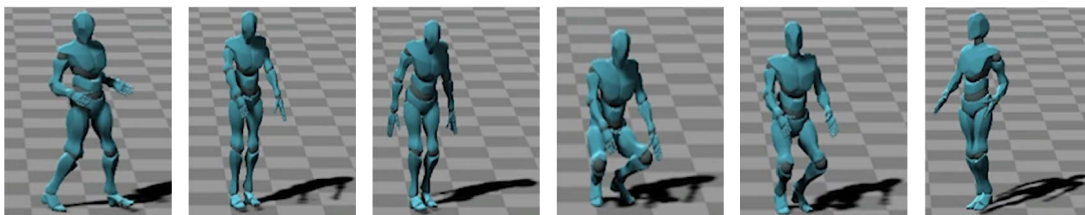


Cyber-physischer Zwilling – Framework zur Generierung menschlicher Bewegungsdaten in der Intralogistik

Cyber-physical twin – framework for generating human movement data in intralogistics

Hülya Avsar
Friedrich Niemann
Christopher Reining
Michael ten Hompel

Lehrstuhl für Förder- und Lagerwesen
Fakultät Maschinenbau
Technische Universität Dortmund



Menschliche Bewegungen zu erkennen, sie zu deuten und für die Analyse manueller Prozesse relevanten Aktivitäten zuzuordnen sind zentrale Herausforderungen der Human Activity Recognition (HAR). Diesen Herausforderungen geht das Trainieren eines Klassifikators mit Daten voraus. Die Erstellung dieser Trainingsdatensätze, bestehend aus Datenaufnahme, Annotation und Revision von Zeitreihen, bedingt einen immensen Aufwand. Aus diesem Grund werden HAR-Methoden überwiegend an simplen Alltagssituationen getestet. Um HAR-Methoden auch für komplexe Umgebungen wie die Intralogistik entwickeln zu können, ist eine neue Form der Datensatz-erstellung notwendig. Dieser Beitrag schlägt ein Framework vor, den Aufwand der Datenaufnahme durch Zuhilfenahme cyber-physischer Zwillinge von Menschen zu reduzieren.

[Schlüsselwörter: Cyber-physischer Zwilling, Simulation, Datengenerierung, Menschliche Aktivitätserkennung]

Recognizing human movements, interpreting them and assigning relevant activities for the analysis of manual processes are central challenges of Human Activity Recognition (HAR). These challenges are preceded by training a classifier with data. The creation of these training data sets, consisting of data acquisition, annotation and revision of time series, requires immense effort. For this reason, HAR methods are mainly tested on simple everyday situations. A new form of data set creation is necessary to develop HAR methods for complex environments such as intralogistics. This contribution proposes a framework to reduce the effort of data acquisition by using cyber-physical twins.

[Keywords: Cyber-physical twin, Simulation, Data generation, Human Activity Recognition]

1 HUMAN ACTIVITY RECOGNITION IN DER INTRALOGISTIK

Die zunehmende Dezentralisierung intralogistischer Systeme und die damit einhergehenden internen Veränderungen der manuellen Arbeits- und Bewegungsabläufe bedürfen einer stetigen Anpassung von Zeitdaten an aktuelle Prozesse. Klassische Methoden der Zeitwirtschaft, wie z. B. Refa oder MTM (Methods-Time Measurement), bieten keine wirtschaftliche Möglichkeit, diese Anpassung regelmäßig oder sogar kontinuierlich vorzunehmen [Bin10, S. 874f]. Ein Grund dafür ist der hohe Aufwand für die Erfassung, Analyse und Administration bei gleichzeitig geringem Abdeckungsgrad des Gesamtpersonals [Bin10, S. 875], [Olb93, S. 65], [LDL16, S. 199], [Sat07, S. 84].

Menschliche Aktivitätserkennung, auch bekannt als Human Activity Recognition (HAR), stellt einen neuen Ansatz zur Zeitdatenermittlung dar. HAR verarbeitet Eingangsdaten wie Videos, Bilder oder Beschleunigungswerte körpergetragener Sensoren, um Aktivitäten automatisch zu erkennen und zeitlich zu erfassen [RNM19]. Wissenschaftliche Untersuchungen belegen das Potenzial von HAR für die Intralogistik [FMH16], [FHA16], [RRH18], [RSH18], [RRN20], [AAR21], [KSH21]. Grundvoraussetzung für den Einsatz von HAR ist das Trainieren eines Klassifikators mit sogenannten Trainingsdaten. Sie bestehen aus Sensordaten menschlicher Bewegungen, welche in der realen Umgebung bzw. im Labor aufgenommen werden. Eine Studie zeigt, dass allein für die Datenaufnahme ein Aufwandsverhältnis von 1:15,6 besteht [NRM20]. Das bedeutet, für eine einstündige Aufnahme sind 15,6 Personenstunden notwendig. In der anschließenden

Annotation werden die Sensordaten in Zeitfenster aufgeteilt und gekennzeichnet, auch labeln genannt (siehe Schritt zwei in Abbildung 1). Durch das Labeln werden die Daten maschinenlesbar. Abschließend erfolgt eine Kontrolle (Revision) der Annotation.

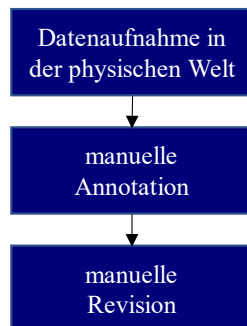


Abbildung 1: Erstellung eines Trainingsdatensatzes für HAR - klassische Schritte der Daten-Pipeline

Mittels vorhandener Bewegungsdaten können neue Trainingsdaten generiert werden, ohne Daten kosten- und zeitintensiv in der physischen Realität aufzunehmen [LGC20]. Dieser Beitrag präsentiert ein Framework zur Generierung von synthetischen Bewegungsdaten mit einer simulationsbasierten Methodik. Dabei wird die erste Stufe (Datenaufnahme) durch ein Framework ersetzt. Statt die menschlichen Bewegungen im Labor oder im realen System aufzunehmen, werden Daten simulativ und somit aufwandsärmer erhoben.

