

## **Nutzung unterschiedlicher Datenformate zum Einsatz von maschinellem Lernen bei Analysen von Körperbewegungen**

Edgar SCHERSTJANOI<sup>1</sup>, Dennis WITTCHEN<sup>2</sup>, Martin SCHMAUDER<sup>1</sup>

<sup>1</sup> *Institut für Arbeitswissenschaft, Technische Universität Dresden  
Marschnerstraße 39, D-01307 Dresden*

<sup>2</sup> *Hochschule für Technik und Wirtschaft Dresden  
Friedrich-List-Platz 1, D-01069 Dresden*

**Kurzfassung:** Die automatisierte Analyse von manuellen Arbeitstätigkeiten benötigt Algorithmen, welche die komplexen Zusammenhänge des menschlichen Bewegungsapparates gezielt verarbeiten können. Methoden des maschinellen Lernens eignen sich für eine solche Modellierung, benötigen dazu aber eine geeignete Auswahl von Trainingsdaten. Verfügbare Datensätze bieten zwar eine vielfältige Auswahl verschiedener Tätigkeiten, jedoch ist die gemeinsame Nutzung aufgrund verschiedener Datenstrukturen, Technologien der Bewegungserfassung und heterogener Austauschformate erschwert. Dieser Beitrag stellt eine flexible Herangehensweise vor, bestehende Datensätze als Grundlage zur Modellierung zu nutzen.

**Schlüsselwörter:** Arbeitsbewegungen, maschinelles Lernen, Trainingsdaten, künstliche Intelligenz

### **1. Einleitung**

Die Kinematik von Arbeitsbewegungen spielt eine zentrale Rolle bei der Umsetzung von arbeitswissenschaftlichen Softwareanwendungen, wie bspw. Programme zur ergonomischen Bewertung von Arbeitsplätzen, Roboternavigation in kollaborativen Mensch-Maschine-Arbeitssystemen oder digitalen Menschmodellen. Die darin wesentlichen Klassifikations- oder Regressionsalgorithmen sind dazu in der Lage Körperhaltungen ergonomisch zu kategorisieren bzw. Bewegungstrajektorien von Körpersegmenten zu prognostizieren.

Immer häufiger finden in diesem Zusammenhang Techniken des maschinellen Lernens (ML) Einsatz, eine Schlüsseltechnologie zur Konstruktion von künstlich intelligenten (KI) Computersystemen. Gemeint sind Algorithmen, die aus einer gegebenen Menge von Trainingsdaten statistische Modelle erzeugen, unter deren Verwendung wiederum Zusammenhänge in Daten erkannt werden, die nicht konkreter Bestandteil der Ausgangsdaten waren. Der Einsatz von ML ist immer dann sinnvoll, wenn Prozesse zu komplex sind, um vollständig formalisiert werden zu können, aber eine ausreichende Datensammlung zur Verfügung steht (Döbel et al. 2018). Bei der Analyse von Körperhaltungen und Arbeitsbewegungen wird diese Komplexität deutlich, vor allem dann, wenn menschliche Beurteilungen der Körperbewegung eine Rolle spielen, wie es bei der Gestaltung von Arbeitsplätzen, Interaktion oder Handhabung von Arbeitsmitteln der Fall ist.

## **2. Bewegungsdaten als Trainingsdaten zum maschinellen Lernen**

Durch Technologien der digitalen Bewegungserfassung ist eine Möglichkeit geboten, Sammlungen von potentiellen Trainingsdaten zu erstellen. Prinzipiell wird zwischen optischen und inertialen Erfassungssystemen unterschieden (Menolotto et al. 2020). Inertiale Systeme verarbeiten den Datenstrom mehrerer am Körper angebrachter Sensoren (Accelerometer und Gyroskop), optische Systeme verwenden wiederum die Erkennung von Markern am Körper (marker-based Motion Capture) oder approximieren den Bewegungsablauf aus einem oder mehreren synchronisierten RGB- oder Tiefenbilddaten (markerless Motion Capture). Unabhängig von der Erfassungstechnologie wird üblicherweise im Ergebnis die Sequenz einer skelettalen Struktur abstrahiert, welche die Kinematik der Bewegung beinhaltet.

