

Effiziente Lern- und Arbeitsprozesse in der Montage dank KI und Bildverarbeitung

Wissenssicherung und -transfer in der manuellen Montage

L. Häge, D. Karelina, C. Jauch

ZUSAMMENFASSUNG Die Produktion steht, wie viele Industriezweige in Deutschland, vor Herausforderungen. Dazu gehören der Fachkräftemangel, rasante Entwicklungen in Wertschöpfungssystemen, demografischer Wandel und die hohe Bereitschaft zum Jobwechsel. Somit geht Erfahrungswissen schneller verloren und effiziente Kompetenzvermittlung wird immer wichtiger. Die Forschung beschäftigt sich daher mit innovativen Strategien, um Erfahrungswissen erhalten und effizient weitergeben zu können.

STICHWÖRTER

Bildverarbeitung, Wissensmanagement, Künstliche Intelligenz

Knowledge retention and transfer in manual assembly

ABSTRACT Similar to other sectors, the manufacturing industry in Germany is faced by challenges. They include a shortage of skilled labor, rapid advancements in value creation systems, demographic shifts, and an increased willingness to change jobs. This means that expertise is lost more rapidly, making the effective transfer of skills a real necessity. Research is centering on innovative strategies to retain expertise and facilitate its efficient transmission.

1 Einleitung

Als Folge des demografischen Wandels macht sich in vielen Branchen ein steigender Fachkräftemangel bemerkbar. Laut einer Konjunkturumfrage des DIHK-(Deutscher Industrie- und Handelskammertag) haben mehr als die Hälfte der 23 000 befragten Unternehmen in Deutschland Schwierigkeiten, offene Stellen mit Fachpersonal zu besetzen. Zusätzlich kann ein Fachkräftemangel neben volkswirtschaftlichen Folgen auch die Innovations- und Wettbewerbsfähigkeit gefährden. Daher gilt es, Lösungen für diese Entwicklung zu finden. Als Gegenmaßnahme wurde neben der attraktiveren Gestaltung der Arbeitsbedingungen die Förderung von Aus- und Weiterbildung von bestehendem und zukünftigem Personal genannt [1]. Das Anlernen von Mitarbeitenden erfordert aber wiederum den Einsatz von Fachpersonal, das ohnehin schon Mehraufgaben übernimmt (Problem 1).

Auch das produzierende Gewerbe sowie die manuelle Montage leiden zunehmend unter der prozentual geringeren Anzahl an qualifizierten Erwerbstätigen und einer erhöhten Mitarbeitendenfluktuation [1, 2]. Dabei besteht die Gefahr, dass dem Unternehmen wertvolles personengebundenes Expertenwissen verloren geht. Diese Form von meist unwiderruflichem Informationsverlust kann gravierende Folgen für den Unternehmenserfolg mit sich bringen [3]. Durch sich wiederholende Arbeitsprozesse entsteht bei erfahrenen Fachleuten wertvolles (Prozess-)Wissen. Dieses Wissen kann für das Einlernen von neuen Arbeitern oder das Optimieren von Arbeitsprozessen weniger erfahrener Mitarbeitenden genutzt werden. Um für das Unternehmen personengebundenes Expertenwissen zugänglich und verwertbar zu machen,

ist eine effiziente Form von Wissensexternalisierung erforderlich (Problem 2).

Zur Erfassung von Experten- oder Erfahrungswissen werden derzeit personenfokussierte Ansätze (wie Wissenstransfers oder -gemeinschaften) oder rein software- beziehungsweise datengetriebene Ansätze (zum Beispiel Wissensdatenbanken, Wikis, Suchmaschinen) genutzt. Dies birgt drei wesentliche Nachteile: Erstens kann Erfahrungswissen explizit und implizit vorliegen. Selbst wenn Fachleute also bereit sind, ihr Wissen zu teilen, bedeutet dies nicht automatisch, dass sie alles verfügbare Wissen mitteilen können. Oft sind ihnen subtile Nuancen ihrer Arbeit nicht bewusst, die aber wiederum maßgeblich zum Erfolg des Unternehmens beitragen. Infolgedessen kann es passieren, dass unbewusste Feinheiten in den Prozessschritten wie etwa spezielle Handgriffe bei der manuellen Montage, die zu geringerem Ausschuss oder kürzeren Montagezeiten führen, bei der aktuellen Externalisierung von Erfahrungswissen (unwissentlich) vernachlässigt werden (Problem 2.1). Zweitens wird keine Bündelung des Erfahrungswissens vorgenommen. Marginale Unterschiede, welche bei derselben Fachkraft über mehrere Durchläufe hinweg auftreten können, sollten nicht als neuartige Ausführung oder Best Practice interpretiert werden (Problem 2.2). Schließlich sind drittens der Zeitaufwand und die damit einhergehende vernachlässigte Pflege von manuell erstelltem und abgelegtem Erfahrungswissen eine weitere Hürde (Problem 2.3).