Die Güte der Bewegungsdaten wird durch ihre Realitätsnähe bestimmt. Eine realitätsnahe Umsetzung kann durch die Inputdaten beeinflusst werden. Um logistische Bewegungen aus der physischen Welt in die cyber-physische Welt zu überführen, werden hier logistische Datensätze als primärer Input verwendet. Eine beispielhafte Abfolge menschlicher Aktivitäten in einer Simulationsumgebung ist in der Abbildung zu Beginn des vorliegenden Beitrags visualisiert. Neben domänenspezifischen Datensätzen wird die Machbarkeit und der Nutzen zur Einbindung von sekundären Datensätzen aus domänenfremden Bereichen als möglicher Input erforscht. Nach der Einbindung von primären und sekundären Inputdaten werden die menschlichen Bewegungen modifiziert, wodurch synthetische und damit neue Bewegungen entstehen. Durch die Erweiterung der Inputdaten um diese synthetischen Bewegungsdaten werden große Datenmengen erzeugt und auf zusätzliche Datenaufnahmen kann so verzichtet werden. Doch ist der Bedarf, Klassifikatoren mit großen Datenmengen zu trainieren, überhaupt vorhanden?

2 ANFORDERUNGEN AN TRAININGSDATENSÄTZE

Methoden der künstlichen Intelligenz wie neuronale Netze werden mittels Beispieldaten trainiert und können anschließend Muster automatisch wiedererkennen. Doch ihre Fähigkeit, aus Daten zu lernen, ist durch die **Vielfalt**,

das **Volumen** und die **Verlässlichkeit** der Informationen, die zum Trainieren verwendet werden, begrenzt. Wann ein Datensatz diese Anforderungen erfüllt, lässt sich nicht allgemeingültig beantworten, da ihre notwendige Ausprägung von der Sensortechnologie, der Verteilung der Daten, dem Anwendungsfall und dem verwendeten Lernverfahren bestimmt wird. [RuN12, Seite 52 f], [BaM21, Seite 198 f, 211], [NSD18, Seite 17 f], [Alp19, Seite 90 ff], [Lan01]

Unter **Vielfalt** ist die Vielzahl verschiedener Merkmale von *Probanden* und *Aktivitätsbeschreibungen* zu verstehen. Für einen Trainingsdatensatz ist eine *starke Variation von inhärenten Merkmalen der Probanden* wie beispielsweise das Alter, Geschlecht, die Körpergröße, -masse, Vorerkrankungen, Verletzungen, Händigkeit oder die Erfahrung im Ausüben der Aktivitäten relevant. Je nach Merkmalsausprägung unterscheiden sich die Bewegungen der Probanden selbst bei Ausführung der gleichen Aktivität. Beispielsweise bedingt das Alter eines Probanden die motorischen Fähigkeiten Kraft, Ausdauer, Geschicklichkeit, Schnelligkeit, Beweglichkeit und Koordination, welche die Bewegungsabläufe beeinflussen. Nur durch eine starke Variation von inhärenten Merkmalen kann sichergestellt werden, dass trainierte Klassifikatoren weniger voreingenommen und auf neue Probanden anwendbar sind. [CCF13] Dabei stellt die Vereinbarkeit von Datenschutz und der Erfassung annotationsrelevanter Daten eine nicht zu unterschätzende Herausforderung dar. Für die manuelle Annotation der Sensordaten kommen vorwiegend Videos zum Einsatz, die neben personenspezifischen auch firmeninterne Informationen enthalten können.

Die *Aktivitätsbeschreibung* ist abhängig von der Komplexität des Anwendungsfalls und der angestrebten Analysetiefe zu wählen. Sollen menschliche Bewegungen beispielsweise der Person-zur-Ware- bzw. Ware-zur-Person-Kommissionierung zugeordnet oder sollen Handhabungen auf der Granularitätsebene der Arbeitsablauf-Zeitanalyse (MTM) erkannt werden? Je granularer die Bewegungen klassifiziert werden sollen, desto granularer müssen die Aktivitäten im Trainingsdatensatz gelabelt sein. Schlussfolgernd ermöglicht eine detaillierte Vielfalt von *Probanden* und *Aktivitäten* eine präzise Generalisierung eines Klassifikators. Bei jeder Einführung einer neuen Technologie bzw. eines neuen Prozesses, der mit veränderten Bewegungsabläufen einhergeht, müssen neue Daten erhoben und dem Trainingsdatensatz hinzugefügt werden. In einer komplexen Umgebung wie der Intralogistik mit zahlreichen Systemtypen, Hilfsmitteln sowie Verfahren zur Kommissioniererführung, ist eine nahezu endlose Anzahl an Merkmalen möglich.

Das notwendige **Volumen**, welches den Umfang an Testdaten angibt, wird von der Vielfalt des

Anwendungsfalls, also der Dimensionalität, bestimmt. Bei einer geringen Vielfalt und zugleich einem geringen Volumen, können nur für einfache Hypothesenrepräsentationen Aussagen getroffen werden. Die Notwendigkeit für ein hohes Datenvolumen ist durch das nicht-deterministische Verhalten menschlicher Bewegungen bestimmt. Die gleiche Aktivität kann von einem Probanden bei mehrfacher Wiederholung nie exakt gleich ausgeführt werden. Hieraus resultiert, dass Trainingsdatensätze mit einer vergleichsweise geringen Vielfalt und einfachem Anwendungsfall durchschnittlich bereits über ein Volumen von zehn Stunden verfügen [NRM20]. Bei zunehmender Variation der inhärenten Merkmale von Probanden und Granularität der Aktivitätsbeschreibung steigt das notwendige Datenvolumen. Hieraus lässt sich ein **Dilemma** zwischen der Zeitdatenermittlung und dem Einsatz von HAR erkennen: Der zeitliche Aufwand für die Aufnahme einer möglichst großen Beispielmenge steht im Widerspruch zur angestrebten Aufwandssenkung der Zeitdatenerfassung durch HAR. Statt in aufwendigen Verfahren das notwendige Datenvolumen zu erfassen, bedarf es einem neuen Vorgehen, dass sich dem Dilemma zwischen der Zeitdatenermittlung und HAR annimmt. [RuN12, Seite 52 f], [BaM21, Seite 198 f, 211], [Lan01]

Das gesamte Datenvolumen muss einen gewissen Grad an **Verlässlichkeit** aufweisen. Die Verlässlichkeit oder auch Korrektheit ist abhängig von den manuell gelabelten Daten. Diese aufwendigen Prozessschritte, bestehend aus Annotation und Revision, bauen auf der Datenaufnahme auf. Mit der semi-automatischen Annotation wurde bereits ein Verfahren entwickelt, dass die Aufwände von Annotation und Revision signifikant reduziert [RRH18], [RSH18], [RRN20], [AAR21]. Schlussfolgernd bleibt ausschließlich die Datenaufnahme als zeitaufwendiger Faktor übrig.

