

Digitale Ergonomieanalysen mittels tiefenbildbasierter Echtzeiterfassung von Arbeitsbewegungen

Martin SCHMAUDER, Edgar SCHERSTJANOI, Daniel GRÖLLICH

*Professur für Arbeitswissenschaft, Institut für Technische Logistik und
Arbeitssysteme, Fakultät Maschinenwesen, Technische Universität Dresden
Helmholtzstraße 10, D-01069 Dresden*

Kurzfassung: Ergonomische Analysen von Arbeitsprozessen sind ein wichtiger Bestandteil zur Optimierung von Arbeitsbedingungen, die im Wesentlichen von Umweltfaktoren, der Arbeitsplatzgestaltung und nicht zuletzt auch von der konstruktiven Auslegung von Maschinenkomponenten abhängen. Aktuelle Forschungen zu Ergonomieverfahren nutzen die stets wachsende Verfügbarkeit von Methoden der computergestützten Bildverarbeitung, um Evaluationsprozesse objektiver, effektiver und effizienter gestalten zu können. In diesem Beitrag wird ein Verfahren vorgestellt, welches mobile Tiefenbildkameras als Sensoren zur Aufnahme menschlicher Bewegungen nutzt, wodurch eine ergonomische Evaluation von Arbeitsprozessen in Echtzeit, ohne virtuelle Modellierung der Umgebung oder Verwendung von optischen Markern, möglich ist.

Schlüsselwörter: Ergonomiebewertung, Echtzeitanalyse, Tiefenbildkamera, Bewegungstracking, Lendenwirbelsäule, Druckkraft

1. Motivation

Bisherige Methoden zur ergonomischen Analyse werden i.d.R. durch einen geschulten Anwender auf Basis einer Prozessbeobachtung durchgeführt, worin auch die Problematik - die subjektive Einschätzung der Situation - zu finden ist. Dank fortschreitender Technik finden aber auch immer häufiger computergestützte Methoden Einsatz, da dort eindeutig anhand erhobener anthropometrischer Daten klare Entscheidungen getroffen werden können. Zum einen werden oft digitale Menschmodelle in CAD-Umgebungen verwendet, wie beispielsweise „CharAT Ergonomics“ als Erweiterung in der Modellierungssoftware „3d Studio Max“ (Kamusella et al. 2015). Zum anderen schreitet auch die Entwicklung von Sensorik voran, die zur Aufnahme realistischer menschlicher Bewegungen eingesetzt werden kann. In der Softwarelösung „ema“ werden beispielsweise virtuelle Arbeitsbewegungen a priori durch detailreiches Tracking von Körperbewegung modelliert (Fritzsche et al. 2011). Gängige Tracking-Verfahren sind allerdings häufig mit großem Aufwand für die Installation der benötigten Technik im Raum und an Personen in ihren Einsatzmöglichkeiten eingeschränkt. Des Weiteren produzieren sensible Sensoren in der Regel ein hohes Datenvolumen, weswegen eine Nachbearbeitung des Datenstroms notwendig wird. Üblicherweise werden die modellierten und erfassten Daten anschließend genutzt, um software-basierte ergonomische Untersuchungen und Auswertungen durchzuführen.

Dieser Beitrag beschäftigt sich mit einer alternativen Erfassungstechnologie und einer bisher noch nicht softwareseitig umgesetzten Ergonomiemethode. Die Technologie ermöglicht es, Datenströme von mehreren Tiefenbildkameras einer automati-

sierten Auswertung zuzuführen. Das modular und mobil aufgebaute System ist dazu in der Lage, mit geringem Installationsaufwand direkt vor Ort, ohne Verwendung von optischen Markierungen oder räumlichen Lagesensoren in Echtzeit verwertbare Daten zu liefern. Es wird ermöglicht, Körperhaltungen, Raumpositionen von Gliedmaßen und Bewegungen nachzuverfolgen und softwaregestützt zu analysieren. Für die Implementierung zur Erfassung von ergonomischen Belastungsgrößen wurde das haltungsabhängige Verfahren zur Ermittlung der Lendenwirbelbelastung nach Jäger et al. (2010) gewählt. Alle dafür notwendigen Daten können vom entwickelten System aufgenommen werden, sodass im Ergebnis einer Anwendung die Analyse zur Belastung der Lendenwirbelsäule bei Hebetätigkeiten steht.