Zu diesen Themen wird aufgrund der gegenwärtigen Relevanz aktiv geforscht. Ein Beispiel ist das Projekt „Wissensbasierte Produktionshilfe beruhend auf bild- und sensorbasierter Externalisierung von Erfahrungswissen“ (WisPE). Das Projektkonsortium setzt sich aus der Pumacy Technologies AG, dem Fraunhofer-

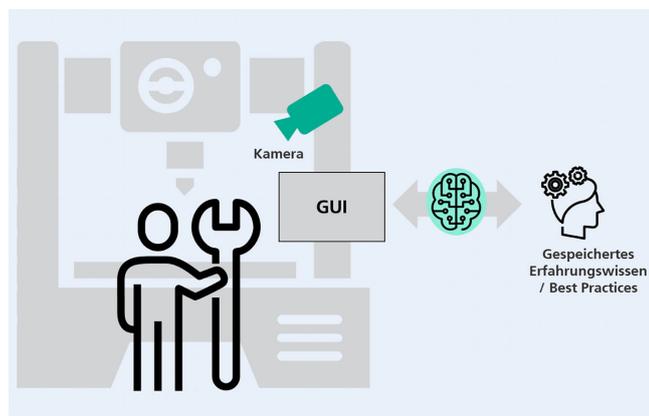


Bild 1 Konzept des Systems zur Erfassung und Weitergabe von Erfahrungswissen. Grafik: Fraunhofer IPA

Institut für Produktionstechnik und Automatisierung IPA sowie MaTec Gummiwerke und WS Wieländer+Schill zusammen. Das Projekt befindet sich in der zweiten Bearbeitungshälfte und wird Mitte des Jahres 2026 abgeschlossen. Ein Projektschwerpunkt des Fraunhofer IPA ist die Externalisierung von Erfahrungswissen mittels kamerabasierter Verfahren, der auch Fokus dieser Arbeit ist.

2 Grundlagen

Die Unterteilung von explizitem und implizitem Wissen und Lernen stammt ursprünglich aus der Sprachforschung zum Erlernen der Muttersprache im Gegensatz zu einer Fremdsprache. Während sprachliche Konstrukte und grammatikalische Regeln beim Erlernen der Muttersprache implizit (etwa durch Zuhören, Nachsprechen) vermittelt werden, werden beim Lernen einer Fremdsprache diese Regeln und Ausnahmefälle zusätzlich explizit gelehrt. Explizites Wissen zeichnet sich dadurch aus, dass es verbal wiedergegeben und kommuniziert werden kann, während implizites Wissen unterbewusst vorhanden ist. Der Träger von implizitem Wissen ist sich somit häufig nicht bewusst, dieses Wissen zu besitzen oder nicht in der Lage, es zu verbalisieren. Bei wiederholten Prozessen, die automatisch von der Hand gehen, wird auf implizites Wissen zurückgegriffen [4]. In der Kompetenzvermittlung wird auch dem Austausch von diesem impliziten Wissen von Erfahrenen/Experten hoher Wert beigemessen. Somit soll nicht nur übermittelt werden, „was“ getan werden soll, sondern auch „wie“ [5]. Implizites Wissen wird oft mit „Know-how“ beschrieben und bildet eine essenzielle Grundlage für die Transferanwendung des Wissens bei ähnlichen Aufgaben [6].

In der manuellen Montage werden neue Mitarbeitende häufig über „learning by doing“ anhand von Montageanleitungen oder über Demonstrationen einer erfahrenen Fachkraft angelernt. Um implizites Wissen von Erfahrenen auch ohne deren Anwesenheit zugänglich zu machen, ist eine Form von Wissensexternalisierung erforderlich. Die meisten bisherigen Ansätze, implizites Wissen für die Weitergabe zu externalisieren, basieren auf verbaler Kommunikation in Form von Interviews oder Prozessbeschreibungen mit und durch Experten. Das externalisierte Expertenwissen wird dabei meist zeitaufwendig in Form von Ontologien oder logischen Regelabfolgen aufbereitet [6, 7].

In der Produktion kommen weitere Randbedingungen rund um die Wissensvermittlung hinzu. So existiert dort häufig unzu-

reichendes Wissensmanagement hinsichtlich der Form und Struktur, in der Wissen vorliegt. Zudem wird bei der Verwaltung von Wissen selten auf die Bedürfnisse der Nutzenden geachtet. Es gibt viele theoretische Ansätze, Wissensmanagement in unterschiedlichen Bereichen wie der Prozessplanung oder Qualitätskontrolle anzuwenden und digital abzubilden. In den meisten Fällen fehlen jedoch die Digitalisierung des Wissens sowie ein universeller Zugang und die Möglichkeit, Wissen zu teilen oder Updates durchzuführen. Dies ist insbesondere für innovative Entwicklungen im Bereich Smart Manufacturing von großer Bedeutung [8].

3 Konzept

Aus der beschriebenen Ausgangssituation kristallisiert sich der Bedarf heraus, ein System zu konzipieren und umzusetzen, das sowohl automatisiert bei der Externalisierung des Erfahrungswissens als auch bei der Kompetenzvermittlung basierend auf dem Erfahrungswissen unterstützt. Um dem zeitintensiven Anlernen neuen Personals durch Fachkräfte entgegenzuwirken, soll ein Assistenzsystem entwickelt werden, das weniger erfahrenen Mitarbeitenden den selbstständigen Einstieg in den Montageprozess erleichtern soll (Lösungsansatz zu Problem 1). Das Konzept des Vorhabens ist in **Bild 1** skizziert.

