung des Ergebnisses des ersten Segmentierungsschrittes (1) in den zweiten Schritt (2), bei welchem Ruhephasen (in Grün) und Übergangsphasen (in Rot) voneinander getrennt werden.	116
5.5	Verdichtung der verschiedenen Komponenten zur Beschreibung von physischer, menschlicher Leistung und Kletterkönnen aus verwandten Arbeiten auf die drei in dieser Arbeit aggregierten Kernfähigkeiten.	118
5.6	Übersicht über die Zuordnung der Features aus den Übergangsphasen auf die drei verdichteten Kernfähigkeiten Kraft P , Ausdauer E und Kontrolle C	120
5.7	Übersicht über die Zuordnung der Features aus den Ruhephasen auf die drei verdichteten Kernfähigkeiten Kraft P , Ausdauer E und Kontrolle C	122
5.8	Übersicht über die Zuordnung der Features aus den Ruhe- und Übergangsphasen auf die drei aggregierten Kernfähigkeiten Kraft P , Ausdauer E und Kontrolle C	124
5.9	Histogramm der Korrelation von nach Routenschwierigkeit gelabelten Features mit farblicher Hervorhebung der zugehörigen Oberkategorie der Features.	128
5.10	Prozentuale Verteilung der drei Oberkategorien der Features innerhalb der unterschiedlichen, mithilfe von Feature-Korrelation erstellten Feature-Sets zur Klassifizierung von Routen nach ihrer Schwierigkeit: <i>Subset 1</i> , <i>Subset 2</i> und <i>Alle Features</i>	128
5.11	Verteilung der Features innerhalb ihrer Kategorien nach ihrer auf Korrelation basierten Wertigkeit für die zwei unterschiedlichen Zuordnungen auf Routenschwierigkeit (1) und auf unterschiedliche Studienteilnehmer (2).	129
5.12	Histogramm der Korrelation von nach Studienteilnehmern gelabelten Features mit farblicher Hervorhebung der zugehörigen Feature-Kategorie.	129
5.13	Prozentuale Verteilung der drei Feature-Kategorien innerhalb der unterschiedlichen, mithilfe von Feature-Korrelation erstellten Feature-Sets zur Klassifizierung nach Studienteilnehmern: <i>Subset 1</i> , <i>Subset 2</i> und <i>Alle Features</i>	130

5.14	Klassifizierungsergebnisse des 10fCV-basierten Experiments, aufgeschlüsselt nach den gelabelten Schwierigkeitsbereichen Blau (B), Schweden (S) und Grün (G) und geordnet nach genutztem Klassifikator sowie nach genutztem Feature-Set.	133
5.15	Durchschnittliche Erkennungsrate für die automatisierte Erkennung von Schwierigkeitsgraden (1) und für die Unterscheidung einzelner Studienteilnehmer (2), aufgeschlüsselt nach Feature-Sets und Validierungsverfahren.	133
5.16	Konfusionsmatrizen der mithilfe von überwachtem Lernen voneinander unterschiedenen 16 Studienteilnehmer für alle Kombinationen aus Feature-Set und Klassifikatoren im manuellen Setup während der 10fCV.	135

Literaturverzeichnis

- [1] P. Aagaard. Training-Induced Changes in Neural Function. *Exercise and Sport Sciences Reviews*, 31(2):61–67, 2003.
- [2] J. P. Abt, T. C. Sell, Y. Chu, M. Lovalekar, R. G. Burdett, and S. M. Lephart. Running Kinematics and Shock Absorption do not Change After Brief Exhaustive Running. *The Journal of Strength & Conditioning Research*, 25:1479–1485, 2011.
- [3] R. Agrawal, C. Faloutsos, and A. Swami. Efficient Similarity Search in Sequence Databases. *Foundations of Data Organization and Algorithms*, pages 69–84, 1993.
- [4] U.-V. Albrecht, M. Höhn, and U. von Jan. Chancen und Risiken von Gesundheits-Apps (CHARISMHA). *Chancen und Risiken von Gesundheits-Apps (CHARISMHA)*. Hannover: Medizinische Hochschule Hannover, 2016.
- [5] H. Aljazzar and S. Leue. K*: A Heuristic Search Algorithm for Finding the k Shortest Paths. *Artificial Intelligence*, 175(18):2129–2154, 2011.
- [6] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz. Human Activity Recognition on Smartphones using a Multiclass Hardware-Friendly Support Vector Machine. In *International Workshop on Ambient Assisted Living*, pages 216–223. Springer, 2012.
- [7] M. Ankerst, M. M. Breunig, H.-P. Kriegel, and J. Sander. OPTICS: Ordering Points to Identify the Clustering Structure. In *ACM Sigmod record*, volume 28, pages 49–60. ACM, 1999.
- [8] I. Ar and Y. S. Akgul. A Computerized Recognition System for the Home-based Physiotherapy Exercises Using an RGBD Camera. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(6):1160–1171, 2014.
- [9] S. Asteriadis, A. Chatzitofis, D. Zarpalas, D. S. Alexiadis, and P. Daras. Estimating Human Motion from Multiple Kinect Sensors. In *Proceedings of the 6th International Conference on Computer Vision/Computer Graphics Collaboration Techniques and Applications*, page 3. ACM, 2013.

