



TECHNISCHE
UNIVERSITÄT
WIEN

DISSERTATION

Spatial Augmented Reality in der industriellen Großgeräte- und Baustellenmontage: Entwicklung eines Assistenzsystems mit dynamischer In-situ-Projektion und nutzeradaptiver Gesteninteraktion

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines

Doktors der technischen Wissenschaften

unter der Leitung von

Univ.-Prof. Dr.-Ing. Sebastian Schlund

Technische Universität Wien, Institut für Managementwissenschaften
Forschungsbereich Mensch-Maschine-Interaktion

unter Begutachtung von

Univ.-Prof. Dr.-Ing. Oliver Riedel

Universität Stuttgart, Institut für Arbeitswissenschaft und Technologiemanagement (IAT)

Univ.-Prof. Dr.-Ing. Robert Weidner

Universität Innsbruck, Institut für Mechatronik, Professur für Fertigungstechnik

eingereicht an der Technischen Universität Wien

Fakultät für Maschinenwesen und Betriebswissenschaften

von

Patrick Rupprecht, MSc MSc MA

Matr.-Nr.: 01446160

Aigen, im Dezember 2021

Patrick Rupprecht



TECHNISCHE
UNIVERSITÄT
WIEN

Ich habe zur Kenntnis genommen, dass ich zur Drucklegung meiner Arbeit unter der Bezeichnung

Dissertation

nur mit Bewilligung der Prüfungskommission berechtigt bin.

Ich erkläre weiters Eides statt, dass ich meine Dissertation nach den anerkannten Grundsätzen für wissenschaftliche Abhandlungen selbstständig ausgeführt habe und alle verwendeten Hilfsmittel, insbesondere die zugrunde gelegte Literatur, genannt habe.

Weiters erkläre ich, dass ich dieses Dissertationsthema bisher weder im In- noch Ausland (einer Beurteilerin/ einem Beurteiler zur Begutachtung) in irgendeiner Form als Prüfungsarbeit vorgelegt habe und dass diese Arbeit mit der von den Begutachtern beurteilten Arbeit inhaltlich übereinstimmt.

Aigen, im Dezember 2021

Patrick Rupprecht

Vorwort und Danksagung

Die Arbeit entstand während meiner Tätigkeit als wissenschaftlicher Mitarbeiter am Institut für Managementwissenschaften im Forschungsbereich Mensch-Maschine-Interaktion (Stiftungsprofessur für Industrie 4.0 – „Human Centered Cyber Physical Production and Assembly Systems“) an der Technischen Universität Wien.

Ich möchte mich hiermit bei all meinen Kollegen und Kolleginnen des Fachbereichs „MMI“, des Instituts „IMW“ und der TU Wien-Pilotfabrik für die spannende Zeit bedanken.

Ein besonderer Dank gebührt meinem Doktorvater Univ.-Prof. Dr.-Ing. Sebastian Schlund, als Ideengeber und Mentor im gesamten Doktorat sowie als kritischer Begutachter dieser Arbeit. Danke für dein ehrliches und konstruktives Feedback zur vorliegenden Dissertation.

Des Weiteren möchte ich mich bei den Gutachtern Univ.-Prof. Dr.-Ing. Oliver Riedel und Univ.-Prof. Dr.-Ing. Robert Weidner für das Feedback zur Optimierung der Arbeit bedanken.

Herrn Dr. Walter Mayrhofer gilt ein besonderer Dank für die zahlreichen Ideen zum Demonstrator und für die fachlichen Diskussionen als Vis-a-vis-Sitznachbar. Ebenso gilt ein großes Dankeschön dem Team „DYPRO“, Hans Küffner-McCauley, Majesa Trimmel, David Kostolani und Michael Hornacek, für die tatkräftige Unterstützung beim Aufbau des Demonstrators sowie bei der Durchführung der Experimente in der Pilotfabrik.

Abschließend bedanke ich mich ganz herzlich bei meinen Eltern und Großeltern, meiner Schwester und im Speziellen bei meiner Freundin Natalie für die Unterstützung und Motivation während des Doktorats sowie beim Verfassen dieser Dissertation.