Die Anzulernenden werden über eine grafische Benutzeroberfläche (GUI) durch die Produktionsschritte geleitet. Die Reihenfolge der einzelnen Schritte des Montageprozesses wird dabei mittels vorab trainierter Aktivitätserkennung kamerabasiert überprüft. Auffälligkeiten, wie vergessene Schritte oder eine abweichende Reihenfolge, werden den Lernenden über die GUI über multimodales Feedback mitgeteilt. Für den Vergleich der Aktivitätsreihenfolge wird bestehendes explizites Wissen in Form von Montageanleitungen des Unternehmens herangezogen. In der Montage gibt es abseits des expliziten Wissens oft auch verschiedene Handgriffe, die die Ausführung einer Aktivität erleichtern oder die Ergonomie verbessern können. Um die Lernenden dabei zu unterstützen, eine eigene Routine zu finden, sollen ihre Ausführungen zusätzlich mit Best Practices von Experten verglichen werden.

Die Externalisierung des impliziten Erfahrungswissens, in diesem Fall der Handgriffe und Ausführungen der Aktivitäten, soll automatisiert abgedeckt werden. Dazu zeichnet eine am Arbeitsplatz montierte Kamera den Ablauf nicht-invasiv (also störungsfrei) auf. Für die Testaufnahmen wurde eine herkömmliche Logitech-C920-Webcam verwendet, die circa einen Meter über der Werkbank in Vogelperspektive befestigt wurde. Die genaue Höhe richtet sich nach dem Aufbau des Arbeitsplatzes der Anwendungspartner und des benötigten Sichtfelds, um alle Aktivitäten ausreichend abzubilden. Die Experten werden aufgefordert, während der Aufnahme ihre natürlichen Arbeitsabläufe durchzuführen. Durch diese Herangehensweise soll wertvolles implizites Erfahrungswissen darüber, „wie“ einzelne Aktivitäten ausgeführt werden, automatisch externalisiert werden. Dadurch können Best Practices erstellt werden, ohne die Fachkräfte in ihrem Arbeitsfluss zu unterbrechen.

Diese Best Practices dienen als Vergleichsgrundlage für das Ausführungsfeedback des Assistenzsystems. Das übergeordnete Ziel ist, spezielle Handgriffe oder Ausführungen, die sich die Mitarbeitenden über die Jahre angeeignet haben, für den Wissensaustausch zugänglich zu machen (Lösungsansatz Problem 2.1). Durch das mehrmalige Ausführen können die Best Practices der

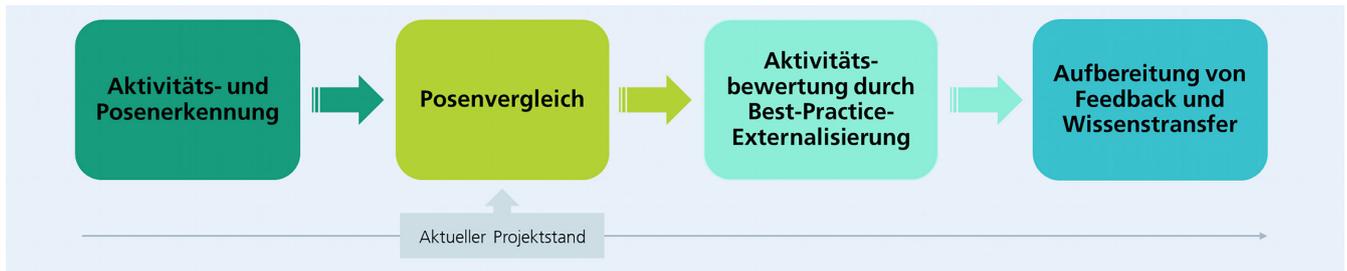


Bild 2 Flowchart des chronologischen Verlaufs der Projektthemen. Grafik: Fraunhofer IPA

einzelnen Aktivitäten gebündelt und Ausreißer entfernt werden. Grundsätzlich können mehrere Best Practices einer Aktivität existieren. Der implementierte Posenvergleich soll sicherstellen, dass sich diese Best Practices auch deutlich unterscheiden (Lösungsansatz zu Problem 2.2). Bei mehreren ermittelten Ausführungsoptionen können diese getestet und die passende Variante ausgewählt werden. Dieses individuelle Feedback ermöglicht den Lernenden, sich neue Kompetenzen anzueignen, ohne auf die Begleitung eines erfahrenen Kollegen angewiesen zu sein. Das Projekt ist in modulare Themenabschnitte eingeteilt, die aufeinander aufbauen, um das übergeordnete Ziel des Wissenstransfers zu ermöglichen. Der chronologische Verlauf der Themenabschnitte sowie der aktuelle Projektstand sind in **Bild 2** dargestellt.

4 Umsetzung

Die zwei Hauptkomponenten, die für die Umsetzung dieses Assistenzsystems trainiert und entwickelt werden, sind eine Aktivitätserkennung basierend auf künstlicher Intelligenz (KI) und ein Analyseverfahren für den Posenvergleich. Die technischen Hintergründe sowie der Forschungsstand und bestehende Herausforderung werden im Folgenden beleuchtet.

4.1 Aktivitätserkennung

Aktivitätserkennung ist ein entscheidender Baustein, um Systeme zu befähigen, Verhalten und Ereignisse in einer Szene zu verstehen. Ausgangspunkt ist ein Machine-Learning-Modell (ML-Modell), das zum Klassifizieren von Aktivitäten trainiert wird. Sensoren erfassen die Eingangswerte des Modells. So wird beispielsweise bei der sogenannten kamerabasierten Aktivitätserkennung die Szene über eine Kamera(-Sensorik) aufgenommen. Als Ausgabewert liefert das Modell die Aktivitätsklasse.