- [10] L. Bao and S. Intille. Activity Recognition from User-annotated Acceleration Data. *Pervasive Computing*, pages 1–17, 2004.
- [11] C. Bauer. On the (in-) Accuracy of GPS Measures of Smartphones: a Study of Running Tracking Applications. In *Proceedings of International Conference on Advances in Mobile Computing & Multimedia*, page 335. ACM, 2013.
- [12] H. Bay, T. Tuytelaars, and L. Van Gool. SURF: Speeded Up Robust Features. In *European conference on computer vision*, pages 404–417. Springer, 2006.
- [13] J. Bergstra and Y. Bengio. Random Search for Hyper-Parameter Optimization. *Journal of Machine Learning Research*, 13:281–305, 2012.
- [14] J. S. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl. Algorithms for Hyper-Parameter Optimization. In *Advances in Neural Information Processing Systems*, pages 2546–2554, 2011.
- [15] D. J. Berndt and J. Clifford. Using Dynamic Time Warping to find Patterns in Time Series. In *KDD workshop*, volume 10, pages 359–370. Seattle, WA, 1994.
- [16] D. P. Bertsekas, D. P. Bertsekas, D. P. Bertsekas, and D. P. Bertsekas. *Dynamic Programming and Optimal Control*, volume 1. Athena Scientific Belmont, MA, 1995.
- [17] D. P. Bertsekas and J. N. Tsitsiklis. Neuro-Dynamic Programming: an Overview. In *Decision and Control, 1995., Proceedings of the 34th IEEE Conference on*, volume 1, pages 560–564. IEEE, 1995.
- [18] U. Blanke and B. Schiele. Daily Routine Recognition through Activity Spotting. *Location and Context Awareness*, pages 192–206, 2009.
- [19] U. Blanke and B. Schiele. Remember and Transfer what you have Learned - Recognizing Composite Activities based on Activity Spotting. In *Wearable Computers (ISWC), 2010 International Symposium on*, pages 1–8. IEEE, 2010.
- [20] A. Bobick and J. Davis. Real-time Recognition of Activity using Temporal Templates. In *Applications of Computer Vision, 1996. WACV'96., Proceedings 3rd IEEE Workshop on*, pages 39–42. IEEE, 1996.
- [21] C. V. Bouten, K. T. Koekkoek, M. Verduin, R. Kodde, and J. D. Janssen. A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity. *IEEE Transactions on Biomedical Engineering*, 44(3):136–147, 1997.

- [22] L. Breiman. Random Forests. *Machine learning*, 45(1):5–32, 2001.
- [23] M. U. Bromba and H. Ziegler. Application hints for Savitzky-Golay digital smoothing filters. *Analytical Chemistry*, 53(11):1583–1586, 1981.
- [24] A. Bulling, U. Blanke, and B. Schiele. A Tutorial on Human Activity Recognition using Body-worn Inertial Sensors. *ACM Computing Surveys (CSUR)*, 46(3):33, 2014.
- [25] A. Bulling and D. Roggen. Recognition of Visual Memory Recall Processes using Eye Movement Analysis. In *Proceedings of the 13th International Conference on Ubiquitous Computing*, pages 455–464. ACM, 2011.
- [26] A. Bulling, C. Weichel, and H. Gellersen. EyeContext: Recognition of High-Level Contextual Cues from Human Visual Behaviour. In *Proceedings of the SigCHI Conference on Human Factors in Computing Systems*, pages 305–308. ACM, 2013.
- [27] L. Busoniu, R. Babuska, B. De Schutter, and D. Ernst. *Reinforcement Learning and Dynamic Programming Using Function Approximators*, volume 39. CRC Press, 2010.
- [28] S. Butterworth. On the Theory of Filter Amplifiers. *Wireless Engineer*, 7:536–541, 1930.
- [29] K. J. Calfas and W. C. Taylor. Effects of Physical Activity on Psychological Variables in Adolescents. *Pediatric exercise science*, 6:406–423, 1994.
- [30] E. Cambria. Affective Computing and Sentiment Analysis. *IEEE Intelligent Systems*, 31:102–107, 2016.
- [31] B. P. Carlin and T. A. Louis. *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman and Hall/CRC, 2010.
- [32] P. Casale, O. Pujol, and P. Radeva. Human Activity Recognition from Accelerometer Data using a Wearable Device. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 289–296. Springer, 2011.
- [33] G. C. Cawley. Leave-one-out Cross-Validation based Model Selection Criteria for Weighted LS-SVMs. In *Neural Networks, 2006. IJCNN'06. International Joint Conference on*, pages 1661–1668. IEEE, 2006.
- [34] K.-H. Chang, M. Y. Chen, and J. Canny. Tracking Free-Weight Exercises. In *International Conference on Ubiquitous Computing*, pages 19–37. Springer, 2007.

- [35] O. Chapelle, V. Vapnik, O. Bousquet, and S. Mukherjee. Choosing Multiple Parameters for Support Vector Machines. *Machine learning*, 46:131–159, 2002.
- [36] Y. Chen, J. Bi, and J. Z. Wang. MILES: Multiple-Instance Learning via Embedded Instance Selection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):1931–1947, 2006.
- [37] S. Corazza, L. Muendermann, A. Chaudhari, T. Demattio, C. Cobelli, and T. P. Andriacchi. A Markerless Motion Capture System to Study Musculoskeletal Biomechanics: Visual Hull and Simulated Annealing Approach. *Annals of Biomedical Engineering*, 34:1019–1029, 2006.
- [38] R. De Maesschalck, D. Jouan-Rimbaud, and D. L. Massart. The Mahalanobis Distance. *Chemometrics and Intelligent Laboratory Systems*, 50(1):1–18, 2000.
- [39] D. De Ridder and R. P. Duin. Sammon’s Mapping using Neural Networks: a Comparison. *Pattern Recognition Letters*, 18(11-13):1307–1316, 1997.
- [40] S. Dernbach, B. Das, N. C. Krishnan, B. L. Thomas, and D. J. Cook. Simple and Complex Activity Recognition Through Smart Phones. In *Intelligent Environments (IE), 2012 8th International Conference on*, pages 214–221. IEEE, 2012.
- [41] T. Dietterich. Overfitting and Undercomputing in Machine Learning. *ACM Computing Surveys (CSUR)*, 27(3):326–327, 1995.
- [42] H. Ding, L. Shangguan, Z. Yang, J. Han, Z. Zhou, P. Yang, W. Xi, and J. Zhao. Femo: A Platform for Free-Weight Exercise Monitoring with RFIDs. In *Proceedings of the 13th ACM Conference on Embedded Networked Sensor Systems*, pages 141–154. ACM, 2015.
- [43] P. Domingos. A Unified Bias-Variance Decomposition. In *Proceedings of 17th International Conference on Machine Learning*, pages 231–238, 2000.
- [44] P. Domingos. A Few Useful Things to Know about Machine Learning. *Communications of the ACM*, 55(10):78–87, 2012.
- [45] J. Donahue, L. Anne Hendricks, S. Guadarrama, M. Rohrbach, S. Venugopalan, K. Saenko, and T. Darrell. Long-term Recurrent Convolutional Networks for Visual Recognition and Description. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2625–2634, 2015.