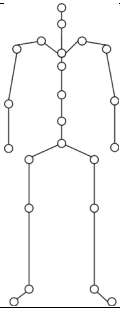
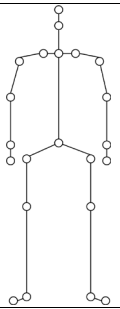
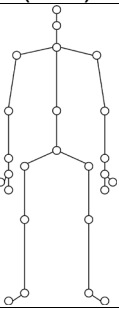
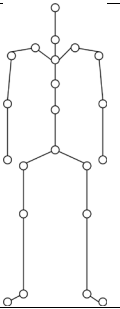
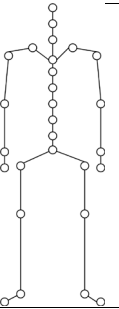
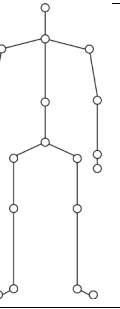
Eine Herausforderung bei der Entwicklung von ML-basierten Systemen besteht in der gezielten Auswahl von Daten, da der Inhalt den Umfang des modellierbaren Wissens bestimmt. In Bezug auf Körperbewegungen ist die Güte eines Datensatzes demzufolge wesentlich von der Auswahl der enthaltenen Aktivitäten, dem ProbandInnenkollektiv sowie der Dauer, Komplexität und Qualität des Bewegungsablaufes bestimmt. Oftmals werden für unterschiedliche Forschungsziele jeweils neue Daten aufgenommen, was mitunter hohe monetäre und zeitliche Aufwände erfordert. Die Gründe hierfür liegen vor allem in der Inkompatibilität oder fehlenden Interoperabilität bereits vorhandener, frei verfügbarer Datensätze.

In dieser Arbeit wurde untersucht, wie bestehende Datensätze kombiniert werden können, um sie für ML im Rahmen von arbeitswissenschaftlichen Untersuchungen nutzbar zu machen. Zu diesem Zweck musste einerseits eine Analyse bestehender Datensätze erfolgen, andererseits eine Methodik zur Konjugation entworfen werden. Beide Aspekte werden in den folgenden Abschnitten vorgestellt.

## **3. Analyse verfügbarer Datensätze mit menschlichen Körperbewegungen**

Eine systematische, semi-automatisierte Recherche ergab, dass mehr als 50 frei zugängliche Datensätze existieren, die sich prinzipiell für das Training einer KI für Bewegungsmodellierungen von Arbeitstätigkeiten eignen. Die Mehrheit dieser Datensätze beinhaltet Bewegungsabläufe und Gesten wie "Gehen", "Laufen" und "Reden". Zwei der Datensätze wurden gezielt für arbeitswissenschaftliche Zwecke erstellt: LARa - zur Detektion von manuellen Transporttätigkeiten in der Intralogistik (Niemann et al. 2020) und AndyData - zur ergonomischen Analyse von Montagearbeiten (Maurice et al. 2018). In Tabelle 1 ist exemplarisch ein Ausschnitt der im Rahmen der Recherche gefundenen und untersuchten Datensätze mit ausgewählten Vergleichskriterien aufgeführt.

**Tabelle 1:** Übersicht eines Ausschnittes der Datensätze, die durch eine semi-automatisierte Recherche identifiziert wurden und deren Eigenschaften in Bezug auf Inhalt, Format, Erfassungstechnologie und skelettale Struktur

	<b>AndyData</b> (Maurice et al. 2018)	<b>LARa</b> (Niemann et al. 2020)	<b>NTU RGBD</b> (Liu et al. 2019)	<b>Mirrored Human</b> (Fang et al. 2021)	<b>TUM Kitchen</b> (Tenorth et al. 2009)	<b>Kinder Gator</b> (Aloba et al. 2020)
<b>Tätigkeiten</b>	Montage	Transport	Alltägliche	u.A. Sport	Kochen	Gestiken
<b>Datenformat</b>	u.A. mvn	csv	txt	simpl	bvh	csv
<b>Erfassungs- technologie</b>	Inertial	Optisch, Marker- basiert	Tiefenbild kamera (ToF)	20 RGB Kameras	4 RGB Kameras	Tiefenbild kamera
<b>Skelettale Struktur</b>						

Es wird deutlich, dass mit den verschiedenen Erfassungstechnologien auch jeweils verschiedene skelettale Strukturen in diversen Datenformaten gegeben sind. MirroredHuman bietet bspw. eine detaillierte Darstellung der Wirbelsäule, wohingegen in NTU RGBD auch Daten zu Fingerspitzen und Daumen enthalten sind. Diese Heterogenität erschwert die interoperable Verwendung. Dennoch ist erkennbar, dass alle Skelette eine bestimmte Menge von Gelenkpunkten teilen: Fuß, Knie, Hüfte, Schulter, Ellenbogen, Hand und Kopf.