3 STATE OF THE ART - CYBER-PHYSISCHER ZWILLING

Cyber-Physische Systeme (CPS) sind im Zuge der Digitalisierung entstanden und bilden die Basis für die erfolgreiche Umsetzung von Industrie 4.0-Ansätzen. Dabei können Objekte, Geräte, Gebäude, Verkehrsmittel, aber auch Produktionsanlagen, Logistikkomponenten etc. über das Internet kommunizieren und Internetdienste nutzen. Die Umwelt wird mit entsprechender Sensorik erfasst. Diese Daten werden ausgewertet und gespeichert. Die Einwirkung der CPS auf die physische Welt geschieht mittels Aktoren / Menschen, die über Mensch-Maschine-Schnittstellen mit diesen CPS kommunizieren. Mit Hilfe der CPS kann sich die sogenannte intelligente Fabrik (engl. Smart Factory) dezentral selbst und echtzeitnah organisieren. Echtzeitfähige Daten ermöglichen die Verschmelzung der realen Welt mit der virtuellen Welt. Folgende Beispiele basieren auf CPS: Smart Home, Smart Building und Smart Mobility. [Bau17]

Als ein Teil des CPS stellt der cyber-physische Zwilling eine Neuentwicklung und eine Schlüsseltechnologie dar. Das Konzept des cyber-physischen Zwillings setzt sich aus dem cyber Zwilling und dem physischen Zwilling zusammen, die über den gesamten Lebenszyklus miteinander interagieren. Bei dem physischen Zwilling handelt es sich um die reale Komponente, die um zusätzliche Sensoren und Aktoren erweitert wird. Somit können physische Umwelteinflüsse wahrgenommen werden. Sie dient in erster Linie dazu serviceorientierte, zuverlässige und nachhaltige Produkte zu erzeugen. Mit Hilfe eines cyber-physischen Zwillings können maßgeschneiderte Echtzeitdaten erfasst und benutzerorientiert in Simulationen genutzt werden. Dies ermöglicht die Vorhersage der Zuverlässigkeit oder Nachhaltigkeit, gibt einen Eindruck über die zukünftige Leistung und ermöglicht die Anpassung des zukünftigen Verhaltens. Dabei werden unter gegenseitiger Beeinflussung die Komponenten des cyber-physischen Zwillings und des physischen Zwillings sowie die bidirektionale Kommunikation zwischen Cyber- und physischem Zwilling gleichzeitig und zielgerichtet entwickelt. Der cyber-physische Zwilling dient zur Analyse und Vorhersage, sowie der möglichen Verhaltensanpassungen des physischen Zwillings. [CzA20] Im Folgenden werden Beispiele zur simulationsbasierten Generierung von menschlichen Bewegungen vorgestellt. Dabei handelt es sich um einen cyber-physischen Zwilling, da eine realistische Darstellung der Bewegung in einer bereitgestellten Umgebung simuliert wird.

In [ChB16] werden reale Bewegungsdaten von Probanden, die auf einem Laufband mit unterschiedlichen Geschwindigkeitsstufen gehen bzw. laufen, mittels der Motion Capturing (MoCap) Software Vicon Nexus aufgenommen und vorverarbeitet. Diese Aufnahmen liefern die notwendigen Informationen, um 3D-Skelettmodelle für die Probanden zu erstellen. Die Open-Source-Software Blender wird für die Steuerung von 3D-Charakteren durch die realen MoCap-Daten verwendet. Das Open-Source-Paket MakeHuman ermöglicht die Erstellung von Avataren. Als Avatare werden in Simulationen dargestellte Personen bezeichnet, die Stellvertreter einer echten Person in der physischen Welt sind [Tüm07]. Das Ziel ist die Erzeugung von Gangsequenzen mit Berücksichtigung der kontrollierbaren Störfaktoren. Kontrollierbare Störfaktoren werden zum Beispiel als Variationen in Kleidung, Beleuchtung, Pose und Umgebung beschrieben.

In [ZBF10] wird eine Software-Pipeline zur Generierung von 3D-Animationen unter Verwendung der MoCap-Daten und menschlichen Modellen (engl. human shape models) vorgeschlagen. Die Pipeline integriert die zwei Softwaretools Maya und MotionBuilder. Hier wird das Problem der Skelett-Inkompatibilität zwischen den MoCap-Daten, menschlichen Modellen und der Animationssoftware behandelt. Diese Inkompatibilität der

Skelettmodelle führt dazu, dass synthetisch erzeugte Animationen unrealistisch ausschauen.

In [LGC20] wird ein Simulationsframework für das Trainieren und Validieren von humanzentrierten Szenarien im Rahmen von intelligenten Verkehrssystemen präsentiert. Dieses Simulationsframework wurde entwickelt, da selten vorkommende menschliche Aktivitäten in sensorbasierten Datensätzen unterrepräsentiert sind. Das führt dazu, dass hochmoderne Algorithmen zur Schätzung der menschlichen Pose solche Aktivitäten nicht erkennen. Die im Rahmen von [LGC20] durchgeführten Laborversuche zeigen, dass ein Rekurrentes Neuronales Netz nur mit simulierten Daten basierend auf MoCap-Daten und 3D-Avataren trainierbar ist und eine nahezu perfekte Leistung bei der Klassifizierung dieser menschlichen Aktivitäten auf realen Daten erreicht wird.

Im Rahmen dieses Beitrags wird eine ähnliche Herangehensweise wie in [ChB16], [LGC20] und [ZBF10] zur Generierung von synthetischen Bewegungsdaten unter Verwendung von MoCap-Daten als Input verfolgt. Dabei ist der Motivationsgrund analog zu [LGC20].

4 FRAMEWORK – DATENGENERIERUNG MENSCHLICHER BEWEGUNGEN

In Abbildung 2 sind die Schritte des Frameworks dargestellt. Diese Darstellung wird im weiteren Verlauf des Beitrags schrittweise vorgestellt. Dabei wird der eingerahmte Bereich betrachtet, welcher die synthetische Datengenerierung des Beitrags darstellt.