- [46] A. Ebert, M. T. Beck, A. Mattausch, L. Belzner, and C. L. Popien. Qualitative Assessment of Recurrent Human Motion. In *Signal Processing Conference (EUSIPCO), 2017 25th European*, pages 316–320. IEEE, 2017.
- [47] A. Ebert, F. Dorfmeister, M. Maier, and C. Linnhoff-Popien. EMMA: A Context-Aware Middleware for Energy Management on Mobile Devices. In *Advances in Human-oriented and Personalized Mechanisms, Technologies, and Services (CENTRIC), 2014 7th International Conference on*, pages 48–53, 2014.
- [48] A. Ebert, F. Dorfmeister, M. Maier, and C. Linnhoff-Popien. Context-, Resource-, and User-Aware Provision of Services on Mobile Devices. In *International Journal On Advances in Networks and Services*, pages 81–91, 2015.
- [49] A. Ebert, S. Feld, and F. Dorfmeister. Segmented and Directional Impact Detection for Parked Vehicles using Mobile Devices. In *Systems, Signals and Image Processing (IWSSIP), 2016 International Conference on*, pages 1–4. IEEE, 2016.
- [50] A. Ebert, S. Feld, and M. Werner. Nuanced Collision Detection for Parked Vehicles. In *Ortsbezogene Anwendungen und Dienste (LBAS), 2015 12. GI/ITG KuVS Fachgespräch*, pages 37–44. Logos Verlag Berlin, 2015.
- [51] A. Ebert, M. Kiermeier, C. Marouane, and C. Linnhoff-Popien. Sensx: About Sensing and Assessment of Complex Human Motion. In *Networking, Sensing and Control (ICNSC), 2017 IEEE 14th International Conference on*, pages 327–332. IEEE, 2017.
- [52] A. Ebert, C. Marouane, B. Rott, and M. Werner. KeyPocket-Improving Security and Usability for Provider Independent Login Architectures with Mobile Devices. In *International Conference on Security and Privacy in Communication Systems*, pages 41–57. Springer, 2015.
- [53] A. Ebert, C. Marouane, C. Ungnadner, and A. Klein. An Open, Labeled Dataset for Analysis and Assessment of Human Motion. In *Wireless Mobile Communication and Healthcare (MobiHealth), 2017 7th EAI Conference on*, pages 1–8. Springer, 2017.
- [54] A. Ebert, K. Schmid, C. Marouane, and C. Linnhoff-Popien. Automated Recognition and Difficulty Assessment of Boulder Routes. In *IoT Technologies for Healthcare (HealthyIoT), 2017 4th EAI Conference on*, pages 62–68. Springer, 2017.
- [55] S. R. Eddy. Hidden Markov Models. *Current Opinion in Structural Biology*, 6(3):361–365, 1996.

- [56] T. Eiter and H. Mannila. Computing Discrete Fréchet Distance. Technical report, Tech. Report CD-TR 94/64, Information Systems Department, Technical University of Vienna, 1994.
- [57] A. Elgammal, R. Duraiswami, D. Harwood, and L. S. Davis. Background and Foreground Modeling using Nonparametric Kernel Density Estimation for Visual Surveillance. *Proceedings of the IEEE*, 90:1151–1163, 2002.
- [58] M. Ester, H.-P. Kriegel, J. Sander, X. Xu, et al. A Density-based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Kdd*, volume 96, pages 226–231, 1996.
- [59] D. J. Fagnant and K. Kockelman. Preparing a Nation for Autonomous Vehicles: Opportunities, Barriers and Policy Recommendations. *Transportation Research Part A: Policy and Practice*, 77:167–181, 2015.
- [60] M. Fanchini, F. Violette, F. M. Impellizzeri, and N. A. Maffiuletti. Differences in Climbing-specific Strength between Boulder and Lead Rock Climbers. *The Journal of Strength & Conditioning Research*, 27:310–314, 2013.
- [61] C. Feeken, M. Wasmann, W. Heuten, D. Ennenga, H. Müller, and S. Boll. ClimbingAssist: Direct Vibro-tactile Feedback on Climbing Technique. In *UbiComp Adjunct*, 2016.
- [62] D. Figo, P. C. Diniz, D. R. Ferreira, and J. M. Cardoso. Preprocessing Techniques for Context Recognition from Accelerometer Data. *Personal and Ubiquitous Computing*, 14(7):645–662, 2010.
- [63] G. A. Fink. *Markov Models for Pattern Recognition: From Theory to Applications*. Springer Science & Business Media, 2014.
- [64] G. Flecken and D. Heise-Flecken. *Klettern in der Halle*. Meyer & Meyer Verlag, 2015.
- [65] G. D. Forney. The Viterbi Algorithm. *Proceedings of the IEEE*, 61:268–278, 1973.
- [66] E. Frank, M. Hall, and B. Pfahringer. Locally weighted naive bayes. In *19th Conference in Uncertainty in Artificial Intelligence*, pages 249–256. Morgan Kaufmann, 2003.
- [67] N. Friedman, D. Geiger, and M. Goldszmidt. Bayesian Network Classifiers. *Machine learning*, 29:131–163, 1997.
- [68] D. M. Gavrilu. The Visual Analysis of Human Movement: A Survey. *Computer Vision and Image Understanding*, 73:82–98, 1999.