Besonders die Aktivitätserkennung beim Menschen (Human Activity Recognition, HAR) eröffnet zahlreiche Anwendungsmöglichkeiten. Durch den Einsatz von Sensorik und ML können menschliche Aktivitäten erfasst und klassifiziert werden, auch in Echtzeit. Die kamerabasierte HAR liefert, vereinfacht ausgedrückt, Antwort auf die Frage, was eine oder mehrere Personen im aufgenommenen Kamerasichtfeld zu einem bestimmten Zeitpunkt oder innerhalb eines bestimmten Zeitraums tun. So kann beispielsweise im Gesundheitswesen das Wohlbefinden alter oder kranker Menschen beobachtet werden, um Stürze zu erkennen und zu verhindern [9].

Ein konkretes Beispiel für die Anwendung der Aktivitätserkennung in der Industrie ist die Assistenz bei Montageprozessen, um Abweichungen vom vorgesehenen Ablauf frühzeitig festzustellen und entsprechend zu reagieren. Das System kann hier erkennen, ob eine Arbeitsabfolge korrekt durchgeführt oder ein

Schritt vergessen wurde und dem Werker Feedback geben. Dies stärkt nicht nur die Effizienz und verringert die Zahl der benötigten Nacharbeiten, sondern fördert auch das Sicherheitsgefühl der Mitarbeitenden, indem potenzielle Fehlerquellen frühzeitig identifiziert werden.

Aktivitätserkennung ermöglicht auch zu dokumentieren, wie Experten den Prozess durchführen. Es könnten verschiedene Durchführungsreihenfolgen der Experten ermittelt werden, also eine Externalisierung des Erfahrungswissen zur Reihenfolge. Zusätzlich können kombiniert mit der Analyse und dem Vergleich der Tätigkeitsausführung weitere Best Practices ermittelt werden, die für Schulungszwecke und die Weitergabe von Wissen an neue Mitarbeitende von Bedeutung sind.

Obwohl es zahlreiche Anwendungsfelder für HAR gibt, in denen auch Forschung betrieben wird, bestehen in diesem Gebiet einige Herausforderungen. Vor allem komplexe Aktivitäten, die aus mehreren Teilaktivitäten zusammengesetzt sind, treten im industriellen Kontext häufig auf. Auch mangelt es an ausreichenden Datensätzen für industrielle Anwendungen, was den Labelingaufwand für neue Prozesse und Aktivitäten erhöht. Um die Aufwände für dieses Labeln zu minimieren, könnte Prelabeling als Verfahren betrachtet werden, bei dem erste (grobe) Labels (semi-)automatisch generiert werden.

4.2 Posenvergleich

In vielen Anwendungsbereichen stellt sich nicht nur die Frage, ob eine Aktivität ausgeführt wurde, sondern auch wie sie ausgeführt wurde. Menschliche Aktivitätsevaluierung (HAE) erweitert HAR um die Bewertung der erkannten Aktivitäten basierend auf einem Qualitätsmaßstab. Im Leistungssport und in der Rehabilitation gibt es Bestrebungen, Bewegungen mittels HAE automatisch auszuwerten und sie hinsichtlich ihrer Ausführungen zu bewerten. Diese Bewertungen können zur Bestimmung von Leistungsentwicklungen, Platzierungsvorhersagen oder Trainingsanpassungen genutzt werden [10–12]. Ansätze, durch HAE chirurgische Ausführungen automatisch zu bewerten [13, 14], eröffnen die Möglichkeit, HAE in der Kompetenzvermittlung weiterer motorischer Fähigkeiten einzusetzen.

Bei einem Großteil der bisherigen Arbeiten wurden menschliche Bewegungen kamerabasiert mithilfe KI-basierter Posenerkennung (PE) abgebildet. PE-Modelle detektieren die Koordinaten einer vordefinierten Anzahl an Gelenkspunkten (und deren Verbindungen) von menschlichen Körpern in einzelnen Bildern und Videoausschnitten. Dabei kann eine PE auch trotz teilweiser Verdeckung der Handgelenkspunkte diese zuverlässig in den Bilddaten schätzen. Bei Aktivitäten handelt es sich um charakteristische, raumzeitliche Abfolgen einzelner Posen. Diese raumzeitlichen Veränderungen können sowohl durch Trajektorien oder

durch relative Abstände der Gelenke und Körperteile sowie deren Winkel(-rotation) als auch durch eine Vielzahl an selbstdefinierter Merkmale charakterisiert werden. Neben einer akkuraten HAR ist die robuste Darstellung menschlicher Posen und ihre raumzeitliche Veränderung eine weitere Herausforderung für HAE [15].

Bevor Bewegungen bewertet werden können, müssen Unterschiede in den jeweiligen Ausführungen erkannt werden. Abhängig von der gewählten raumzeitlichen Repräsentation der Aktivitäten durch Posen oder vordefinierte Merkmale können dazu verschiedene Vergleichsmetriken gewählt werden. Dafür werden in der Literatur Zeitreihenanalysen und ML-Ansätze verwendet.