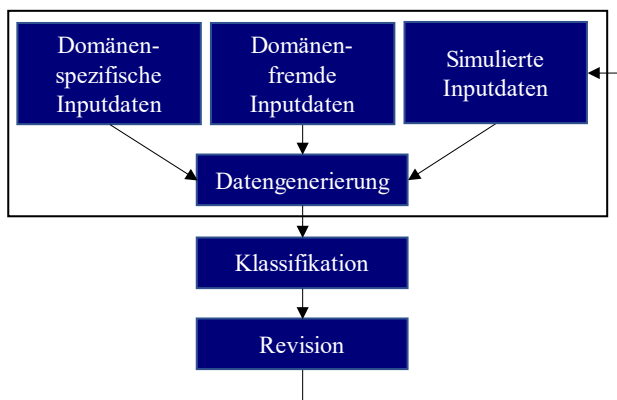


Abbildung 1: Framework zur Datengenerierung menschlicher Bewegungen und dessen Einordnung in die Trainingsdatensatzerstellung

Die Rückführung simulierter Daten als erneuten Dateninput führt zu einem Kreislauf-Effekt. Der Einsatz von Inputdaten aus domänenfremden Bereichen ist im Rahmen des Beitrags nicht notwendig, wird aber der Vollständigkeit halber mit in die Abbildung aufgenommen.

4.1 INPUTDATEN

Die Basis für das Framework bilden Daten menschlicher Bewegungen, welche aus verschiedenen Quellen herangezogen werden können. Dabei lassen sich grundsätzlich zwei Datenquellen unterscheiden. Zum einen können Sensordaten von Aufnahmen aus der physischen Welt stammen, die in einer realen Umgebung wie beispielsweise dem Lager oder aus realitätsnahen Laborumgebungen aufgenommen wurden. Als Sensoren eignen sich unter anderem inertielle Messeinheiten (IMUs), Tiefensensor oder ein optisches MoCap-System. Zum anderen können Daten aus der virtuellen Welt, also simulierte Daten, als Basis für das Framework dienen.

Wie lassen sich die notwendigen Daten beschaffen? Die aufwandsärmste Form der Datenbeschaffung liegt in der Verwendung vorhandener domänenspezifischer oder domänenfremder Datensätze. Unter domänenspezifisch werden Datensätze verstanden, die im logistischen Kontext aufgenommen worden sind. Entsprechend werden Datensätze als domänenfremd bezeichnet, die nicht im Kontext der Logistik aufgenommen sind. Hierbei müssen neben den verwendeten Sensoren ebenso die aufgezeichneten Bewegungen und annotierten Aktivitäten berücksichtigt werden. Bei der Verwendung domänenfremder Datensätze ist die Übereinstimmung der Aktivitäten zwischen dem Trainingsdatensatz und dem zu klassifizierenden Szenario zu bestimmen. Bei nur teilweisen Übereinstimmungen können einzelne Aufnahmen aus einem Datensatz verwendet werden. Je genauer die Übereinstimmung zwischen den Aktivitäten ist, desto realitätsnäher lassen sich in den nachfolgenden Schritten neue Bewegungen generieren. Als Alternative zu der Nutzung vorhandener Datensätze besteht die Möglichkeit, eigene Aufnahmen durchzuführen. Hierbei wäre die Übereinstimmung der Aktivitäten zwar am größten, gleichzeitig würde die eigene Datenaufnahme mit einem erhöhten Aufwand einhergehen. Im Vergleich mit dem klassischen Vorgehen der Datensatzerstellung würde jedoch nur ein Bruchteil von Daten benötigt. Daher wäre der Einsatz dieses Frameworks in Kombination mit eigenen Aufnahmen weiterhin deutlich aufwandsärmer als das klassische Vorgehen.

In diesem Beitrag wird eine Machbarkeitsstudie beispielhaft an MoCap-Daten durchgeführt. Eine Übertragung auf andere Sensordaten ist denkbar. Die notwendigen MoCap-Inputdaten müssen für die Abbildung intralogistischer Tätigkeiten Greif-, Handhabungs- und Beinbewegungen beinhalten. Eine Analyse von über 62 HAR-Datensätzen liefert 15 frei zugängliche Datensätze, die über MoCap-Daten verfügen [NRM20]. Die Aktivitäten von vier dieser 15 Datensätze umfassen unter anderem Sportaktivitäten, Tanzen oder Kochen und haben daher keine Relevanz für die Logistik. Neun Datensätze passen teilweise, da sie Beinbewegungen, wie Gehen, Stehen oder Sitzen

beinhalten. Mit AndyData-lab-one [MMI19], [MMA19] und LARa [NRA20], [NRM20] fokussieren lediglich zwei Datensätze den Fachbereich Intralogistik. Im Vergleich zu AndyData-lab-one umfasst LARa deutlich komplexere Handhabungsbewegungen. LARa beinhaltet Bewegungen aus Kommissionier- und Verpackungsprozessen, die unter realitätsnahen Bedingungen in einer Laborumgebung aufgenommen wurden [NRM20]. Schlussfolgernd dienen die MoCap-Daten von LARa als ausreichenden Input für die nachfolgenden Schritte.

4.2 DATENVORVERARBEITUNG

Eine simulationsbasierte Methodik für die synthetische Datenerweiterung bzw. -generierung setzt die Vorverarbeitung der Inputdaten voraus. Die Auswahl der Simulationssoftware bei der Umsetzung kann je nach Bedarf variieren.

Im Rahmen dieses Beitrags wird folgende Herangehensweise für die synthetische Datenerweiterung mittels simulationsbasierter Methodiken verfolgt: Aus dem MoCap-Datensatz LARa, bestehend aus insgesamt 379 Zwei-Minuten-Aufnahmen von 14 Probanden und dessen Erweiterung um zwei Probanden, werden exemplarisch Aufnahmen gewählt und in eine Simulationsumgebung integriert. Das Framework kann analog auf den gesamten LARa-Datensatz angewendet werden. In Abbildung 3 sind die Schritte der Datenvorverarbeitung abgebildet.