2. Tiefenbildsensoren als Marker-freie Alternative zur Ergonomieanalyse

Zur Generierung eines Tiefenbildes werden im Rahmen des entwickelten Erfassungssystems Kameras der Marke ASUS (XTion Pro Live) verwendet, bestehend aus Infrarot(IR)-Emitter, IR-Kamera, herkömmlicher RGB-Kamera und einem Softwarepaket zur Generierung von Skelettdaten in aufgenommenen Bildsequenzen. Der IR-Emitter projiziert ein positionseindeutiges Muster in das Sichtfeld der IR-Kamera, welche wiederum durch die räumliche Verzerrung des Musters Rückschlüsse zu Entfernungen für jeden Bildpunkt liefert. In Abbildung 1 sind Beispielbilder zur Verdeutlichung der Funktionsweise dargestellt. Segmentierungsalgorithmen unterscheiden Objekte, die räumlich voneinander getrennt sind anhand ihrer Bewegung. Entsprechen ein segmentiertes Objekt und dessen Bewegung, einer menschlichen Gestalt, so wird algorithmisch versucht, Skelettpunkte darin einzuordnen (vgl. Abb. 1, rechts).

Der IR-Emitter ist dazu in der Lage das Muster bis zu einer Entfernung von ~5m zu projizieren, so dass die Muster von der Kamera erkannt und verarbeitet werden können. Nur wenn das projizierte IR-Muster von der IR-Kamera erkannt wird, erfolgt auch eine Skelettgenerierung. Das generierte Skelett besteht aus 15 Knotenpunkten, darunter Kopf, Hals, Torso und jeweils für beide Körperhälften Schulter, Ellenbogen, Hand, Hüfte, Knie und Fußpunkte. Jeder Punkt der Skelettstruktur wird durch seine räumlichen Koordinaten, Orientierung und einem Wert zur Vertrauenswürdigkeit der Generierung beschrieben. Genannte Konfidenz sinkt, wenn beispielsweise Körperabschnitte verdeckt werden bzw. teilweise oder kurzzeitig den Aufnahmebereich verlassen. In solchen Fällen werden anthropometrische Annahmen getroffen, die sich auf erfolgreich erkannte Körperteile beziehen.

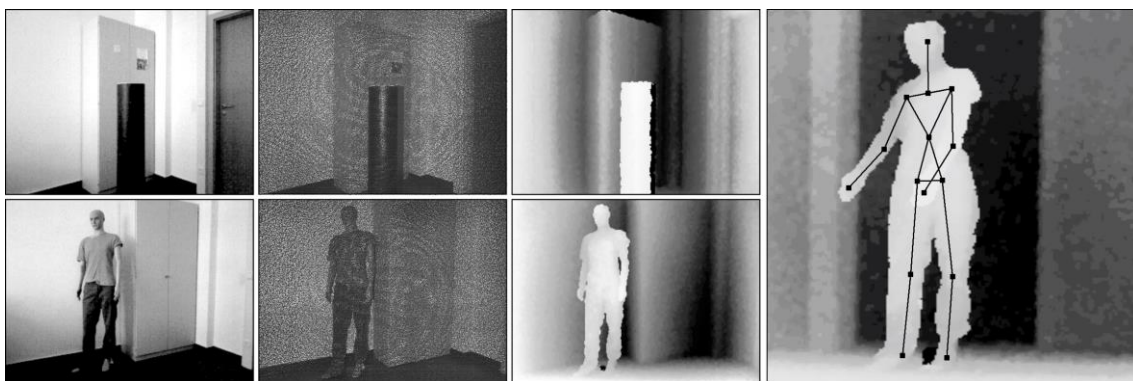


Abbildung 1: Verdeutlichung zur Funktionsweise einer Tiefenbildkamera. Links: Bild im Sichtfeld der Kamera. Mitte: Überlagerung des Bildes mit dem projizierten IR-Muster und entstandenes Tiefenbild. Rechts: Generiertes Skelett im Tiefenbild.

Mit dem Ziel eine Aussage über die Genauigkeit des generierten Skelettes treffen zu können, wurden am Beispiel des Kniegelenks Untersuchungen durchgeführt. Hierfür wurde das Bein einer Testperson unter Einsatz einer medizinischen Gelenk-Orthese in fest definierten Kniewinkeln fixiert. Anschließend erfolgten Aufnahmen aus drei Perspektiven mit jeweils zwei Körperhaltungen. Die Perspektiven umfassten frontale und seitliche Aufnahmen, wobei bei letzterem einmal das zu untersuchende Bein vom anderen Bein verdeckt war und in einem weiteren Fall von der Kamera erfassbar war. Die Ermittlung der Gelenkstellung (Winkel) erfolgte mit der Tiefenbildkamera auf Basis des generierten Skelettes. Diese wurden anschließend mit den tatsächlich eingestellten Gelenkwinkeln abgeglichen, um den Fehlerwert zu ermitteln. Abbildung 2 zeigt anhand eines Beispiels den gemessenen Kniewinkel für eine Zeitspanne von 10s (30 Frames/Sekunde) bei einer Einstellung von 120°.