-
- [69] D. Goddard and U. Neumann. *Performance Rock Climbing*. Stackpole Books, 1993.
- [70] P. M. Granitto, C. Furlanello, F. Biasioli, and F. Gasperi. Recursive Feature Elimination with Random Forest for PTR-MS Analysis of Agroindustrial Products. *Chemometrics and Intelligent Laboratory Systems*, 83(2):83–90, 2006.
- [71] E. Guenterberg, S. Ostadabbas, H. Ghasemzadeh, and R. Jafari. An Automatic Segmentation Technique in Body Sensor Networks based-on Signal Energy. In *Proceedings of the Fourth International Conference on Body Area Networks*, page 21. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2009.
- [72] I. Guyon and A. Elisseeff. An Introduction to Variable and Feature Selection. *Journal of machine learning research*, 3(Mar):1157–1182, 2003.
- [73] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten. The weka data mining software: An update. *ACM SIGKDD Explorations Newsletter*, 11:10–18, 2009.
- [74] M. A. Hall. Correlation-based Feature Selection for Machine Learning. *PHD Thesis, University of Waikato*, 1999.
- [75] M. A. Hall and L. A. Smith. Feature Selection for Machine Learning: Comparing a Correlation-based Filter Approach to the Wrapper. In *FLAIRS Conference*, volume 1999, pages 235–239, 1999.
- [76] T. Hastie, R. Tibshirani, and J. Friedman. Unsupervised Learning. In *The Elements of Statistical Learning*, pages 485–585. Springer, 2009.
- [77] W. G. Hawkins. Fourier Transform Resampling: Theory and Application. *IEEE Transactions on Nuclear Science*, 44:1543–1551, 1997.
- [78] S. S. Haykin, S. S. Haykin, S. S. Haykin, and S. S. Haykin. *Neural Networks and Learning Machines*, volume 3. Pearson Upper Saddle River, NJ, USA:, 2009.
- [79] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf. Support Vector Machines. *IEEE Intelligent Systems and their applications*, 13(4):18–28, 1998.
- [80] P. S. Heckbert and M. Garland. Survey of Polygonal Surface Simplification Algorithms. Technical report, Carnegie-Mellon University Pittsburgh, PA School of Computer Science, 1997.

- [81] J. R. Hershey and P. A. Olsen. Approximating the Kullback Leibler Divergence between Gaussian Mixture Models. In *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, volume 4, pages IV–317. IEEE, 2007.
- [82] G. Holmes, A. Donkin, and I. H. Witten. Weka: A Machine Learning Workbench. In *Intelligent Information Systems, 1994. Proceedings of the 1994 Second Australian and New Zealand Conference on*, pages 357–361. IEEE, 1994.
- [83] E. J. Hörst. *Training for Climbing: The Definitive Guide to Improve your Climbing Performance*. Globe Pequot, 2003.
- [84] J. Hunter and N. McIntosh. Knowledge-Based Event Detection in Complex Time Series Data. In *Joint European Conference on Artificial Intelligence in Medicine and Medical Decision Making*, pages 271–280. Springer, 1999.
- [85] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge. Comparing Images Using the Hausdorff Distance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):850–863, 1993.
- [86] F. Hutter, H. H. Hoos, and K. Leyton-Brown. Sequential Model-based Optimization for General Algorithm Configuration. In *International Conference on Learning and Intelligent Optimization*, pages 507–523. Springer, 2011.
- [87] T. Huynh, M. Fritz, and B. Schiele. Discovery of Activity Patterns using Topic Models. In *Proceedings of the 10th International Conference on Ubiquitous Computing*, pages 10–19. ACM, 2008.
- [88] T. Huynh and B. Schiele. Analyzing Features for Activity Recognition. In *Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-aware Services: Usages and Technologies*, pages 159–163. ACM, 2005.
- [89] S. S. Intille, L. Bao, E. M. Tapia, and J. Rondoni. Acquiring in Situ Training Data for Context-aware Ubiquitous Computing Applications. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 1–8. ACM, 2004.
- [90] F. Itakura. Minimum Prediction Residual Principle applied to Speech Recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 23(1):67–72, 1975.
- [91] W. Jiang and Z. Yin. Human Activity Recognition using Wearable Sensors by Deep Convolutional Neural Networks. In *Proceedings of the*

- 23rd ACM International Conference on Multimedia*, pages 1307–1310. ACM, 2015.
- [92] I. T. Jolliffe. Principal Component Analysis and Factor Analysis. In *Principal Component Analysis*, pages 115–128. Springer, 1986.
- [93] B. H. Jones, M. W. Bovee, J. M. Harris III, and D. N. Cowan. Intrinsic Risk Factors for Exercise-Related Injuries Among Male and Female Army Trainees. *The American Journal of Sports Medicine*, 21:705–710, 1993.
- [94] H. Kagermann. Chancen von Industrie 4.0 nutzen. In *Handbuch Industrie 4.0 Bd. 4*, pages 237–248. Springer, 2017.
- [95] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler. Implementation of a Real-time Human Movement Classifier using a Triaxial Accelerometer for Ambulatory Monitoring. *IEEE Transactions on Information Technology in Biomedicine*, 10(1):156–167, 2006.
- [96] G. N. Karystinos and D. A. Pados. On Overfitting, Generalization, and Randomly Expanded Training Sets. *IEEE Transactions on Neural Networks*, 11(5):1050–1057, 2000.
- [97] J. M. Keller, M. R. Gray, and J. A. Givens. A Fuzzy k-Nearest Neighbor Algorithm. *IEEE Transactions on Systems, Man, and Cybernetics*, 4:580–585, 1985.
- [98] E. Keogh, S. Chu, D. Hart, and M. Pazzani. An Online Algorithm for Segmenting Time Series. In *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, pages 289–296. IEEE, 2001.
- [99] E. Keogh, S. Chu, D. Hart, and M. Pazzani. Segmenting Time Series: A Survey and Novel Approach. In *Data Mining in Time Series Databases*, pages 1–21. World Scientific, 2004.
- [100] E. Keogh and C. A. Ratanamahatana. Exact Indexing of Dynamic Time Warping. *Knowledge and Information Systems*, 7(3):358–386, 2005.
- [101] N. Kern, B. Schiele, and A. Schmidt. Recognizing Context for Annotating a Live Life Recording. *Personal and Ubiquitous Computing*, 11(4):251–263, 2007.
- [102] R. Keys. Cubic Convolution Interpolation for Digital Image Processing. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 29(6):1153–1160, 1981.