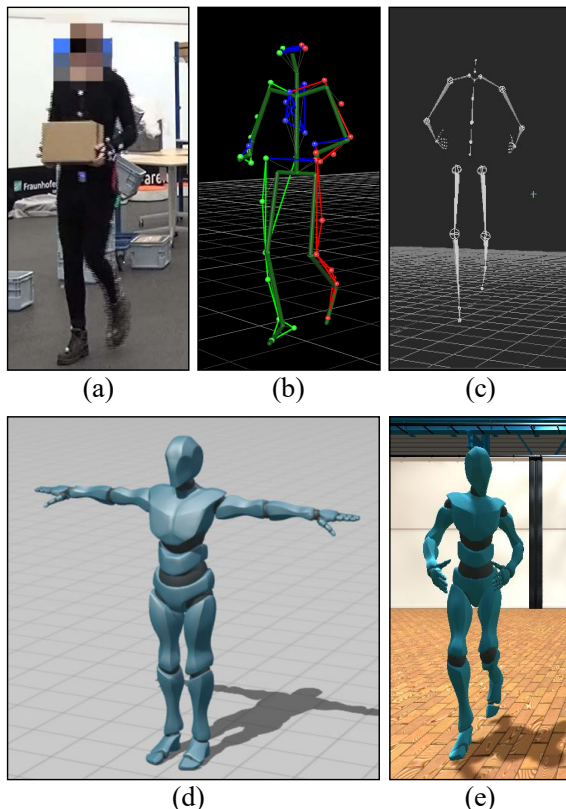


Abbildung 2: Pipeline für die Datenvorverarbeitung (in Anlehnung an [ChB16])

In Abbildung 3 (a) ist der Proband in der physischen Welt mit dem MoCap-Anzug und die darauf angebrachten Marker zu sehen. Dabei ist ein LogistikszENARIO in der Forschungshalle des Lehrstuhls für Förder- und Lagerwesen nachgebaut worden, um die Daten über die Aktivitäten eines Logistikmitarbeiters erfassen zu können. Die Daten wurden in der Forschungshalle mit dem MoCap-System erhoben. Die 3D-Darstellung in Vicon Nexus ist in Abbildung 3 (b) abgebildet. Die geometrischen Informationen über die Positionen der Marker werden in c3d-Dateien gespeichert. Um die MoCap-Daten in eine Simulationsumgebung integrieren zu können, müssen die c3d-Dateien konvertiert werden. Die c3d-Dateien beinhalten die Positionen der an den Probanden angebrachten Marker. Diese Positionsdaten werden in die Software MotionBuilder von AutoDesk importiert. Dafür wird die T-Pose, auch Referenzpose genannt, des Probanden aus dem MoCap-System importiert, um eine Actor-Vorlage zu erstellen. Hierzu werden die Markerpunkte einem Actor zugeordnet und anschließend als fbx-Datei gespeichert. Diese Konvertierung dient dazu, aus den c3d-Dateien ein Format zu gewinnen, das in eine Simulationsumgebung wie Unity importiert werden kann. Gängige Animationsformate sind u. a. fbx und bvh. In Abbildung 3 (c) ist das geschätzte Skelettmodell aus MotionBuilder abgebildet. Das Skelettmodell kann mit einem aus Adobe Mixamo heruntergeladenen 3D-Animationscharakter verknüpft werden. In Abbildung 3 (d) ist der 3D-Animationscharakter in der T-Pose abgebildet. Nach der Konvertierung von c3d zu fbx wird die fbx-Datei in Unity importiert, da der cyber-physische Zwilling der Forschungshalle in Unity erstellt ist. Somit werden die in MotionBuilder animierten Charaktere in die cyber-physische Forschungshalle in Unity integriert. In Abbildung 3 (e) ist der von den MoCap-Daten gesteuerte 3D-Animationscharakter in Unity zu sehen.

4.3 VERÄNDERUNG MENSCHLICHER BEWEGUNGEN

An dieser Stelle wird die Veränderung der vorhandenen MoCap-Daten in einer Simulationsumgebung vorgeschlagen. Die Aufnahmen zeigen, dass die MoCap-Daten sowohl auf einen männlichen 3D-Charakter als auch auf einen weiblichen 3D-Charakter in einer Simulationsumgebung übertragen werden können. Die zweite und dritte Spalte in Abbildung 4 stellen die Umsetzung in Unity dar. Hier sind zwei Avatare männlich und weiblich abgebildet, die mit den gleichen MoCap-Daten aus der physischen Welt (Spalte 1 in Abbildung 4) gesteuert werden. Mittels des Klassifikators kann überprüft werden, ob die Auswahl der 3D-Charakteren mit unterschiedlichen Ausprägungen sich auf die Güte des Klassifikators auswirken. Der Klassifikator wurde für die semi-automatische Annotation entwickelt [AAR21].