Diese Arbeit widme ich meinem Großvater Franz Ritter.

Kurzfassung

Die industrielle Großgeräte- und Baustellenmontage ist aufgrund der großen Montageobjekte, der verschiedenen Bereitstellungsflächen und der geringen Losgrößen geprägt durch einen hohen manuellen Arbeitsanteil, einen geringen Automatisierungsgrad und den Bedarf an qualifizierten Mitarbeiter*innen. Der große Arbeitsraum dieser Montageform erfordert, über die Kenntnis der einzelnen Montagetätigkeiten hinausgehend, ebenso das Wissen über die Standorte der einzelnen Montageobjekte und Komponenten sowie die Information über deren definierte Montagepositionen. Demnach besteht der Bedarf an einem geeigneten Assistenzsystem zur Informationsbereitstellung, das klar strukturierte Arbeitsanweisungen über einen großen Arbeitsbereich anzeigt, den Menschen gezielt durch den Montageprozess führt und durch eine intuitive Interaktionsmöglichkeit gesteuert werden kann.

Aktuelle Forschungen beschäftigen sich mit kognitiven Assistenzsystemen zur effektiven Informationsbereitstellung und zur Verbesserung der Mensch-Maschine-Interaktion, allen voran unter dem Einsatz von Spatial Augmented Reality, welche bereits in der Kleingerätemontage erhebliches Potenzial zeigt. Da in der Großgerätemontage aktuell noch kaum vergleichbare Assistenzsysteme eingesetzt werden und bestehende Konzepte sich auf Montagestationen mit kleinem Arbeitsbereich und statischen Projektoren beziehen, stellt die Entwicklung eines Spatial Augmented Reality Assistenzsystems mit dynamischer Projektion und nutzeradaptiver Steuerung speziell für die Anforderungen der Großgeräte- und Baustellenmontage den Kern der vorliegenden Forschungsarbeit dar.

Dazu wurde ein aus der Veranstaltungstechnik etabliertes dynamisches Projektionssystem für die Anwendung in der industriellen Montage adaptiert, ein Spatial Augmented Reality Assistenzsystem in einem iterativen Designprozess entwickelt und mit Demonstratoren in der Pilotfabrik der TU Wien getestet. Dabei wurden verschiedene Designelemente der intuitiven und projektionsgerechten Informationsgestaltung, der dynamischen In-situ-Projektion und Informationsführung sowie der nutzeradaptiven Gesteninteraktion mittels Kamerasystemen und Deep Learning evaluiert. Die Usability, kognitive Arbeitsbelastung sowie Prozesseffizienz wurde anhand von Experimenten bewertet und der Mehrwert im Vergleich zu herkömmlichen Assistenzsystemen (PC-Terminal, Tablet) festgestellt. Durch die gezielte In-situ-Projektion von Arbeitsinformationen und dynamischer Führung mittels des Projektionssystems konnte eine erhebliche Verbesserung der Montage- und Lokalisierungszeit erreicht und gleichzeitig die kognitive Arbeitsbelastung verringert werden. Des Weiteren zeigte sich, dass der Einsatz des Spatial Augmented Reality Assistenzsystems eine gute Usability aufweist und auf verschiedenste Anwendungsfälle in der Großgeräte- und Baustellenmontage skalierbar ist.

Die Arbeit präsentiert einen wissenschaftlichen und praxisrelevanten Beitrag zur Gestaltung eines Spatial Augmented Reality Assistenzsystems mit dynamischer In-situ-Projektion zur Informationsbereitstellung für die industrielle Großgeräte- und Baustellenmontage.

Schlüsselwörter: Spatial Augmented Reality, dynamische In-situ-Projektion, nutzeradaptive Gesteninteraktion, industrielle Großgeräte- und Baustellenmontage

Abstract

Industrial large-scale equipment and site assembly is characterized by a high degree of manual work, a low level of automation and the need for qualified employees due to the large assembly objects, the various supply areas and small batch sizes. The resulting large workspace requires, in addition to the knowledge of the individual assembly activities, also the information about the locations of the individual assembly objects and components as well as the specific assembly positions. Accordingly, there is a need for a suitable information assistance system that dynamically displays clearly structured work instructions over a large work area, guides humans through the assembly process and can be controlled through an intuitive interaction possibility.