Eine besondere Herausforderung beim Vergleich von Posen ist die natürliche Varianz von Bewegungen innerhalb einer Aktivität. Somit ist die Vorverarbeitung der rohen Gelenkpunkt-Koordinaten beim Posenvergleich essenziell. Um tatsächliche Bewegungsunterschiede zu detektieren, müssen die Repräsentationen invariant zu räumlicher und zeitlicher Verschiebung sein. Andernfalls werden sehr ähnliche Bewegungen durch Variation in der Position oder des Abstandes der Kamera sowie unterschiedliche Körperproportionen als auch Ausführungsgeschwindigkeiten deutlich unterschiedlich abgebildet. Mit dem Ziel, dem Einfluss einer räumlichen Verschiebung entgegenzuwirken, werden die rohen Daten z-normalisiert. Um Zeitreihen unabhängig von ihrer Geschwindigkeit zu vergleichen, ist der Einsatz von Dynamic Time Warping (DTW) die etablierteste Methode in der Literatur [12, 15].

Die automatische Bewertung von erkannten Aktivitäten und deren Merkmalen scheitert oft an der unzureichenden Vergleichsgrundlage: Es existieren kaum umfassende, öffentliche Datensätze für HAE. Das hat mehrere Gründe: Zum einen sind Aktivitäten meist domänenspezifisch. Ein genereller Datensatz kann die Komplexität einzelner Anwendungsfälle also kaum abdecken. Des Weiteren erfordern Vergleichsbewertungen wiederum Expertenwissen, was kostspielig und zeitintensiv ist. Videoaufnahmen enthalten zudem oft sensible personen- oder unternehmensbezogene Daten, die dem Datenschutz unterliegen. Auch existiert kein genereller Maßstab, ab wann sich eine Bewegung von einer anderen unterscheidet. Somit muss im Einzelfall definiert werden, wie viel Abweichung innerhalb einer Aktivitätsklasse als natürliche Varianz gewertet wird. Insbesondere für komplexe Aktivitäten ist eine gewisse Abweichung zu erwarten, selbst wenn dieselbe Person die Bewegungsabläufe mehrmals ausführt [15].

5 Vorgehen und Ergebnisse

Das am Fraunhofer IPA entwickelte System besteht aus zwei Hauptkomponenten: Zum einen werden die ausgeführten Prozessschritte mittels Aktivitätserkennung auf ihre Reihenfolge hin überprüft. Für den Abgleich der Aktivitätsreihenfolge wird bestehendes explizites Wissen in Form von Montageanleitungen des Unternehmens herangezogen. Zum anderen wird implizites Expertenwissen automatisch externalisiert, abgespeichert und an lernende Mitarbeitende mittels HAE weitergegeben. Die automatische Erfassung von Experten-Best-Practices, also des impliziten Wissens, erfolgt über mehrmalige Aufnahmen des natürlichen Arbeitsprozesses der erfahrenen Mitarbeitenden durch eine Kamera in Vogelperspektive. Mittels KI-Modellen werden die zweidimensionalen Koordinaten der Handposen bestimmt und als Zeitreihen erfasst. Durch das wiederholte Ausführen werden die Posendaten

der einzelnen Aktivitäten gebündelt und Ausreißer entfernt. Um weiteren Zeitaufwand durch Experten zu vermeiden, indem sie Ausführungen manuell als Best Practice identifizieren, liegt dem Projekt die Annahme zugrunde, dass sich jeder langjährige Mitarbeitende Best Practices angeeignet hat.

Im Anschluss sollen zudem wichtige Kennzahlen wie die Ausschussquote und die Montagezeiten verwendet werden, um Best Practices zu identifizieren. Grundsätzlich können mehrere Best Practices einer Aktivität existieren. Der implementierte Posenvergleich soll sicherstellen, dass sich diese Best Practices auch deutlich voneinander unterscheiden. Die Best Practices dienen als Vergleichsgrundlage für HAE.

Der für HAE verwendete Posenvergleich basiert auf etablierten Methoden wie der z-Normalisierung für die räumliche und DTW für die zeitliche Invarianz. Somit sollen unterschiedliche Lokalisationen oder Größen der Hände sowie die Ausführungsgeschwindigkeit den Vergleich der jeweiligen Ausführungen nicht wesentlich beeinflussen. Des Weiteren wird ein gleitender Mittelwert auf den Zeitreihen angewandt, um kleinere Abweichungen und Schätzungsfehler des KI-Posenmodells auszugleichen. Durch den DTW-Algorithmus können Abschnitte unterschiedlich langer Zeitreihen auf ähnliche Muster untersucht werden. Dazu berechnet DTW die jeweiligen Abstände zweier Elemente der Zeitreihen und minimiert diese. **Bild 3** zeigt zwei Zeitreihen der jeweiligen rechten Zeigefingerspitze bei verschiedenen Ausführungen einer Aktivität aufgeteilt nach den normalisierten x-Koordinaten (Dimension 1) und normalisierten y-Koordinaten (Dimension 2).

Die Verbindungen zwischen den Zeitreihen symbolisieren, welche Elementpaare der Ausführungen gemäß des DTW-Algorithmus ähnliche Muster aufweisen. Teilsequenzen, die Muster aufweisen, welche in der anderen Zeitreihe nicht vorhanden sind, werden oft auf einzelne Elemente der anderen Zeitreihe abgebildet. Die so entstehenden dreieckigen Formen symbolisieren Unterschiede in den Zeitreihen. Anhand des „Warping Path“, der aus dem DTW-Algorithmus entsteht, können Zeitabschnitte identifiziert werden, in denen die Bewegungsabläufe von einem definierten Schwellenwert abweichen.