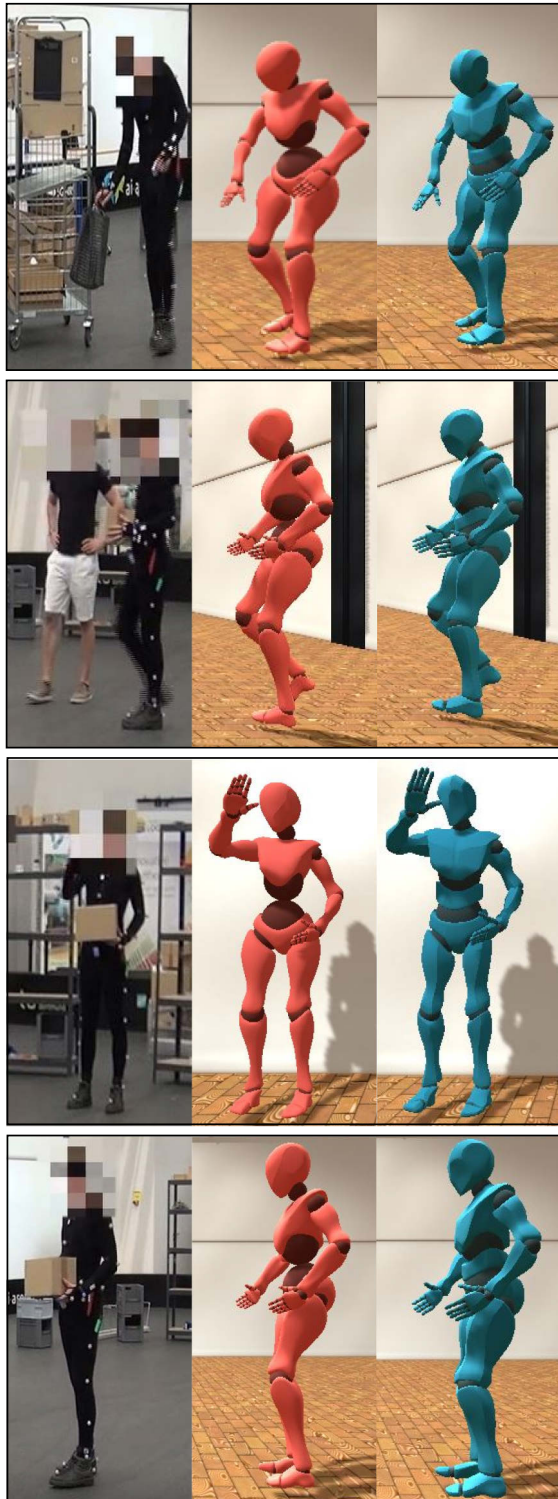


Abbildung 4: *Physische Welt vs. cyber-physischer Zwilling*

Zusätzlich können die menschlichen Bewegungen in der Simulationsumgebung erweitert werden. In Vicon Nexus wurde pro Aufnahme ein Proband in der physischen Welt aufgenommen. In der Simulationsumgebung kann mehr als ein Avatar gleichzeitig animiert werden (siehe die Abbildung 5). Für die Animation wurden je Avatar zwei MoCap-Aufnahmen eines Probanden verwendet.

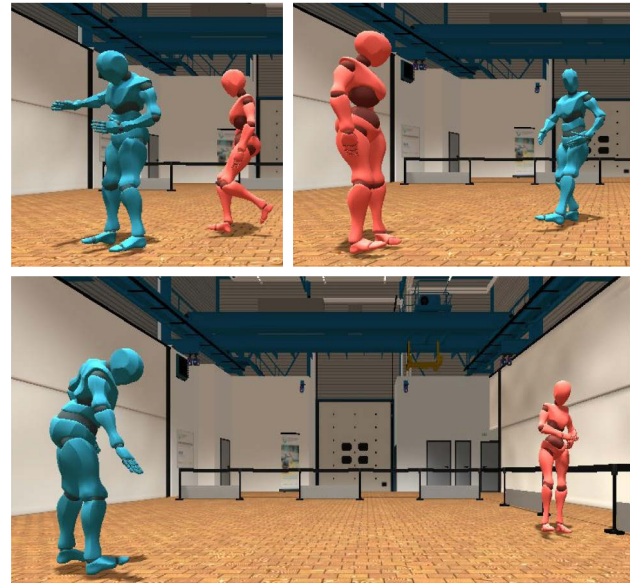


Abbildung 5: *Visualisierung generierter Aktivitäten und Erweiterung der physischen Welt im cyber-physischen Zwilling*

Eine weitere Möglichkeit der Generierung von menschlichen Bewegungen bietet der Animator Controller in Unity. Mit einem Animator Controller können eine Reihe von Animationsclips mit den entsprechenden Animationsübergängen erstellt und verwaltet werden. Die Animationsdateien können in MotionBuilder nach bestimmten Aktivitäten wie Laufen, Stehen, etc. aufgesplittet werden. Die aufgesplitteten Animationsdateien können in Unity mittels dem Animator Controller zu einer neuen Bewegungsabfolge verknüpft werden.

4.4 DATENUMWANDELUNG UND -EXPORT

Sowohl MotionBuilder als auch Unity bieten keine automatische Möglichkeit zur Datenrückführung in den Klassifikator. Aus diesem Grund erfordert die Konvertierung der exportierten fbx-Dateien in csv-Dateien ein eigenständig programmiertes Skript. Das Skript kann entweder über die Python Schnittstelle in MotionBuilder oder über die C# Schnittstelle in Unity programmiert werden. Im Rahmen dieses Beitrags wurden frei zugängliche Python-Skripte zur Datenkonvertierung in MotionBuilder getestet. Kein Python-Skript konvertierte die fbx-Dateien vollständig richtig in csv-Dateien. Aus diesem Grund muss für das vorgestellte Framework ein neues Skript erstellt werden.

4.5 ANNOTATION UND REVISION

Nachdem die generierten Daten exportiert wurden, findet die manuelle Annotation und Revision statt. Diese beiden Schritte sind unabhängig vom zuvor betrachteten Framework. Alternativ zur klassisch manuellen Annotation kann die semi-automatische Annotation nach Reining und Moya [RRH18] verwendet werden.

Die revidierten Daten werden als Inputdaten zur Generierung neuer Daten zurückgeführt (siehe Abbildung 2). Dadurch schließt sich der Kreislauf der in diesem Beitrag vorgestellten Trainingsdatengenerierung.

5 ZUSAMMENFASSUNG UND AUSBLICK

Bei dem beschriebenen Framework handelt es sich um einen cyber-physischen Zwilling, mit dem eine realistische Darstellung der Bewegung in einer bereitgestellten Umgebung simuliert wird. Basierend darauf wurden menschliche Bewegungen generiert. Die virtuell erzeugten und klassifizierten Daten dienen als Input für die Generierung neuer Daten, wodurch der Kreislauf der Datengenerierung geschlossen wird. Mit der Aufwandsreduktion der Datensatzerstellung liefert das Framework einen notwendigen Baustein einer neuen zeitwirtschaftlichen Methode, welche die Verbesserungspotenziale klassischer Methoden hebt.