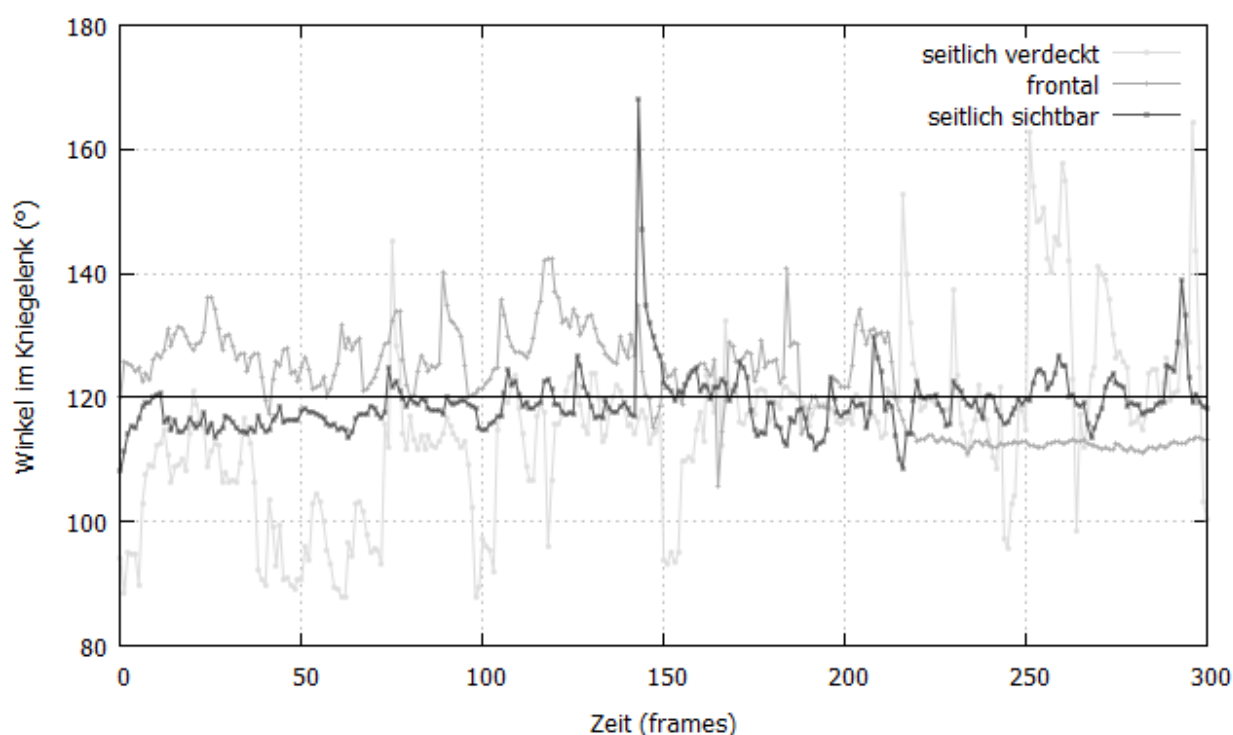


Abbildung 2: Messung eines fixierten Kniegelenks aus drei Perspektiven. In der Perspektive „seitlich verdeckt“ (hellgrau) ist das Kniegelenk durch das andere Bein verdeckt. Das Kniegelenk wurde auf 120° fixiert.

Der durchschnittliche Fehler (Abstand zum Erwartungswert der Gelenkorthese) einer Skeletterzeugung betrug in diesem Beispiel ~7°. Dieser Durchschnitt bezieht alle im Versuch vorgenommenen Aufnahmesituationen mit ein, auch solche, in denen das zu messende Kniegelenk durch das andere Bein verdeckt war. Im Idealfall konnten durchschnittliche Fehlerwerte von ~4° erreicht werden.