Die Lernenden werden während des Montageprozesses ebenfalls in Vogelperspektive aufgenommen und erhalten vom System Echtzeit-Feedback über die ausgeführte Prozessreihenfolge. Wenn sie die Prozessreihenfolge korrekt eingehalten haben, erfolgt nach jeder durchgeführten Aktivität personalisiertes Feedback über alternative Ausführungsmöglichkeiten. Dies erfolgt durch einen Posenvergleich der aufgenommenen Aktivität und der zugehörigen Best Practices, indem Zeitabschnitte identifiziert werden, in denen die Ausführungen der Lernenden Unterschiede aufweisen. Diese Zeitintervalle werden genutzt, um die Ausführungsunterschiede visuell anhand von Videoausschnitten darzustellen und dem Nutzer Feedback zu geben. Beispiele von unterschiedlichen Ausführungen der Aktivität „Schraube mittels Innensechskantschlüssels festziehen“ sind in **Bild 4** dargestellt: „Ausführung A“ zeigt das Festziehen der Schraube mittels der langen Seite des Innensechskantschlüssels, während bei „Ausführung B“ die Schraube mit der kurzen Seite des Werkzeugs festgezogen wird. Durch den implementierten Posenvergleich wurde ein deutlicher Unterschied der beiden Ausführungen festgestellt.

Obwohl die Erkennung von unterschiedlichen Ausführungen einiger Montageaktivitäten vielversprechende erste Ergebnisse liefert, gibt es noch einige Herausforderungen zu bewältigen. Eine offene Frage ist, wie viele und welche Gelenkpunkte der Hand-

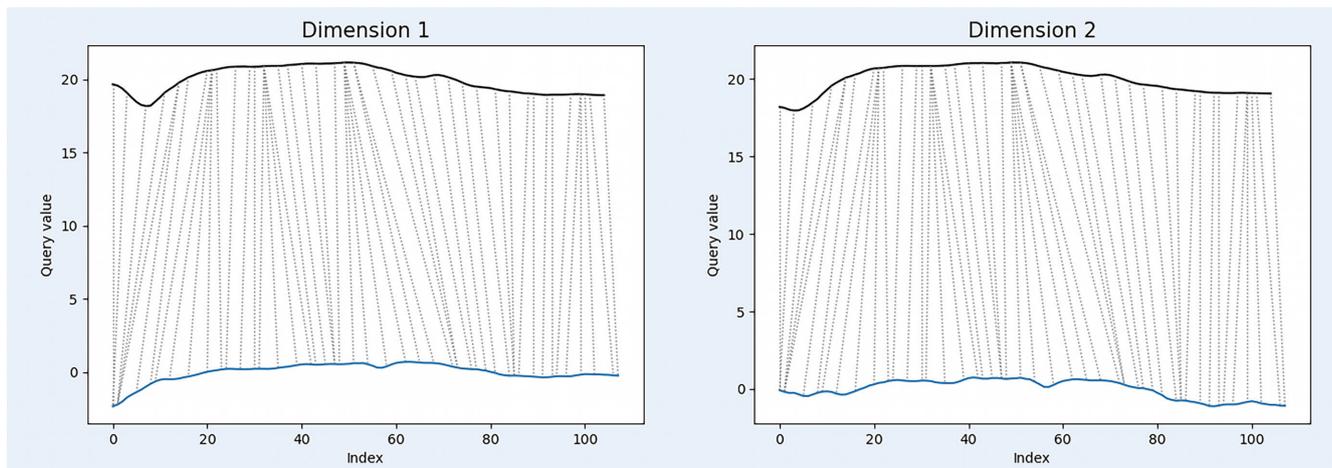


Bild 3 Darstellung zweier verschiedener Ausführungen einer Aktivität mittels Zeitreihen. Dimension 1 bildet jeweils die normalisierten x-Koordinaten und Dimension 2 die normalisierten y-Koordinaten der jeweiligen rechten Zeigefingerspitze ab. Grafik: Fraunhofer IPA