Aus dem vorgestellten Framework ergeben sich folgende Ansätze für die Grundlagenforschung:

- In zukünftigen Forschungsbeiträgen wird der Nutzen des Datenkreislaufs überprüft. Sind Bewegungen, welche aus generierten Daten generiert werden, ausreichend realitätsnah, um Bewegungen aus einem realen Lager zu erkennen?
- Zusätzlich zu Simulationen als Werkzeug zum Ändern von menschlichen Bewegungen können weitere Methoden untersucht werden. Sowohl existierende Methoden des maschinellen Lernens als auch neu entwickelte Methoden stehen dabei im Fokus.
- Sind generierte Daten ausreichend, um ein für die Aktivitätsklassifikation von Realdaten geeignetes neuronales Netz zu trainieren? Diese Frage kann durch einen Vergleich von drei neuronalen Netzen beantwortet werden, welche mit verschiedenen Datensätzen trainiert wurden (1. LARa, 2. generierte Daten, 3. Kombination aus LARa und generierten Daten).
- Die Tatsache, dass für diesen Denkansatz mehr als eine Umsetzungsmöglichkeit existiert, weist eine gewisse Relevanz auf. Beispielsweise können andere Softwareprodukte (u. a. Blender) oder andere Sensoren (u. a. IMUs, RGB-Kameras) eingesetzt werden. Neben dem Datensatz LARa kann das Framework mit weiteren Datensätzen als Inputdaten erweitert werden. Es ist die Frage zu klären, in welchem Maße aus domänenfremden Datensätzen generierten Daten für die Erkennung von logistischen Aktivitäten verwendet werden können. Zur Beantwortung dieser Frage kann der Ansatz von Kirchhof [KSR21] herangezogen werden.

- Im aktuellen Framework werden ausschließlich die Bewegungen des Menschen in der Simulationsumgebung überführt. Neben dem Menschen könnten weitere Datenströme von Hilfsmitteln oder Orten verwendet werden. Unter dem sogenannten Kontext fallen z. B. die Position von Verpackungstischen, Kommissionierwagen, Regalen, Handscannern oder Computern.

DANKSAGUNG

Die Arbeit an dieser Publikation wurde von der Deutschen Forschungsgemeinschaft (DFG) im Rahmen des Projekts HO2463/14-2 „Transfer Learning for Human Activity Recognition in Logistics“ unterstützt.

LITERATUR

- [AAR21] Avsar, Hülya; Altermann, Erik; Reining, Christopher; Rueda, Fernando Moya; Fink, Gernot A.; ten Hompel, Michael: *Benchmarking Annotation Procedures for Multi-channel Time Series HAR Dataset*. In: IEEE International Conference on Pervasive Computing and Communications (PerCom). Kassel, 2021. — DOI 10.1109/PerComWorkshops51409.2021.9431062, S. 6
- [Alp19] Alpaydin, Ethem: *Maschinelles Lernen*, De Gruyter Studium. 2. Auflage. Berlin; Boston: De Gruyter Oldenbourg, 2019. — ISBN 978-3-11-061788-7
- [BaM21] Barton, T.; Müller, C. (Hrsg.): *Künstliche Intelligenz in der Anwendung: rechtliche Aspekte, Anwendungspotenziale und Einsatzszenarien, Angewandte Wirtschaftsinformatik*. Wiesbaden: Springer Vierweg, 2021. — DOI 10.1007/978-3-658-30936-7. — ISBN 978-3-658-30936-7
- [Bau17] Bauernhansl, Thomas: *Die Vierte Industrielle Revolution – Der Weg in ein wertschaffendes Produktionsparadigma*. In: Vogel-Heuser, B.; Bauernhansl, T.; ten Hompel, M. (Hrsg.): *Handbuch Industrie 4.0 Bd.4*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2017. — DOI 10.1007/978-3-662-53254-6_1. — ISBN 978-3-662-53253-9, S. 1–31
- [Bin10] Binner, Hartmut F.: *Handbuch der prozessorientierten Arbeitsorganisation: Methoden und Werkzeuge zur Umsetzung, REFA-Fachbuchreihe Unternehmensentwicklung*. 4. Aufl. München: Hanser, 2010. — ISBN 978-3-446-42641-2.
- [CCF13] Chaquet, Jose M.; Carmona, Enrique J.; Fernández-Caballero, Antonio: *A survey of*

video datasets for human action and activity recognition. In: Computer Vision and Image Understanding Bd. 117 (2013), Nr. 6, S. 633–659. — DOI 10.1016/j.cviu.2013.01.013