Diese Messungen zeigen, welche Genauigkeit beim Einsatz von Tiefenbildbasierten Skelettdaten zu erwarten ist. Selbstverständlich beziehen sich nicht alle Ergonomieverfahren auf Winkelgrößen im Kniegelenk, dennoch soll der Versuch als exemplarisch verstanden werden, da darin das Prinzip der Generierung von Skelettpunkten, deren Position und die daraus resultierenden anthropometrischen Informationen enthalten sind.

3. Erweiterung der Messtechnik für Ergonomieanalysen

Einige der im vorangegangenen Kapitel beschriebenen Eigenschaften einer Tiefenbildkamera schränken den Einsatz im Arbeitsumfeld ein, insbesondere Ungenauigkeiten bei verdeckten Gliedmaßen und der verhältnismäßig begrenzte Aufnahmebereich. Aus diesen Gründen war ein wesentlicher Bestandteil der Entwicklungsarbeit, eine Möglichkeit des parallelen Einsatzes mehrerer Tiefenbildsensoren zu schaffen, welche durch Motorik in Ihrer räumlichen Lage und Ausrichtung angepasst werden können. Die Erweiterung zielt dabei auf die Vergrößerung des Erfassungsbereiches und auf die Verbesserung der Genauigkeit von generierten Skelettdaten ab.

3.1 Optimierung der Skelettgenerierung durch Einsatz mehrerer Kamerasysteme

Das entwickelte Verfahren sieht den gleichzeitigen Einsatz mehrerer Tiefenkameras vor, die einen synchronen Datenstrom erzeugen und auswerten. Sind mehrere Tiefenkameras in den Aufnahmeprozess eingebunden, liegt der Gedanke nahe so auch Ungenauigkeiten in den Gelenkmessungen kompensieren zu können, da entweder Durchschnittswerte berechnet oder günstigere Kamerapositionen priorisiert werden können. Werden IR-Muster von mehreren Kameras auf das gleiche Motiv projiziert, entstehen stärker fehlerbelastete Tiefenbilder, wie in Abbildung 3 (links) verdeutlicht. Inwiefern die Qualität eines Tiefenbildes Auswirkungen auf die Qualität der Skelettgenerierung hat wird in Abbildung 3 (rechts) dargestellt. In diesem Versuch wurde das fixierte Kniegelenk über eine Zeitspanne von 30s (900 Frames) gemessen. In den ersten 10s (300 Frames) wurde im Sichtfeld nur ein IR-Muster emittiert, in den folgenden 10s ein zweites und schlussendlich ein drittes.

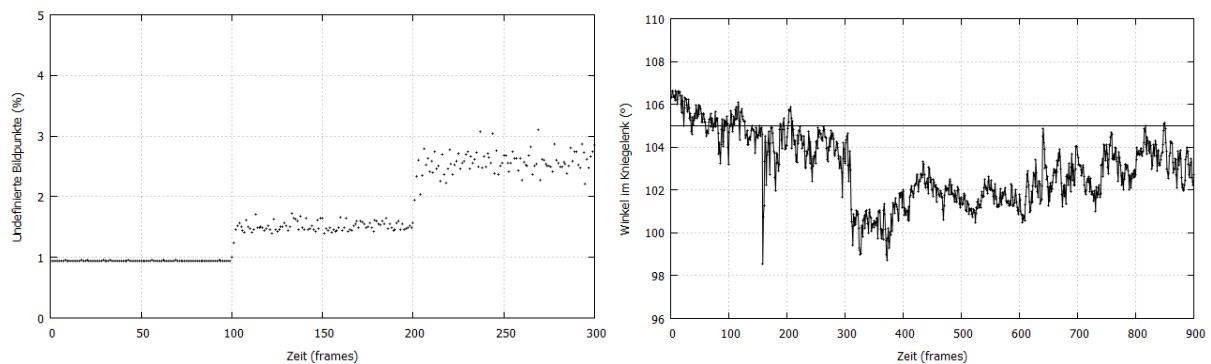


Abbildung 3: Links: Vergleich von Bildstörungen beim Einsatz mehrerer Kameras. Zum Zeitpunkt 100 wurde ein IR-Muster der zweiten, zum Zeitpunkt 200 das einer dritten Kamera aktiviert. Rechts: Vergleich von Störungen bei der Messung von Gelenkwinkeln bei Einsatz mehrerer Kameras. Ein Kniegelenk wurde dazu auf 105° mit einer Orthese fixiert. Zum Zeitpunkt 300 wurde ein IR-Muster im Aufnahmebereich hinzugefügt, zum Zeitpunkt 600 ein drittes.