Current research is addressing cognitive assistance systems for the effective provision of information and the improvement of human-machine interaction, especially with the use of Spatial Augmented Reality, which already shows considerable potential in the assembly of small devices.

Since comparable assistance systems are currently barely used in large-scale equipment assembly and existing concepts refer to assembly stations with a small workspace and static projectors, the development of a spatial augmented reality assistance system with dynamic projection and user-adaptive control, especially for the requirements of large-scale equipment and site assembly, represents the central core of the thesis.

Therefore, a dynamic projection system established in entertainment technology was adapted for the application in industrial assembly, a Spatial Augmented Reality assistance system was developed in an iterative design process and tested with demonstrators in the pilot factory of the TU Wien. In the process, various elements of intuitive and projection-oriented information designs, dynamic in-situ projection and information guidance, as well as user-adaptive gesture interaction using camera systems and Deep Learning were evaluated. The usability, cognitive workload as well as process efficiency of the systems were evaluated by means of experiments and the added value compared to conventional systems (PC terminal, tablet) was determined. Through the targeted in-situ projection of work information and dynamic guidance by means of a projection system, a significant improvement in assembly and localization time of mounting positions was achieved, while at the same time reducing cognitive workload. Furthermore, it was shown that the application of the spatial augmented reality assistance system has a good usability and is scalable to various use cases in the assembly of large-scale equipment and site assembly. The work presents a scientific as well as practical contribution to the design of a Spatial Augmented Reality assistance system with dynamic in-situ projection for information provision for industrial large-scale equipment and site assembly.

Keywords: Spatial Augmented Reality, dynamic in-situ projection, user-adaptive gesture interaction, industrial site assembly

Inhaltsverzeichnis

Abkürzungsverzeichnis	IX
1. Einleitung	1
1.1 Ausgangssituation	1
1.2 Kontext und Relevanz des Themas	4
1.2.1 Informationsbereitstellung in der industriellen Großgeräte- und Baustellenmontage	4
1.2.2 Dynamisches Projektionssystem	7
1.2.3 Nutzeradaptive Mensch-Maschine-Interaktion mit natürlichen Gesten	8
1.3 Forschungsbedarf und Zielsetzung	10
1.4 Forschungsmethodik – Design Science Research	12
1.5 Aufbau und Struktur der Arbeit	16
2. Grundlagen und Stand der Wissenschaft	18
2.1 Systematische Literaturrecherche	18
2.2 Nutzeradaptive Gestaltung und Assembly 4.0	23
2.2.1 Menschzentrierte Gestaltung und Operator 4.0	23
2.2.2 Adaptive Einstellmöglichkeiten und Assembly 4.0	27
2.2.3 Vision der nutzeradaptiven Gestaltung in der Großgerätemontage	32
2.3 Spatial Augmented Reality zur Informationsbereitstellung	34
2.3.1 Spatial Augmented Reality in der Montage	36
2.3.2 Informationsbereitstellung mit Spatial Augmented Reality	41
2.3.3 Dynamische In-situ-Projektion und Informationsführung	45
2.3.4 Echtzeitfähige Interaktion mit Computer Vision und Deep Learning	48
2.4 Nutzeradaptive Mensch-Maschine-Interaktion	54
2.4.1 Natürliche Mensch-Maschine-Interaktion	56
2.4.2 Nutzeradaptive Gesteninteraktion	57
2.5 Zusammenfassung der Literaturrecherche	62
3. Spatial Augmented Reality Assistenzsystem mit dynami-scher In-situ-Projektion und nutzeradaptiver Gesteninteraktion	64
3.1 Konzeptidee – Informationsassistentz in der Baustellenmontage	64