- [103] A. Khan, S. Mellor, E. Berlin, R. Thompson, R. McNaney, P. Olivier, and T. Plötz. Beyond Activity Recognition: Skill Assessment from Accelerometer Data. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 1155–1166. ACM, 2015.
- [104] E. Kim, S. Helal, and D. Cook. Human Activity Recognition and Pattern Discovery. *IEEE Pervasive Computing*, 9(1), 2010.
- [105] I. Kim, S. Im, E. Hong, S. C. Ahn, and H.-G. Kim. ADL Classification Using Triaxial Accelerometers and RFID. In *International Conference on Ubiquitous Computing Convergence Technology*, 2007.
- [106] R. Kohavi et al. A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. In *IJCAI*, volume 14, pages 1137–1145. Montreal, Canada, 1995.
- [107] R. Kohavi and G. H. John. Wrappers for Feature Subset Selection. *Artificial intelligence*, 97(1-2):273–324, 1997.
- [108] T. Komura, B. Lam, R. W. Lau, and H. Leung. e-Learning Martial Arts. In *International Conference on Web-Based Learning*, pages 239–248. Springer, 2006.
- [109] J. P. Koplan, D. S. Siscovick, and G. M. Goldbaum. The Risks of Exercise: A Public Health View of Injuries and Hazards. *Public Health Reports*, 100:189, 1985.
- [110] H. Koskimaki, V. Huikari, P. Siirtola, P. Laurinen, and J. Roning. Activity Recognition Using a Wrist-Worn Inertial Measurement Unit: A Case Study for Industrial Assembly Lines. In *Control and Automation, 2009. MED'09. 17th Mediterranean Conference on*, pages 401–405. IEEE, 2009.
- [111] F. Kosmalla, F. Daiber, and A. Krüger. Climbsense: Automatic Climbing Route Recognition using Wrist-worn Inertia Measurement Units. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 2033–2042. ACM, 2015.
- [112] L. Kotthoff, C. Thornton, H. H. Hoos, F. Hutter, and K. Leyton-Brown. Auto-WEKA 2.0: Automatic Model Selection and Hyperparameter Optimization in WEKA. *The Journal of Machine Learning Research*, 18:826–830, 2017.
- [113] S. Kozina, M. Lustrek, and M. Gams. Dynamic Signal Segmentation for Activity Recognition. In *Proceedings of International Joint Conference on Artificial Intelligence, Barcelona, Spain*, volume 1622, page 1522, 2011.

- [114] A. F. Kramer and K. I. Erickson. Capitalizing on Cortical Plasticity: Influence of Physical Activity on Cognition and Brain Function. *Trends in cognitive sciences*, 11:342–348, 2007.
- [115] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet Classification with Deep Convolutional Neural Networks. In *Advances in Neural Information Processing Systems*, pages 1097–1105, 2012.
- [116] H. Kuehne, H. Jhuang, E. Garrote, T. Poggio, and T. Serre. HMDB: A Large Video Database for Human Motion Recognition. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 2556–2563. IEEE, 2011.
- [117] C. Ladha, N. Y. Hammerla, P. Olivier, and T. Plötz. ClimbAX: Skill Assessment for Climbing Enthusiasts. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pages 235–244. ACM, 2013.
- [118] O. D. Lara and M. A. Labrador. A Survey on Human Activity Recognition using Wearable Sensors. *IEEE Communications Surveys and Tutorials*, 15:1192–1209, 2013.
- [119] J. Larsen, L. K. Hansen, C. Svarer, and M. Ohlsson. Design and Regularization of Neural Networks: the Optimal Use of a Validation Set. In *Neural Networks for Signal Processing [1996] VI. Proceedings of the 1996 IEEE Signal Processing Society Workshop*, pages 62–71. IEEE, 1996.
- [120] J. R. Leaf, J. L. Keating, and G. S. Kolt. Injury in the Australian Sport of Calisthenics: A Prospective Study. *Australian Journal of Physiotherapy*, 49:123–130, 2003.
- [121] J. Lee Rodgers and W. A. Nicewander. Thirteen Ways to Look at the Correlation Coefficient. *The American Statistician*, 42(1):59–66, 1988.
- [122] A. Liaw, M. Wiener, et al. Classification and Regression by Random Forest. *R news*, 2(3):18–22, 2002.
- [123] S. Lloyd. Least Squares Quantization in PCM. *IEEE Transactions on Information Theory*, 28(2):129–137, 1982.
- [124] Y. Lv, Y. Duan, W. Kang, Z. Li, and F.-Y. Wang. Traffic Flow Prediction with Big Data: a Deep Learning Approach. *IEEE Transactions on Intelligent Transportation Systems*, 16:865–873, 2015.
- [125] M. Maier and F. Dorfmeister. Fine-Grained Activity Recognition of Pedestrians Travelling by Subway. In *International Conference on Mobile Computing, Applications, and Services*, pages 122–139. Springer, 2013.