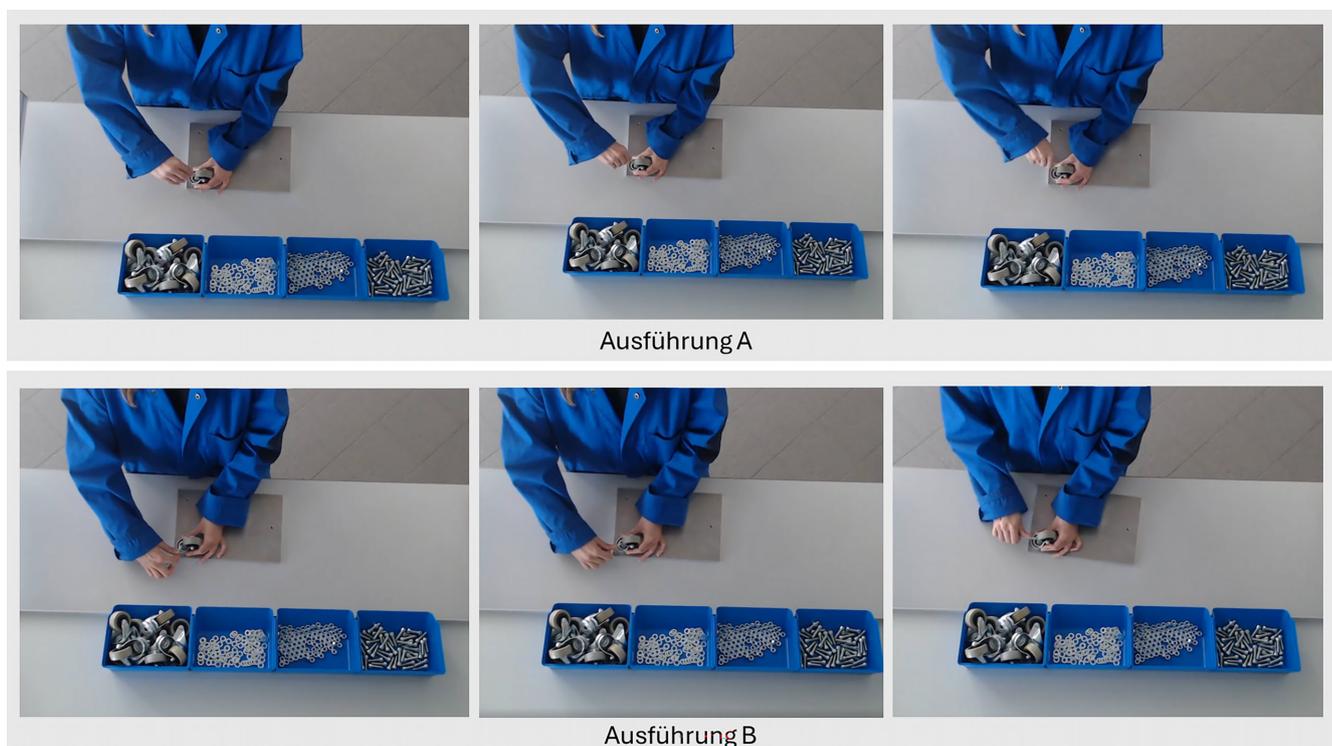


Bild 4 Darstellung zweier verschiedener Ausführungen der Aktivität „Festziehen einer Schraube mittels Innensechskantschlüssels“. Fotos: Fraunhofer IPA

posen für Montagetätigkeiten relevant sind. Je mehr einzelne Gelenkpunkte für den Vergleich herangezogen werden, umso schwieriger wird es, ähnliche Bewegungen als diese zu erkennen, da mehrdimensionale Daten mehr Rauschen und natürliche Varianz enthalten.

Eine weitere Herausforderung stellt der Zeitreihenvergleich von ähnlichen Aktivitäten dar, in denen ungeplante Abweichungen auftreten. Ein mögliches Beispiel wäre das Herunterfallen einer Schraube. Selbst wenn die Ausführung ansonsten identisch zum Best Practice ist, würde der Posenvergleich anhand der zusätzlichen Bewegung einen Unterschied detektieren. Einfaches Glätten der Zeitreihen beziehungsweise das Entfernen einzelner Ausreißer reicht hierfür nicht aus. Auch die Frage, ab welcher Abweichung eine Ausführung unterschiedlich zu einer anderen

ist, stellt einen aktuellen Forschungsgegenstand dar. Beispielsweise resultieren, trotz zeitunabhängigen Vergleichs mittels DTW, unterschiedlich beschleunigte Ausführungsgeschwindigkeiten und oszillierende Bewegungen (wie das Eindrehen einer Schraube per Hand) in großen Abweichungen der Zeitreihen, obwohl die Bewegungen ähnlich sind. Zudem sind Prozesse wie die Datenpflege (Problem 2.3) und die Erweiterbarkeit des Erfahrungswissens durch die Anwendenden noch nicht hinreichend gelöst.

6 Fazit und Ausblick

Der vorgestellte Ansatz zielt darauf ab, den Wissenstransfer in der manuellen Montage effizienter zu gestalten und den Verlust von personengebundenem Erfahrungswissen zu reduzieren. Die

automatisierte Erfassung und Vermittlung von Best Practices verkürzt die Einarbeitungszeit neuer Mitarbeitender und steigert die Qualität ihrer Arbeit sowie die Qualität des Einarbeitungsprozesses selbst. Zugleich bleibt der zwischenmenschliche Austausch von hoher Bedeutung, da Fragen und spezielle Situationen oft im direkten Dialog geklärt werden können. Der Ansatz ergänzt den persönlichen Wissenstransfer somit durch eine technische Lösung, die durch den Einsatz von posebasierten Verfahren datenschutzkonform gestaltet werden kann.

Der aktuelle Stand enthält Verbesserungspotenzial: Einerseits führen unerwartete Ereignisse (wie die erwähnte fallengelassene Schraube) oder komplexe Bewegungsabläufe zu ungenauen Abbildungen der Posen. In der Folge kann dies zu Problemen in der weiteren Verarbeitungskette führen, sodass Abweichungen nicht korrekt identifiziert werden können. Andererseits stellt sich die Frage nach der Definition relevanter Bewegungsbereiche, um natürliche Bewegungsvariationen nicht als Fehler zu werten. Es fehlt eine eindeutige Definition, ab wann eine Bewegung als abweichend gilt. Es besteht somit weiterer Forschungsbedarf, geeignete Schwellwerte zu bestimmen und diese idealerweise sogar kontinuierlich an neue Anforderungen anzupassen. Ausführliche, quantitative Evaluierungen der Ergebnisse sind im weiteren Verlauf des Projektes geplant, vor allem hinsichtlich des tatsächlichen Wissenstransfers.