3.2 Technische Gestaltung und Aufbau des Spatial Augmented Reality Assistenzsystems	68
3.2.1 Dynamisches Projektionssystem	69
3.2.2 Visuelles Interaktionssystem	79
3.3 Gestaltung der dynamischen Informationsbereitstellung und Mensch-Maschine-Interaktion	84
3.3.1 Intuitive und projektionsoptimierte Informationsgestaltung	85
3.3.2 Dynamische In-situ-Projektion und Informationsführung	89
3.3.3 Nutzeradaptive Gesteninteraktion	95
4. Demonstration von Spatial Augmented Reality in der industriellen Großgeräte- und Baustellenmontage	100
4.1 Demonstratoren und Use Cases	101
4.2 Projektion von Arbeitsinformationen im industriellen Umfeld	103
4.2.1 Herausforderungen bei projizierten Arbeitsinformationen	103
4.2.2 Projektionsoptimierte Arbeitsinformationen	107
4.3 Dynamische In-situ-Projektion und Informationsführung	112
4.3.1 Dynamische In-situ-Projektion der Montageposition	113
4.3.2 Dynamische In-situ-Projektion der Entnahme- und Montageposition	115
4.4 Nutzeradaptive und natürliche Gesteninteraktion	116
4.4.1 Nutzeradaptive Interaktion durch Positionserkennung des Menschen	117
4.4.2 Nutzeradaptive Gesteninteraktion	117
5. Evaluierung des Spatial Augmented Reality Assistenzsystems	120
5.1 Evaluierungsmethoden	120
5.1.1 Usability, Technologieakzeptanz und kognitive Arbeitsbelastung	121
5.1.2 Montage- und Taskzeit (General Assembly Task Model)	123
5.1.3 Fehlerrate	127
5.1.4 Qualitative Evaluierungsmethoden	127
5.2 Experimente und Use Cases	128
5.2.1 Experiment 1 – Fan-Cowl-Montage: Informationsbereitstellung über ein Projektionssystem, Interaktion mittels projizierten Buttons	129
5.2.2 Experiment 2 – Fan-Cowl-Montage: Dynamische In-situ-Projektion mit Informationsführung zur Montageposition, Interaktion mittels Gestensteuerung	132

5.2.3 Experiment 3 – Spielzeugbaustein-Montage: Dynamische In-situ-Projektion mit Informationsführung zur Entnahme- und Montageposition, Interaktion mittels Gestensteuerung	140
5.3 Zusammenfassung der Ergebnisse der Evaluierung	149
6. Diskussion der Ergebnisse	153
6.1 Erkenntnisse aus dem Design der Artefakte	153
6.2 Erkenntnisse aus der Gestaltung der dynamischen Informationsbereitstellung	156
6.3 Erkenntnisse aus der Gestaltung der Gesteninteraktion	158
6.4 Limitationen der Ergebnisse und weiterer Forschungsbedarf	160
6.5 Kritische Diskussion in Bezug auf die Forschungsfragen	163
7. Dissemination und Zusammenfassung	164
Abbildungsverzeichnis	166
Literaturverzeichnis	171
Anhang	187

Abkürzungsverzeichnis

AR	Augmented Reality
CNN	Convolutional Neural Network
GATM	General Assembly Task Model
GUI	Grafical User-Interface
FOV	Field of View
fps	frames per second
H-CPS	Human-Cyber-Physical System
HCI	Human-Computer Interaction
HITL	Human-in-the-Loop
HITM	Human-in-the-Mesh
NASA-(R)TLX	NASA-(Raw) Task Load Index
NHMI	Natürliche Mensch-Maschine-Schnittstelle
HMD	Head-Mounted Displays
mAP	Mean Average Precision
MMI	Mensch-Maschine-Interaktion
MTM	Methods Time Measurement
SAR	Spatial Augmented Reality
SLR	Systematischen Literaturrecherche
SUS	System Usability Scale
TAM	Technology Acceptance Model
TCT	Task Completion Time

1. Einleitung

“Together with the real-time interaction feature, the dynamic projection system, which is adaptively controlled by movements and natural gestures of the user, represents the ‘Adaptive Spatial Augmented Reality System’ in industrial site assembly” (Rupprecht et al., 2021).