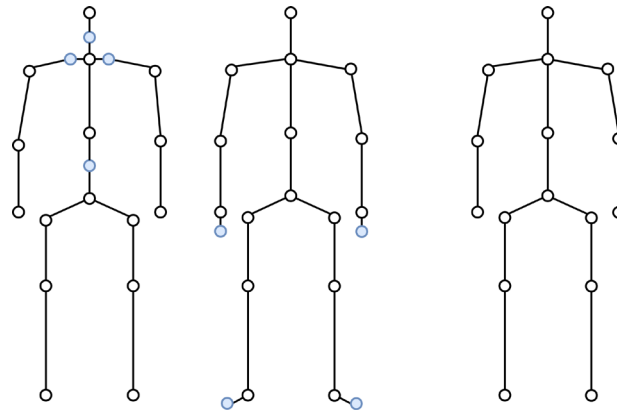
#### 4. Gemeinsame Nutzung verschiedener Datenformate

Das Zusammenführen von unterschiedlichen Datensätzen erhöht potentiell die Informationsvielfalt eines Trainingsdatensatzes und somit auch die Generalisierbarkeit des ML-Modells. Die Kombination von heterogenen Datensätzen ist dabei eine Herausforderung. Eine Möglichkeit besteht darin, durch Konvertierung aggregierte einheitliche Datensätze zu schaffen.

Demgegenüber steht der hier aufgeführte Ansatz, die inhaltliche Struktur eines Datensatzes semantisch zu beschreiben und diesen somit zur weiteren Verarbeitung verfügbar zu machen. Ausgangspunkt dafür, ist ein eindeutig definiertes Vokabular, mit dem die strukturellen und inhaltlichen Eigenschaften eines Datenformats beschrieben werden. Das umfasst Informationen zur Datenstruktur (z.B. tabellarischer oder hierarchischer Aufbau), Art der Information (z.B. Position oder Orientierung) zeitliche und räumliche Skalierungen, ggf. Segmentlängen sowie Anzahl von und Beziehungen zwischen Gelenkpunkten. Eine solche Semantik ermöglicht die einheitliche Interpretation.

Werden nur Gelenkpunkte verwendet, die als Teilmenge der ausgewählten skelettalen Strukturen vorliegen, ist die aggregierte Datenmenge ohne weitere

Verarbeitung für den Einsatz von ML nutzbar. In Abbildung 1 ist diese Neuordnung illustriert.



**Abbildung 1:** Schematische Darstellung, welche Gelenkpunkte bei der Aggregation enthalten bleiben (rechts), wenn die Teilmenge beider skelettaler Strukturen (links, Mitte) gebildet wird.

Es wird deutlich, dass die resultierende Skelettstruktur reduziert ist. Dennoch kann, je nach Einsatzzweck, die verbleibende Menge von Körpersegmenten als Trainingsdatensatz genutzt werden, bspw. um Laufgeschwindigkeiten, Armbewegungen oder Oberkörperdrehungen zu verarbeiten.

## 5. Schlussfolgerung und Ausblick

Es existiert eine Vielzahl an Datensätzen mit skelettalen Bewegungsdaten, die zur KI-Modellierung von Körperbewegungen genutzt werden und somit auch für arbeitswissenschaftliche Zwecke hilfreich sind. Die Heterogenität der Datenformate und skelettalen Strukturen stehen einer Datensatz-übergreifenden Verwendung für das Training einer KI im Weg. Durch die Verwendung eines deskriptiven Vokabulars zur Annotation wird hier die Möglichkeit vorgestellt, eine individuelle Auswahl von Datensätzen als Trainingsmenge für die Implementierung und Validierung von ML-Algorithmen zu nutzen. Die auf diese Weise durchgeführte Aggregation mehrerer Datensätze ermöglicht umfangreiche Modelle und spart Ressourcen, die andernfalls für umständliche Konvertierungen oder Aufnahme neuer Datensätze aufgewendet werden müssten.

Die Reduktion von Gelenkpunkten führt jedoch dazu, dass nach Aggregation auch der Einsatzzweck unter Umständen eingeschränkt ist. Wird bspw. auf die Daten der Position der Fußspitzen verzichtet, liefern ML-Algorithmen womöglich auch nur unzureichende Ergebnisse bei Ganganalysen. Um diesen Informationsverlust zu kompensieren, steht in Aussicht, generative Algorithmen einzusetzen, die dazu in der Lage sind, fehlende Daten anhand eines digitalen Menschmodells oder weiterer ML-Modelle zu reproduzieren. Eine solche Weiterentwicklung steigert die Flexibilität und ermöglicht eine KI-Modellierung mit einer umfangreicheren Auswahl an Datensätzen.