Das entwickelte Verfahren kann in zukünftigen Arbeiten dazu genutzt werden, die optimale Reihenfolge von Produktionsschritten zu ermitteln. Dazu werden die einzelnen, in sich optimierten Prozessschritte in einem Wissensgraphen erfasst, der wiederum zwischen verschiedenen Personen abgeglichen und optimiert wird. Ein weiterer Aspekt ist die Berücksichtigung ergonomischer Aspekte in der automatisierten Prozessanalyse. So können ungünstige Körperhaltungen oder repetitive Bewegungen frühzeitig erkannt und präventive Maßnahmen ergriffen werden. Durch die Erkennung von Abweichungen in den einzelnen Gelenkposen kann zwischen ergonomisch günstiger Haltung und Fehlhaltungen unterschieden werden.

FÖRDERNACHWEIS

Das Projekt „Wissensbasierte Produktionshilfe beruhend auf bild- und sensorbasierter Externalisierung von Erfahrungswissen“ („WisPE“) wird im Rahmen des Programms „Zukunft der Arbeit“ vom Bundesministerium Forschung, Technologie und Raumfahrt (BMFTR) und der EU über den Europäischen Sozialfonds Plus (ESF Plus) gefördert und vom Projektträger Karlsruhe (PTKA) betreut.

LITERATUR

- [1] Deutscher Industrie- und Handelskammertag e.V. (DIHK): Fachkräfteengpässe schon über Vorkrisenniveau. DIHK-Report Fachkräfte 2021. Stand: 2021. Internet: <https://www.dihk.de/resource/blob/61638/9bde58258a88d4fce8cda7e2>

ef300b9c/dihk-report-fachkraefteversicherung-2021-data.pdf. Zugriff am 11.08.2025, S. 4–27

- [2] Bundesministerium für Arbeit und Soziales (BMAS): Langfristprojektion des Fachkräftebedarfs in Deutschland, 2021–2040. Szenario „Fortschrittliche Arbeitswelt“ (Annahmensetzung nach dem Koalitionsvertrag von 2021). Stand: 2023. Internet: <https://www.bmas.de/DE/Service/Publikationen/Forschungsberichte/fb-617-langfristprojektion-des-fachkraeftebedarfs.html>. Zugriff am 11.08.2025
- [3] Mariano, S.: Mitigating the Disruptive Consequences of Knowledge Loss in Organizational Settings: Knowledge Loss Clusters and Potential Organizational Interventions. *European Conference on Knowledge Management* 24 (2023) 1, pp. 872–880
- [4] Hulstijn, J. H.: Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition* 27 (2005) 02, pp. 129–140
- [5] Neuweg, G.: Tacit knowing and implicit learning. In: Fischer, M.; Boreham, N.; Nyhan, B. (eds.): *European perspectives on learning at work. The acquisition of work process knowledge*. Luxembourg: Office for Official Publications of the European Communities 2004, pp. 130–147
- [6] Johnson, T. L.; Fletcher, S. R.; Baker, W. et al.: How and why we need to capture tacit knowledge in manufacturing: Case studies of visual inspection. *Applied ergonomics* 74 (2019), pp. 1–9
- [7] Titah, M.; Aitouche, S.; Mouss, M. D. et al.: Externalising and reusing of tacit knowledge in manufacturing task. *International Journal of Knowledge Management Studies* 8 (2017) 3/4, p. 351
- [8] Feng, S. C.; Bernstein, W. Z.; Hedberg, T. et al.: Towards Knowledge Management for Smart Manufacturing. *ASME Journal of computing and information science in engineering* 17 (2017) 3, #JCISE-16-2051, doi.org/10.1115/1.4037178
- [9] Kaur, H.; Rani, V.; Kumar, M.: Human activity recognition: A comprehensive review. *Expert Systems* 41 (2024) 11, doi.org/10.1111/exsy.13680
- [10] Pirsiavash, H.; Vondrick, C.; Torralba, A.: Assessing the Quality of Actions. In: Fleet, D.; Pajdla, T.; Schiele, B. et al. (eds.): *Computer Vision – ECCV 2014*. Cham: Springer International Publishing 2014, pp. 556–571
- [11] Patrona, F.; Chatzitofis, A.; Zarpalas, D. et al.: Motion analysis: Action detection, recognition and evaluation based on motion capture data. *Pattern Recognition* 76 (2018), pp. 612–622
- [12] Yan, M.; Liu, X.; Li, Z. et al.: Evaluation of Human Action Based on Feature-Weighted Dynamic Time Warping. *Applied Sciences* 14 (2024) 23, #11130
- [13] Zia, A.; Sharma, Y.; Bettadapura, V. et al.: Automated video-based assessment of surgical skills for training and evaluation in medical schools. *International journal of computer assisted radiology and surgery* 11 (2016) 9, pp. 1623–1636
- [14] Zia, A.; Essa, I.: Automated surgical skill assessment in RMIS training. *International journal of computer assisted radiology and surgery* 13 (2018) 5, pp. 731–739
- [15] Lei, Q.; Du, J.-X.; Zhang, H.-B. et al.: A Survey of Vision-Based Human Action Evaluation Methods. *Sensors (Basel, Switzerland)* 19 (2019) 19, #4129, doi.org/10.3390/s19194129

Laura Häge, M. Sc. 
laura.haegel@ipa.fraunhofer.de

Dascha, Karelina, M. Sc. 

Jauch, Christian, M. Sc. 

Fraunhofer-Institut für Produktionstechnik und Automatisierung IPA 
Nobelstr. 12, 70569 Stuttgart
www.ipa.fraunhofer.de

LIZENZ



Dieser Fachaufsatz steht unter der Lizenz Creative Commons Namensnennung 4.0 International (CC BY 4.0)