Die vorliegende Forschungsarbeit beschäftigt sich mit der Entwicklung eines Spatial Augmented Reality Assistenzsystems mit dynamischer In-situ-Projektion und nutzeradaptiver Gesteninteraktion im Kontext der industriellen Großgeräte- und Baustellenmontage. Dabei wird ein dynamisches Projektionssystem für die Anforderungen der Montage adaptiert und eine nutzeradaptive Gesteninteraktion mittels Kamerasystemen und Bilderkennung implementiert. Des Weiteren wird durch die Gestaltung der Informationsbereitstellung und Mensch-Maschine-Interaktion ein Mehrwert für Mensch und Unternehmen geschaffen.

Spatial Augmented Reality (SAR), deutsch „**räumlich erweiterte Realität**“, verzichtet auf das Tragen von Devices am Körper (Funk, 2016; Mengoni et al., 2018) und verwendet stattdessen statische Projektionssysteme für die Informationsanzeige (Funk, 2016). Jedoch erst durch die Integration der dynamischen Komponente der Projektion, durch ein steuerbares Spiegelsystem und die nutzeradaptive Interaktion trägt das Assistenzsystem hohes Potenzial für den Einsatz zur Informationsbereitstellung in der Großgeräte- und Baustellenmontage (Rupprecht et al., 2020; Rupprecht et al., 2021).

1.1 Ausgangssituation

Der steigende technologische und organisatorische Wandel sowie die Zunahme an Dynamik, Komplexität und Varianz von Produkten und Prozessen, ausgelöst von individuell gestaltbaren Produkten, kürzeren Lebenszyklen und neuen Technologien, stellt Produktionsunternehmen und deren Mitarbeiter*innen vor allem in der Montage vor neue Herausforderungen (Dombrowski et al., 2013; Mayrhofer, Kames & Schlund, 2019; OECD, 2019; Spath, 2013; Vogel-Heuser et al., 2017; Wichmann et al., 2019; Wiendahl et al., 2007).

Globale Trends in den Bereichen Digitalisierung und Technologie, wie das „Internet der Dinge (IoT)“ (Atzori et al., 2010; Stankovic, 2014), cyberphysische Systeme (acatech, 2011; Niesen et al., 2016) und cyberphysische Produktionssysteme (Monostori, 2014), verändern aktuell große Bereiche der Arbeitswelt in produzierenden Unternehmen. Vor allem trägt der Einsatz von modernen Technologien und Assistenzsystemen, wie (Spatial) Augmented Reality, kollaborativen Robotern und „Machine & Deep Learning“, großes Potenzial zur Unterstützung des Menschen bei der Ausführung der Arbeitstätigkeiten (Wichmann et al., 2019).

Arbeitsplätze mit simplen, repetitiven Tätigkeiten (Schweißen, Pick & Place etc.) werden heutzutage schon mit Automatisierungslösungen ausgestattet, wobei vollautomatische Industrieroboter die menschliche Arbeitskraft teils komplett ersetzen (IFR International Federation of Robotics, 2019).

In der **manuellen Montage mit komplexen Tätigkeiten** ist eine Automatisierung jedoch mit hohem Aufwand und Kosten verbunden und beeinträchtigt dadurch die notwendige Flexibilität. Demnach ist der Einsatz von qualifizierten Beschäftigten, welche dynamisch auf Änderungen reagieren und somit die Erfordernis der Flexibilität und Komplexität bewerkstelligen können, unabdingbar (Gorecky et al., 2014; Jensen & Alting, 2006; Lotter & Wiendahl, 2012). Der Mensch stellt demnach mit seinen spezifischen physischen und kognitiven Fähigkeiten wie Kreativität, Problemlösungskompetenz und Flexibilität trotz aufstrebender Automatisierung eine zukünftige Rolle in der Montage dar (Latos et al., 2018; Reinhart, 2017; Spath, 2013; Weisner et al., 2016).