## 6. Literatur

- Aloba A, Flores G, Woodward J, Shaw A, Castonguay A, Cuba I, Dong Y, Jain E, Anthony L (2018) Kinder-Gator: The UF Kinect Database of Child and Adult Motion. In Eurographics (Short Papers), 13-16.
- Döbel I, Leis M, Vogelsang MM, Neustroev D, Petzka H, Rüping S, Welz J (2018) Maschinelles Lernen–Kompetenzen, Anwendungen und Forschungsbedarf. Sankt Augustin: Fraunhofer-Gesellschaft (IAS, IMW, Zentrale).
- Fang Q, Shuai Q, Dong J, Bao H, Zhou X (2021) Reconstructing 3D Human Pose by Watching Humans in the Mirror. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 12814-12823.
- Liu J, Shahroudy A, Perez M, Wang G, Duan L Y, Kot AC (2019) NTU RGB+ D 120: A large-scale benchmark for 3d human activity understanding. IEEE transactions on pattern analysis and machine intelligence, 42(10), 2684-2701.
- Maurice P, Malaisé A, Amiot C, Paris N, Richard GJ, Rochel O, Ivaldi S (2019) Human movement and ergonomics: An industry-oriented dataset for collaborative robotics. The International Journal of Robotics Research, 38(14), 1529-1537.
- Menolotto M, Komaris D S, Tedesco S, O'Flynn B, Walsh M (2020) Motion capture technology in industrial applications: A systematic review. Sensors, 20(19), 5687.
- Apparies RJ, Riniolo TC, Porges SW (1998) A psychophysiological investigation of the effects of driver longer-combination vehicles. Ergonomics 41:581-592.
- Niemann F, Reining C, Moya Rueda F, Nair NR, Steffens JA, Fink GA, Ten Hompel M (2020) LARa: Creating a Dataset for Human Activity Recognition in Logistics Using Semantic Attributes. Sensors, 20(15), 4083.
- Tenorth M, Bandouch J, Beetz M. (2009) The TUM kitchen data set of everyday manipulation activities for motion tracking and action recognition. In: 2009 IEEE 12th International Conference on Computer Vision Workshops, ICCV Workshops, 1089-1096.



Gesellschaft für  
Arbeitswissenschaft e.V.

## Technologie und Bildung in hybriden Arbeitswelten

68. Kongress der  
Gesellschaft für Arbeitswissenschaft

Otto-von-Guericke-Universität Magdeburg

Fraunhofer-Institut für Fabrikbetrieb und  
Fabrikautomatisierung IFF, Magdeburg

02. – 04. März 2022

---

## GfA-Press

---

**Bericht zum 68. Arbeitswissenschaftlichen Kongress vom 02. – 04. März 2022**

**Otto-von Guericke-Universität Magdeburg;**

**Fraunhofer-Institut für Fabrikbetrieb und -automatisierung IFF, Magdeburg**

Herausgegeben von der Gesellschaft für Arbeitswissenschaft e.V.

Sankt Augustin: GfA-Press, 2022

ISBN 978-3-936804-31-7

NE: Gesellschaft für Arbeitswissenschaft: Jahresdokumentation

Als Manuskript zusammengestellt. Diese Jahresdokumentation ist nur in der Geschäftsstelle (s. u.) erhältlich.

Alle Rechte vorbehalten.

© **GfA-Press, Sankt Augustin**

**Schriftleitung: Prof. Dr. Rolf Ellegast**

im Auftrag der Gesellschaft für Arbeitswissenschaft e.V.

Ohne ausdrückliche Genehmigung der Gesellschaft für Arbeitswissenschaft e.V. ist es nicht gestattet:

- den Kongressband oder Teile daraus in irgendeiner Form (durch Fotokopie, Mikrofilm oder ein anderes Verfahren) zu vervielfältigen,
- den Kongressband oder Teile daraus in Print- und/oder Nonprint-Medien (Webseiten, Blog, Social Media) zu verbreiten.

Die Verantwortung für die Inhalte der Beiträge tragen alleine die jeweiligen Verfasser; die GfA haftet nicht für die weitere Verwendung der darin enthaltenen Angaben.

### **Geschäftsstelle der GfA**

Simone John, Tel.: +49 (0)30 1300-13003

Alte Heerstraße 111, D-53757 Sankt Augustin

[info@gesellschaft-fuer-arbeitswissenschaft.de](mailto:info@gesellschaft-fuer-arbeitswissenschaft.de) · [www.gesellschaft-fuer-arbeitswissenschaft.de](http://www.gesellschaft-fuer-arbeitswissenschaft.de)

### **Screen design und Umsetzung**

© 2022 fröse multimedia, Frank Fröse

[office@internetkundenservice.de](mailto:office@internetkundenservice.de) · [www.internetkundenservice.de](http://www.internetkundenservice.de)