Es wird ersichtlich, dass die Qualität der Gelenkerfassung zwar beeinträchtigt wird, aber relativ zur Bildstörung nur geringe Auswirkungen hat. Der durchschnittlich gemessene Winkel verfälscht sich im schlechtesten Fall um ~5°. Mit dieser Erkenntnis ist es naheliegend Skelettgenerierungen mehrerer Kameras zu vereinen, um entstandene Fehler auszugleichen. In der Arbeit von Kopniak (2015) wird beispielsweise gezeigt, dass, insofern IR-Muster sich nicht überlagern, Verbesserungen von bis zu ~2° zu erwarten sind.

3.2 Vergrößerung des Erfassungsraums durch mobile Sensoren

Durch die Bewegung eines Sensors kann der Aufnahmebereich um den Bewegungsraum erweitert werden. In Kooperation mit der Fa. imk automotive GmbH wurde hierzu eine prototypische Vorrichtung entwickelt, die der Kamera Bewegungsfreiheit in vier Richtungen ermöglicht (seitliches Verfahren, Höhenverstellung, Neigung um die Seitachse und Rotation um die Höhenachse). Jede Richtung wird durch Schrittmotoren angesteuert. Die in Abbildung 4 (links) dargestellte Konstruktion zeigt einen Abschnitt eines Schienensegmentes mit Modul zur Höhen- und Winkelanpassung, auf welches ein mobiler Sensor montiert werden kann. Der mobile Sensor besteht aus Tiefenbildkamera + Motoreinheit und ist jeweils mit Kontrollern zur Kommunikation ausgestattet. Jeder Kamerakontroller kommuniziert die erfassten Tiefenbilder an eine zentrale Recheneinheit, wo für empfangene Bildsequenzen eine Skelettgenerierung vorgenommen wird, insofern menschliche Bewegungen darin zu erkennen sind. Die Kommunikationsstruktur ist schematisch in Abbildung 4 (rechts) dargestellt.

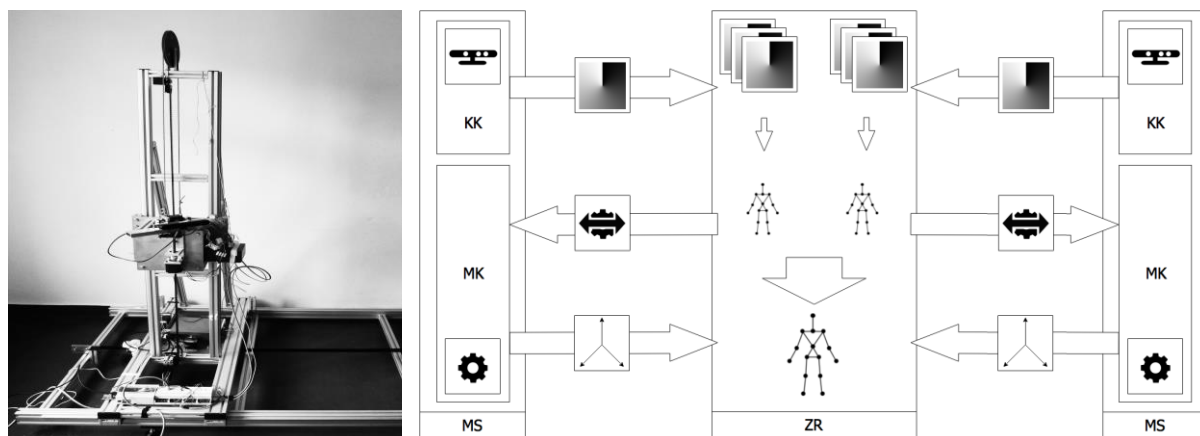


Abbildung 4: Links: Prototyp eines mobilen Schienenelementes mit Höhen-, Seiten- und Neigungs- sowie Rotationsmotoren. Rechts: Kommunikationsstruktur zwischen den Komponenten im System, bestehend aus mobilem Sensor (MS), Tiefenkamera, Motor und Steuereinheiten für Kamerakontroller (KK) und Motorkontroller (MK). Die Kommunikation verläuft über einen zentralen Rechner (ZR).

Erfassen mehrere Sensoren eine Person zur gleichen Zeit, so werden Skelettdaten anhand der Position, die vom Motorkontroller auf Anfrage zurückgegeben wird, normalisiert und im globalen Koordinatensystem zusammengeführt. Bewegt sich die aufgenommene Person bzw. dessen erzeugtes Skelett an die Grenzen des Aufnahmebereiches, empfangen die Motorkontroller ein entsprechendes Steuersignal und verfolgen die aufgenommene Person automatisch nach. Die Position der Skelettpunkte im globalen Koordinatensystem wird dabei permanent rekonstruiert.