Vor allem stellt der Einsatz von **kognitiven Assistenzsystemen** (z. B. Informationsbereitstellung mit Augmented Reality) hohes Potenzial dar, welche den Menschen und dessen individuelle Fähigkeiten bei Arbeitsplätzen mit hohem manuellen Arbeitsanteil unterstützen und ergänzen (Romero et al., 2016; Schlund & Baaij, 2018; Spath, 2013; Weidner et al., 2015). Augmented Reality Systeme, Kamerasysteme u. Ä. sind mittlerweile preisgünstig erhältlich, schnell verfügbar und vor allem industriell einsetzbar. Auch die intelligente und leistungsfähige Verarbeitung von Daten durch spezielle Algorithmen (Schmidt et al., 2015) und der Einsatz von maschinellem Lernen (Lee et al., 2015) zur intelligenten Objekterkennung zu Tracking-, Qualitäts- oder Montagezwecken ist bereits ein essenzieller Faktor im Produktions- und Montagesektor. Schließlich verbessert die Zugänglichkeit zu Softwaresystemen und Hochleistungscomputern die Umsetzung moderner Entwicklungen (Eckert et al., 2019). Der Einsatz von Assistenzsystemen speziell mit Fokus auf den Menschen ist jedoch erst am Anfang und trägt demnach großes Potenzial für weitere Forschungsaktivitäten (IFR International Federation of Robotics, 2019).

Demnach beschäftigen sich aktuelle Forschungen mit kognitiven Assistenzsystemen zur Informationsbereitstellung und zur Verbesserung der Mensch-Maschine-Interaktion. Diese Entwicklungen sind u. a. Informationssysteme mit Touchbildschirmen direkt am Arbeitsplatz, (Gorecky et al., 2013; Günthner et al., 2009; Lušić et al., 2016; Paelke, 2014) Assistenzsysteme in Form von intelligenten Wearables zum Tragen am Körper oder **Augmented Reality (AR) Systeme** (Azuma, 1997), um die bestehende Realität mit Informationen zu erweitern und eine Interaktion zu ermöglichen (Büttner et al., 2017; Gorecky et al., 2013; Schembera-Kneifel & Keil, 2016). Unternehmen und Forschungseinrichtungen stecken aktuell erhebliche Ressourcen in die anwendungsorientierte Entwicklung von Augmented Reality im industriellen Umfeld (Masood & Egger, 2019). Datenbrillen, Head-Mounted Displays (HMD) (Paelke et al., 2015) oder AR-Brillen werden immer leistungsfähiger und leichter, wobei hier jedoch die Herausforderung besteht, dass diese Devices ein kleines Anzeigefeld von nur einigen Zentimetern besitzen sowie der Tragekomfort noch mangelhaft ist. Diese Devices werden vor allem bei der Fernwartung und bei Instandhaltungstätigkeiten, jedoch kaum in der Montage eingesetzt (Heinz et al., 2018; Hořejší, 2015). Weitere AR-Systeme sind u. a. Pick-by-Vision-, Pick-by-Light-, Pick-by-Voice-Systeme, bei denen mit

gezielter visueller oder akustischer Informationsbereitstellung zu einzelnen Montageschritten Anweisungen gegeben werden (Günthner et al., 2009; Picavi GmbH, 2020; Theel, 2015).

Bei **Spatial Augmented Reality Assistenzsystemen** werden bei Montagearbeitsplätzen mittels überkopf-angeordneter Projektoren Arbeitsinformationen am richtigen Ort und in der richtigen Lage (in situ) auf die Arbeitsfläche projiziert und die Mensch-Maschine-Interaktion mittels visueller Erkennung über ein Kamerasystem umgesetzt. Hier bestehen bereits fundierte Forschungsarbeiten im Bereich der Einzelplatzmontagen, wo speziell **Spatial Augmented Reality Assistenzsysteme mit statischen Projektionssystemen** verwendet werden (Bertram et al., 2018; Büttner et al., 2016; Funk, 2016; Sand et al., 2016). Diese SAR-Systeme tragen einen deutlichen Mehrwert zur Verbesserung der Prozessleistung sowie zur Verringerung der Prozesszeit und Fehlerrate bei, was Funk, 2016 in seinen Vergleichen mit Head-Mounted Displays und Papierdokumenten gezeigt hat. Aktuelle vorherrschende SAR-Systeme mit statischen Projektionssystemen können jedoch aktuell nur dort eingesetzt werden, wo ein kleiner Arbeitsbereich vorherrschend ist, wie z. B. bei der Einzelplatzmontage, weshalb sie bei großen Arbeitsräumen an ihre Grenzen stoßen.