- [126] A. Mannini and A. M. Sabatini. Machine Learning Methods for Classifying Human Physical Activity from On-Body Accelerometers. *Sensors*, 10:1154–1175, 2010.
- [127] C. Marouane. *Visuelle Verfahren für ortsbezogene Dienste*. PhD thesis, LMU, 2017.
- [128] C. Marouane and A. Ebert. Enabling Pedometers on Basis of Visual Feature Point Conversion. *1st GI Expert Talk on Localization*, pages 33–35, 2015.
- [129] C. Marouane and A. Ebert. SURFLogo-Mobile Tagging with App Icons. In *International Conference on Mobile Computing, Applications, and Services*, pages 239–256. Springer, 2015.
- [130] C. Marouane, A. Ebert, C. Linnhoff-Popien, and M. Christil. Step and Activity Detection based on the Orientation and Scale Attributes of the SURF Algorithm. In *Indoor Positioning and Indoor Navigation (IPIN), 2016 International Conference on*, pages 1–8. IEEE, 2016.
- [131] C. Marouane, A. Ebert, and B. Rott. Trends und Chancen beim mobilen Einkaufen. In *Marktplätze im Umbruch*, pages 289–298. Springer, 2015.
- [132] J. J. Martin. *Bayesian Decision Problems and Markov chains*. Wiley, 1967.
- [133] M. Mathie, B. G. Celler, N. H. Lovell, and A. Coster. Classification of Basic Daily Movements using a triaxial Accelerometer. *Medical and Biological Engineering and Computing*, 42(5):679–687, 2004.
- [134] A. Menache. *Understanding Motion Capture for Computer Animation and Video Games*. Morgan Kaufmann, 2000.
- [135] D. Miorandi, S. Sicari, F. De Pellegrini, and I. Chlamtac. Internet of things: Vision, Applications and Research Challenges. *Ad hoc networks*, 10:1497–1516, 2012.
- [136] L. C. Molina, L. Belanche, and À. Nebot. Feature Selection Algorithms: A survey and Experimental Evaluation. In *Data Mining, 2002. ICDM 2003. Proceedings. 2002 IEEE International Conference on*, pages 306–313. IEEE, 2002.
- [137] A. Möller, L. Roalter, S. Diewald, J. Scherr, M. Kranz, N. Hammerla, P. Olivier, and T. Plötz. Gymskill: A personal trainer for physical exercises. In *Pervasive Computing and Communications (PerCom), 2012 IEEE International Conference on*, pages 213–220. IEEE, 2012.

- [138] T. K. Moon. The Expectation-Maximization Algorithm. *IEEE Signal Processing Magazine*, 13:47–60, 1996.
- [139] D. Morris, T. S. Saponas, A. Guillory, and I. Kelner. RecoFit: Using a Wearable Sensor to Find, Recognize, and Count Repetitive Exercises. In *Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems*, pages 3225–3234. ACM, 2014.
- [140] K. P. Murphy. Naive Bayes Classifiers. *University of British Columbia*, 18, 2006.
- [141] N. M. Nasrabadi. Pattern Recognition and Machine Learning. *Journal of Electronic Imaging*, 16(4):049901, 2007.
- [142] A. Y. Ng. Preventing Overfitting of Cross-Validation Data. In *ICML*, volume 97, pages 245–253, 1997.
- [143] B. Nham, K. Siangliulue, and S. Yeung. Predicting Mode of Transport from iPhone Accelerometer Data. *Stanford University Class Project*, 2008.
- [144] C. L. Nikias and J. M. Mendel. Signal Processing with Higher-Order Spectra. *IEEE Signal Processing Magazine*, 10(3):10–37, 1993.
- [145] R. S. Olson, R. J. Urbanowicz, P. C. Andrews, N. A. Lavender, J. H. Moore, et al. Automating biomedical data science through tree-based pipeline optimization. In *European Conference on the Applications of Evolutionary Computation*, pages 123–137. Springer, 2016.
- [146] T. E. Paige, D. C. Fiore, and J. D. Houston. Injury in Traditional and Sport Rock Climbing. *Wilderness & Environmental Medicine*, 9:2–7, 1998.
- [147] J. Pansiot, R. C. King, D. G. McIlwraith, B. P. Lo, and G.-Z. Yang. ClimBSN: Climber Performance Monitoring with BSN. In *Medical Devices and Biosensors, 2008. ISSS-MDBS 2008. 5th International Summer School and Symposium on*, pages 33–36. IEEE, 2008.
- [148] G. I. Parisi, S. Magg, and S. Wermter. Human Motion Assessment in Real Time using Recurrent Self-Organization. In *Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on*, pages 71–76. IEEE, 2016.
- [149] H. Peng, F. Long, and C. Ding. Feature Selection Based on Mutual Information Criteria of Max-Dependency, Max-Relevance, and Min-Redundancy. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27(8):1226–1238, 2005.

- [150] I. Pernek, K. A. Hummel, and P. Kokol. Exercise Repetition Detection for Resistance Training based on Smartphones. *Personal and ubiquitous computing*, 17(4):771–782, 2013.
- [151] I. Peter and A. Fyffe. *Handbook of Climbing*. Pelham Books, London, 1997.
- [152] A. Pfister, A. M. West, S. Bronner, and J. A. Noah. Comparative Abilities of Microsoft Kinect and Vicon 3D motion Capture for Gait Analysis. *Journal of Medical Engineering & Technology*, 38:274–280, 2014.
- [153] T. Plötz, N. Y. Hammerla, A. Rozga, A. Reavis, N. Call, and G. D. Abowd. Automatic Assessment of Problem Behavior in Individuals with Developmental Disabilities. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, pages 391–400. ACM, 2012.
- [154] K. Pullen and C. Bregler. Motion Capture Assisted Animation: Texturing and Synthesis. In *ACM Transactions on Graphics (TOG)*, volume 21, pages 501–508. ACM, 2002.
- [155] J. R. Quinlan. *C4. 5: Programs for Machine Learning*. Elsevier, 2014.
- [156] L. R. Rabiner. A tutorial on hidden Markov Models and Selected Applications in Speech Recognition. *Proceedings of the IEEE*, 77(2):257–286, 1989.
- [157] Z. Radak, N. Hart, L. Sarga, E. Koltai, M. Atalay, H. Ohno, and I. Boldogh. Exercise Plays a Preventive Role Against Alzheimer’s Disease. *Journal of Alzheimer’s disease*, 20:777–783, 2010.
- [158] Splitting the Data into Training and Evaluation Data - Amazon Machine Learning. Online, 2018. <https://docs.aws.amazon.com/machine-learning/latest/dg/splitting-the-data-into-training-and-evaluation-data.html>, letzter Abruf: 29.01.2018.
- [159] Apple Watch - Apple (DE). Online, 2018. <https://www.apple.com/de/watch/>, letzter Abruf: 01.02.2018.
- [160] BPM Fitness Studios. Online, 2018. <https://play.google.com/store/apps/details?id=com.technogym.bpmfitnessstudios>, letzter Abruf: 04.07.2018.
- [161] Unfallstatistik - Sicherheit - Bergsport - Deutscher Alpenverein (DAV). Online, 2018. https://www.alpenverein.de/der-dav/presse/presse-aktuell/bergunfallstatistik-2016_aid_30183.html, letzter Abruf: 04.04.2018.