Mit dem entwickelten Messaufbau haben sich die Leistungsmöglichkeiten der verwendeten Tiefenbildkameras deutlich verbessert. Das größere Aufnahmefeld, sowie die steigende Genauigkeit von anthropometrischen Messungen geben ausreichend Motivation für eine Anwendung ergonomischer Analysen von Arbeitstätigkeiten.

4. Umsetzung von Ergonomieverfahren am Beispiel der Erfassung von Lendenwirbelsäulenbelastungen

Ein erster Ansatz wurde mit der softwareseitigen Implementierung eines Verfahrens zur Ermittlung von haltungsabhängigen Druckkräften in der Lendenwirbelsäule nach Jäger et al. (2010) umgesetzt. Aktuell lassen sich mit dem System manuelle Lastenhandhabungsfälle aufnehmen und daraus Belastungen der Lendenwirbelsäule beim Heben oder Umsetzen von Gegenständen über spezielle Algorithmen zeitdiskret erfassen. Dabei werden Abhängigkeiten von Hubweg, Lastmasse und Ausführungsdauer berücksichtigt. Weitere Einflüsse, wie ein- oder beidhändige Ausführung, Heberichtung oder Körperdrehung fließen ebenfalls in die Berechnung ein. Eine Erfassung der genannten Kenngrößen und der Körperhaltung mit Hilfe des oben beschriebenen Messsystems lässt nun Rückschlüsse zu auftretenden Risiken in Folge von Lastenhandhabung über die Zeit (z. B. eine Arbeitsschicht) zu. Die gewonnenen Belastungswerte werden dazu mit Belastbarkeitsgrenzwerten (altersabhängige axiale Kompressionsfestigkeit, sog. „Dortmunder Richtwerte“) verglichen.

Da sich die Körperhaltungen verfahrensseitig größtenteils einfach über relative Positionen der Hände zu Boden, Knie, Hüfte oder Schulter definieren lassen und somit keine genauen Winkelmessungen notwendig sind, ist die automatisierte Auswertung dieser Methode mit dem Tiefenbild-basierten Messsystem sehr gut umsetzbar.

5. Diskussion und Ausblick

Das entwickelte System ist dazu in der Lage, computergestützte ergonomische Analysen ohne die Anbringung von Markern an der arbeitenden Person durchzuführen. Der Einsatz erfolgt mittels Verwendung von Tiefenbildkameras, mit denen sich Daten erheben lassen, die einer Skelettgenerierung zugeführt werden können.

Die Erzeugung der Skelettdaten ist derzeit jedoch nur unter bestimmten Bedingungen durchführbar und auf 15 Gelenkpunkte beschränkt. Ebenso werden derzeit bspw. keine Informationen zur Kopfdrehung bereitgestellt. Da die verwendete Kamera jedoch auch über eine übliche Farbwertkamera verfügt, ist es vorstellbar mittels Gesichtserkennung Informationen dazu zu gewinnen. Untersuchungen, wie auf Basis solcher Datensätze auch eine Kopfdrehung ermittelt werden kann, stehen in Aussicht. Dies würde Spielraum für die Implementierung weiterer Ergonomieverfahren, wie OWAS oder RULA schaffen.

6. Literatur

- Fritzsche L, Jendrusch R, Leidholdt W, Bauer S, Jäckel T, Pirger A (2011) Introducing ema (editor for manual work activities) - A new tool for enhancing accuracy and efficiency of human simulations in digital production planning. Springer, 272-281.
- Jäger M, Göllner R, Jordan C, Theilmeier A, Luttmann A (2010) Belastung der Lendenwirbelsäule beim Heben und Umsetzen von Lasten. GRIN Verlag.
- Kamusella C, Scherstjanoi E, Schmauder M (2015) Ergotyping tools for ergonomic research on human-machine interfaces in Digital Prototyping. IOS Press, 97-107.
- Kopniak P (2015) Motion capture using multiple Kinect controllers. Przegląd Elektrotechniczny, 28-31.
- Lemkens W, Kaur P, Buys K, Slaets P, Tuytelaars T, De Schutter J (2013) Multi RGB-D camera setup for generating large 3D point clouds. IEEE, 1092-1099.