Eine der Hauptherausforderungen liegt jedoch meist nicht bei der erwähnten Verfügbarkeit und immer preisgünstigeren Systemen, sondern vielmehr in der aufwändigen Integration von und Adaptierung an individuelle Unternehmens- und Prozessanforderungen sowie den Bedürfnissen der Mitarbeiter*innen (Rodriguez et al., 2015). Zudem steigt mit jedem neuen System auch die Komplexität – mit Auswirkungen auf die Prozesse sowie die Produktionsarbeit in der Fertigung und Montage (Brinzer et al., 2017; Coletti & Aichner, 2011; Mattsson et al., 2018). Um diesen Herausforderungen entgegenzuwirken, müssen Arbeitsplätze in der Montage iterativ angepasst werden und neue Systeme unter Einbezug der Mitarbeiter*innen nachhaltig implementiert werden (Eckert et al., 2019; Gaham et al., 2015; Hermann et al., 2016). Eine Möglichkeit, menschliche Merkmale sinnvoll in die Gestaltung von Arbeitssystemen und Entwicklungen von Assistenzsystemen zu integrieren, stellt der menschenzentrierte Gestaltungsansatz nach der DIN EN ISO 9241-210 dar, welcher mit einer iterativen Vorgehensweise den definierten Nutzungskontext analysiert, die Anforderungen ermittelt und geeignete und gebrauchstaugliche Gestaltungslösungen entwickelt. In der vorliegenden Arbeit wird dazu die iterative Gestaltungsmethode „Design Science Research“ verwendet, um ein Assistenzsystem für den Kontext der manuellen Montage und entsprechend den Anforderungen der User zu gestalten, zu demonstrieren und zu evaluieren.

In den folgenden Kapiteln werden die Problemstellung und Relevanz des Themas erläutert und die Forschungsbereiche, Forschungsfragen und Ziele abgeleitet. Die Darstellung der Forschungsmethode der Arbeit sowie die Vorgehensweise und Struktur wird anschließend präsentiert.