- [ChB16] Charalambous, Christoforos C.; Bharath, Anil A.: *A data augmentation methodology for training machine/deep learning gait recognition algorithms*. In: arXiv:1610.07570 [cs] (2016)
- [Cza20] Czwick, Cordula; Anderl, Reiner: *Cyber-physical twins - definition, conception and benefit*. In: Procedia CIRP, 2020, S. 584–588. — DOI 10.1016/j.procir.2020.01.070.
- [FHA16] Feldhorst, Sascha; ten Hompel, Michael; Aniol, Sandra: *Human Activity Recognition in der Kommissionierung – Charakterisierung des Kommissionierprozesses als Ausgangsbasis für die Methodenentwicklung*. Logistics Journal: Proceedings, Vol. 2016.
- [FMH16] Feldhorst, Sascha; Masoudenijad, Mojtaba; ten Hompel, Michael; Fink, Gernot A.: *Motion Classification for Analyzing the Order Picking Process using Mobile Sensors - General Concepts, Case Studies and Empirical Evaluation*. In: SCITEPRESS - Science and Technology Publications, 2016. — DOI 10.5220/0005828407060713.
- [KSH21] Kirchhof, Michael; Schmid, Lena; Reining, Christopher; Hompel, Michael ten; Pauly, Markus: *pRSL: Interpretable Multi-label Stacking by Learning Probabilistic Rules*. In: arXiv:2105.13850 [cs, stat] (2021)
- [KSR21] Kirchhof, Michael; Schmid, Lena; Reining, Christopher; ten Hompel, Michael; Pauly, Markus: *Chances of Interpretable Transfer Learning for Human Activity Recognition in Warehousing*. In International Conference on Computational Logistics (ICCL2021) — Status: akzeptiert.
- [Lan01] Laney, Doug: *3-D Data Management: Controlling Data Volume, Velocity and Variety*. In: Application Delivery Strategies by META Group Inc. Bd. 949 (2001)
- [LDL16] Lotter, Bruno; Deuse, Jochen; Lotter, Edwin: *Die Primäre Produktion: Ein praktischer Leitfaden zur verlustfreien Wertschöpfung*, VDI-Buch. 1. Aufl. 2016. Berlin s.l: Springer Berlin, 2016. — ISBN 978-3-662-53211-9.
- [LGC20] Ludl, Dennis; Gulde, Thomas; Curio, Cristobal: *Enhancing Data-Driven Algorithms for Human Pose Estimation and Action Recognition Through Simulation*. In: IEEE Transactions on Intelligent Transportation Systems Bd. 21 (2020), Nr. 9, S. 3990–3999. — DOI 10.1109/TITS.2020.2988504
- [MMA19] Maurice, Pauline; Malaisé, Adrien; Amiot, Clélie; Paris, Nicolas; Richard, Guy-Junior; Rochel, Olivier; Ivaldi, Serena: *Human movement and ergonomics: An industry-oriented dataset for collaborative robotics*. In: The International Journal of Robotics Research Bd. 38 (2019), Nr. 14, S. 1529–1537. — DOI 10.1177/0278364919882089
- [MMI19] Maurice, Pauline; Malaisé, Adrien; Ivaldi, Serena; Rochel, Olivier; Amiot, Clélie; Paris, Nicolas; Richard, Guy-Junior; Fritzsche, Lars: *AndyData-lab-onePerson*. In: Zenodo, 2020. — DOI 10.5281/zenodo.3254403.
- [NRA20] Niemann, Friedrich; Reining, Christopher; Moya Rueda, Fernando; Altermann, Erik; Nair, Nilah Ravi; Steffens, Janine Anika; Fink, Gernot A.; ten Hompel, Michael: *Logistic Activity Recognition Challenge (LARA) – A Motion Capture and Inertial Measurement Dataset*. In: Zenodo, 2020. — DOI 10.5281/zenodo.3862782.
- [NRM20] Niemann, Friedrich; Reining, Christopher; Moya Rueda, Fernando; Nair, Nilah Ravi; Steffens, Janine Anika; Fink, Gernot A.; ten Hompel, Michael: *LARA: Creating a Dataset for Human Activity Recognition in Logistics Using Semantic Attributes*. In: Sensors, 2020, Nr. 15, S. 4083. — DOI 10.3390/s20154083.
- [NSD18] Ng, Annalyn; Soo, Kenneth; Delbrück, M. (Übers.): *Data Science - was ist das eigentlich?! Algorithmen des maschinellen Lernens verständlich erklärt*. Berlin [Heidelberg]: Springer, 2018. — ISBN 978-3-662-56775-3
- [Olb93] Olbrich, Ralf: *Aufbau einer Zeitwirtschaft: Beschreibung der Vorgehensweise*. Köln: Wirtschaftsverlag Bachem, 1993. — ISBN 978-3-89172-258-9.
- [RNM19] Reining, Christopher; Niemann, Friedrich; Moya Rueda, Fernando; Fink, Gernot A.; ten Hompel, Michael: *Human Activity Recognition for Production and Logistics—A Systematic Literature Review*. In: Information Bd. 10 (2019), Nr. 8, S. 245. — DOI 10.3390/info10080245
- [RRH18] Reining, Christopher; Rueda, Fernando; Moya; Hompel, Michael ten; Fink, Gernot A.:

Towards a Framework for Semi-Automated Annotation of Human Order Picking Activities Using Motion Capturing. 2018. — DOI 10.15439/2018F188, S. 817–821

- [RRN20] Reining, Christopher; Rueda, Fernando Moya; Niemann, Friedrich; Fink, Gernot A.; Hompel, Michael ten: *Annotation Performance for multi-channel time series HAR Dataset in Logistics*. In: 2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops). Austin, TX, USA: IEEE, 2020. — DOI 10.1109/PerComWorkshops48775.2020.9156170. S. 1–6
- [RSH18] Reining, Christopher; Schlangen, Michelle; Hissmann, Leon; ten Hompel, Michael; Moya, Fernando; Fink, Gernot A.: *Attribute Representation for Human Activity Recognition of Manual Order Picking Activities*. In: Proceedings of the 5th International Workshop on Sensor-based Activity Recognition and Interaction, *iWOAR '18*. New York, NY, USA: ACM, 2018. — DOI 10.1145/3266157.3266214.
- [RuN12] Russell, Stuart; Norvig, Peter: *Künstliche Intelligenz: ein moderner Ansatz, Always learning*. 3., aktualisierte Aufl. München: Pearson Studium, 2012. — ISBN 978-3-86894-098-5
- [Sat07] Sadowsky, Volker; ten Hompel, Michael: *Beitrag zur analytischen Leistungsermittlung von Kommissioniersystemen*, Logistik für die Praxis. Dortmund: Verl. Praxiswissen, 2007. — ISBN 978-3-89957-057-1.
- [Tüm07] Tümmler, Jörn: *Avatare in Echtzeitsimulationen*. Kassel: Kassel University Press, 2007. — ISBN 978-3-89958-291-8
- [ZBF10] Zhang, Xin; Biswas, David S.; Fan, Guoliang: *A software pipeline for 3D animation generation using mocap data and commercial shape models*. In: Proceedings of the ACM International Conference on Image and Video Retrieval - CIVR '10. Xi'an, China: ACM Press, 2010. — DOI 10.1145/1816041.1816092. — ISBN 978-1-4503-0117-6, S. 350

Hülya Avsar, M.Sc., Research Assistant at the Chair of Materials Handling and Warehousing, Technical University Dortmund. She studied Computer science.

Phone: +49 231 755-4073, Fax: +49 231 755-4768,
E-Mail: huelya.avsar@tu-dortmund.de

Friedrich Niemann, M.Sc., Research Assistant at the Chair of Materials Handling and Warehousing, Technical University Dortmund. He studied Logistics.

Phone: +49 231 755-4548, Fax: +49 231 755-4768,
E-Mail: friedrich.niemann@tu-dortmund.de

Christopher Reining, M.Sc., Research Assistant at the Chair of Materials Handling and Warehousing, Technical University Dortmund. He studied Logistics.

Phone: +49 231 755-3228, Fax: +49 231 755-4768,
E-Mail: christopher.reining@tu-dortmund.de

Prof. Dr. Dr. h. c. Michael ten Hompel, head of the Chair of Materials Handling and Warehousing, Technical University of Dortmund and Managing director of the Fraunhofer Institute for Material Flow and Logistics.

Phone: +49 231 9743-600, Fax: +49 231 755-4768,
E-Mail: michael.tenHompel@tu-dortmund.de

Address: Lehrstuhl für Förder- und Lagerwesen, TU Dortmund, Joseph-von-Fraunhofer-Str. 2-4, 44227 Dortmund, Germany