- [162] Google Playstore - Fitbit Coach | Individuelles Training. Online, 2018. <https://play.google.com/store/apps/details?id=com.fitstar.pt>, letzter Abruf: 01.02.2018.
- [163] Offizielle Webseite für fitbit Aktivitätstracker und mehr. Online, 2018. <https://www.fitbit.com/de/home>, letzter Abruf: 01.02.2018.
- [164] An der Leine der Versicherer. Online, 2015. <https://www.zeit.de/wirtschaft/2015-12/fitnesstracker-krankenkassen-gesundheitsapps>, letzter Abruf: 11.05.2018.
- [165] Intensive Workouts & individuelles Training | Freeletics. Online, 2018. <https://www.freeletics.com/de>, letzter Abruf: 01.02.2018.
- [166] Gartner Says 8.4 Billion Connected "Things" Will Be in Use in 2017, Up 31 Percent From 2016. Online, 2018. <https://www.gartner.com/newsroom/id/3598917>, letzter Abruf: 23.05.2018.
- [167] Google Playstore - Google Fit. Online, 2018. <https://play.google.com/store/apps/details?id=com.google.android.apps.fitness>, letzter Abruf: 01.02.2018.
- [168] Aus der Depression rausklettern - Bouldern als Therapie? Online, 2018. <http://www.fr.de/wissen/gesundheit/trendsport-aus-der-depression-raus-klettern-bouldern-als-therapie-a-1315789>, letzter Abruf: 04.04.2018.
- [169] Klettern und Bouldern: Wie ein Bergsport zum Trendsport wird. Online, 2018. https://www.ispo.com/trends/id_79326674/klettern-und-bouldern-aus-bergsport-wird-fitness-sport.html, letzter Abruf: 04.04.2018.
- [170] Android Developers - Motion Sensors. Online, 2018. https://developer.android.com/guide/topics/sensors/sensors_motion.html, letzter Abruf: 12.02.2018.
- [171] Google Playstore. Online, 2018. <https://play.google.com/store?hl=de>, letzter Abruf: 08.06.2018.
- [172] Enflux Motion Capture Clothing | GetEnflux. Online, 2018. <https://www.getenflux.com/pages/product>, letzter Abruf: 15.01.2018.
- [173] MBientLab Sensors. Online, 2018. <https://mbientlab.com/store/sensors>, letzter Abruf: 15.01.2018.

- [174] Model Selection and Train/Validation/Test Sets - Stanford University | Coursera. Online, 2018. <https://www.coursera.org/learn/machine-learning/lecture/QGKbr/model-selection-and-train-validation-sets>, letzter Abruf: 29.01.2018.
- [175] Beweg Dich, Deutschland! Online, 2016. <https://www.tk.de/centaurus/servlet/contentblob/819848/Datei/74792/TK-Bewegungsstudie-2016-Beweg-dich-Deutschland.pdf>, letzter Abruf: 11.05.2018.
- [176] Ein Boom mit Nebenwirkung. Online, 2018. <http://www.faz.net/aktuell/sport/mehr-sport/trendsport-klettern-ein-boom-mit-nebenwirkung-1881982.html>, letzter Abruf: 04.04.2018.
- [177] Products - Xsens 3D Motion Tracking. Online, 2018. <https://www.xsens.com/products>, letzter Abruf: 15.01.2018.
- [178] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman. Activity Recognition from Accelerometer Data. In *AAAI*, volume 5, pages 1541–1546, 2005.
- [179] E. S. Ristad and P. N. Yianilos. Learning String-edit dDistance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20:522–532, 1998.
- [180] B. Robert, B. White, D. Renter, and R. Larson. Evaluation of Three-dimensional Accelerometers to Monitor and Classify Behavior Patterns in Cattle. *Computers and Electronics in Agriculture*, 67(1):80–84, 2009.
- [181] S. J. Russell, P. Norvig, J. F. Canny, J. M. Malik, and D. D. Edwards. *Artificial Intelligence: a Modern Approach*, volume 2. Prentice Hall Upper Saddle River, 2003.
- [182] H. Sakoe and S. Chiba. Dynamic Programming Algorithm Optimization for Spoken Word Recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 26(1):43–49, 1978.
- [183] S. Salvador and P. Chan. Toward Accurate Dynamic Time Warping in Linear Time and Space. *Intelligent Data Analysis*, 11(5):561–580, 2007.
- [184] A. Savitzky and M. J. Golay. Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical chemistry*, 36(8):1627–1639, 1964.
- [185] R. W. Schafer and L. R. Rabiner. A Digital Signal Processing Approach to Interpolation. *Proceedings of the IEEE*, 61(6):692–702, 1973.

- [186] C. Schläger, A. Ebert, A. Mattausch, and M. Beck. *Enabling Cyber Sovereignty: with Knowledge, Not with National Products*, pages 895–904. Springer Berlin Heidelberg, Berlin, Heidelberg, 2018.
- [187] J. Shawe-Taylor and N. Cristianini. *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.
- [188] M. Shoaib, S. Bosch, O. D. Incel, H. Scholten, and P. J. Havinga. Complex Human Activity Recognition using Smartphone and Wrist-worn Motion Sensors. *Sensors*, 16:426, 2016.
- [189] K. Simonyan and A. Zisserman. Two-Stream Convolutional Networks for Action Recognition in Videos. In *Advances in Neural Information Processing Systems*, pages 568–576, 2014.
- [190] J. Skaloud, A. Bruton, and K.-P. Schwarz. Detection and Filtering of Short-Term ($1/f^\gamma$) Noise in Inertial Sensors. *Navigation*, 46(2):97–107, 1999.
- [191] J. Snoek, H. Larochelle, and R. P. Adams. Practical Bayesian Optimization of Machine Learning Algorithms. In *Advances in Neural Information Processing Systems*, pages 2951–2959, 2012.
- [192] M. Soliman, T. Abiodun, T. Hamouda, J. Zhou, and C.-H. Lung. Smart Home: Integrating Internet of Things with Web Services and Cloud Computing. In *Cloud Computing Technology and Science (CloudCom), 2013 IEEE 5th International Conference on*, volume 2, pages 317–320. IEEE, 2013.
- [193] P. Somol, J. Novovičová, and P. Pudil. Flexible-hybrid Sequential Floating Search in Statistical Feature Selection. *Structural, Syntactic, and Statistical Pattern Recognition*, pages 632–639, 2006.
- [194] K. Soomro, A. R. Zamir, and M. Shah. UCF101: A Dataset of 101 Human Actions Classes from Videos in the Wild. *arXiv preprint arXiv:1212.0402*, 2012.
- [195] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. Dropout: a Simple Way to Prevent Neural Networks from Overfitting. *Journal of Machine Learning Research*, 15(1):1929–1958, 2014.
- [196] I. Steinwart and A. Christmann. *Support Vector Machines*. Springer Science & Business Media, 2008.
- [197] C. Strohrmann, H. Harms, G. Tröster, S. Hensler, and R. Müller. Out of the Lab and into the Woods: Kinematic Analysis in Running using Wearable Sensors. In *Proceedings of the 13th International Conference on Ubiquitous Computing*, pages 119–122. ACM, 2011.

- [198] R. S. Sutton and A. G. Barto. *Reinforcement Learning: An Introduction*, volume 1. MIT Press Cambridge, 1998.
- [199] C. Thornton, F. Hutter, H. H. Hoos, and K. Leyton-Brown. Auto-WEKA: Combined Selection and Hyperparameter Optimization of Classification Algorithms. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 847–855. ACM, 2013.
- [200] T. Van Kasteren, A. Noulas, G. Englebienne, and B. Kröse. Accurate Activity Recognition in a Home Setting. In *Proceedings of the 10th International Conference on Ubiquitous Computing*, pages 1–9. ACM, 2008.
- [201] E. Velloso, A. Bulling, H. Gellersen, W. Ugulino, and H. Fuks. Qualitative Activity Recognition of Weight Lifting Exercises. In *Proceedings of the 4th Augmented Human International Conference*, pages 116–123. ACM, 2013.
- [202] K. Wagstaff, C. Cardie, S. Rogers, S. Schrödl, et al. Constrained k-Means Clustering with Background Knowledge. In *ICML*, volume 1, pages 577–584, 2001.
- [203] M. E. Wall, A. Rechtsteiner, and L. M. Rocha. Singular Value Decomposition and Principal Component Analysis. In *A Practical Approach to Microarray Data Analysis*, pages 91–109. Springer, 2003.
- [204] L. Wang, Y. Qiao, and X. Tang. Action Recognition with Trajectory-Pooled Deep-Convolutional Descriptors. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4305–4314, 2015.
- [205] Z. Wang, F. Hutter, M. Zoghi, D. Matheson, and N. de Freitas. Bayesian Optimization in a Billion Dimensions via Random Embeddings. *Journal of Artificial Intelligence Research*, 55:361–387, 2016.
- [206] P. B. Watts. Physiology of Difficult Rock Climbing. *European journal of applied physiology*, 91:361–372, 2004.
- [207] P. B. Watts, D. T. Martin, and S. Durtschi. Anthropometric Profiles of Elite Male and Female Competitive Sport Rock Climbers. *Journal of sports sciences*, 11:113–117, 1993.
- [208] S. Wold, K. Esbensen, and P. Geladi. Principal Component Analysis. *Chemometrics and Intelligent Laboratory Systems*, 2(1-3):37–52, 1987.
- [209] D. Wright, T. Royle, and T. Marshall. Indoor Rock Climbing: Who Gets Injured? *British Journal of Sports Medicine*, 35:181–185, 2001.

- [210] T. R. Wylie et al. *The Discrete Fréchet Distance with Applications*. PhD thesis, Montana State University-Bozeman, College of Engineering, 2013.
- [211] B. Xu, L. Da Xu, H. Cai, C. Xie, J. Hu, and F. Bu. Ubiquitous Data Accessing Method in IoT-based Information System for Emergency Medical Services. *IEEE Transactions on Industrial Informatics*, 10:1578–1586, 2014.
- [212] X. Yang and Y. L. Tian. Eigenjoints-based Action Recognition using Naive-Bayes-Nearest-Neighbor. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*, pages 14–19. IEEE, 2012.
- [213] B. Yu, D. Gabriel, L. Noble, and K.-N. An. Estimate of the Optimum Cutoff Frequency for the Butterworth Low-Pass Digital Filter. *Journal of Applied Biomechanics*, 15(3):318–329, 1999.
- [214] L. Yujian and L. Bo. A Normalized Levenshtein Distance Metric. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29:1091–1095, 2007.
- [215] A. Zell. *Simulation neuronaler Netze*, volume 1. Addison-Wesley Bonn, 1994.
- [216] M. Zhang and A. A. Sawchuk. Motion Primitive-based Human Activity Recognition using a Bag-of-Features Approach. In *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*, pages 631–640. ACM, 2012.
- [217] A. Zinnen, C. Wojek, and B. Schiele. Multi Activity Recognition based on Bodymodel-derived Primitives. In *International Symposium on Location- and Context-Awareness*, pages 1–18. Springer, 2009.