1.2 Kontext und Relevanz des Themas

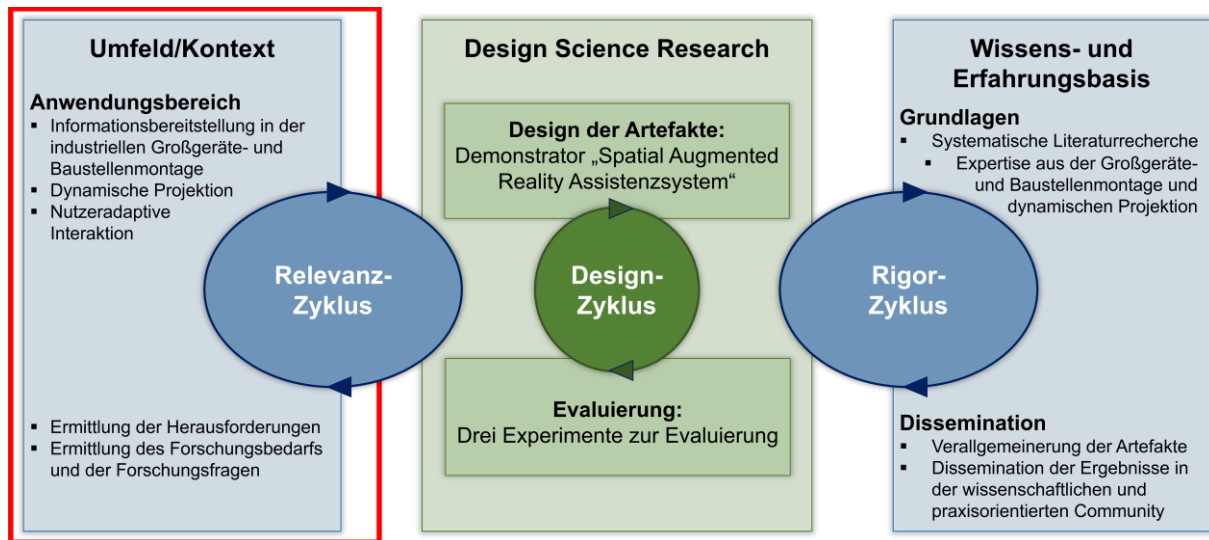


Abbildung 1 | Kontext und Relevanz des Themas, in Anlehnung an Alan R. Hevner et al. (2004)

Die Analyse des Umfelds der Forschungsarbeit (vgl. Abbildung 1) ist ein wesentlicher Teil der Design-Science-Research-Methode (Alan R. Hevner et al., 2004), welche in Kapitel 1.4 detailliert dargestellt wird. Dabei werden die Relevanz des Themas ermittelt, die Problemstellung und Herausforderung der Anwendungsbereiche dargestellt und ein mögliches Lösungspotenzial aufgezeigt, welches über den **Relevanz-Zyklus** in die Forschung eingebracht werden. Die industrielle Großgeräte- und Baustellenmontage sowie dessen Herausforderungen in der Informationsbereitstellung, mögliche Lösungsansätze der dynamischen Projektion sowie der nutzeradaptiven Interaktion stellen dabei den Kontext dar.

1.2.1 Informationsbereitstellung in der industriellen Großgeräte- und Baustellenmontage

Der Anwendungsbereich der Arbeit bezieht sich auf die Charakteristika der manuellen Montage in der produzierenden Industrie und die Unterstützung des Menschen durch den Einsatz eines Assistenzsystem zur Informationsbereitstellung. Im Speziellen wird auf die Montageform der **Großgeräte- und Baustellenmontage** eingegangen, welche hohes Potenzial für die Entwicklung eines Assistenzsystems darstellt.

Industriell hergestellte Produkte bestehen in der Regel aus einer Vielzahl von Einzelteilen, Komponenten und Baugruppen, die zu unterschiedlichen Zeitpunkten mit unterschiedlichen Fertigungsverfahren hergestellt werden und am Ende des Wertschöpfungsprozesses in der Montage zu einem finalen Produkt assembliert werden. Die Aufgabe der Montage ist, aus diesen Teilen ein Produkt höherer Komplexität mit vorgegebenen Funktionen in einer bestimmten Zeit zusammenzubauen (VDI, 1978, 1990; Warnecke et al., 1975). Lotter und Wiendahl (2012) unterscheiden die Montage nach dem Aspekt „Automatisierungsgrad“ und teilen diese in die Varianten manuelle Montage, halbautomatische (hybride) Montage und automatische Montage. Dabei stellt die Großgeräte- und Baustellenmontage die Montageart der manuellen Montage dar, welche durch hohe Flexibilität, hoher Produktvarianz sowie einen

geringen Automatisierungsgrad geprägt ist und im Vergleich zu anderen Formen nur niedrige Stückzahlen und geringe Losgrößen bewältigen kann. Zur Umsetzung der manuellen Tätigkeiten sowie zum Handling der Flexibilität und Varianz ist der Mensch ein entscheidender Handlungsakteur in dieser Form der Montage (Lotter & Wiendahl, 2012).

Die Montage von großen Bauteilen und Objekten (z. B. Flugzeugkomponenten, Zügen, Großmaschinen, Getriebe etc.) wird üblicherweise in der Montageorganisationsform der Baustellenmontage (Abbildung 2) durchgeführt (Lotter & Wiendahl, 2012).

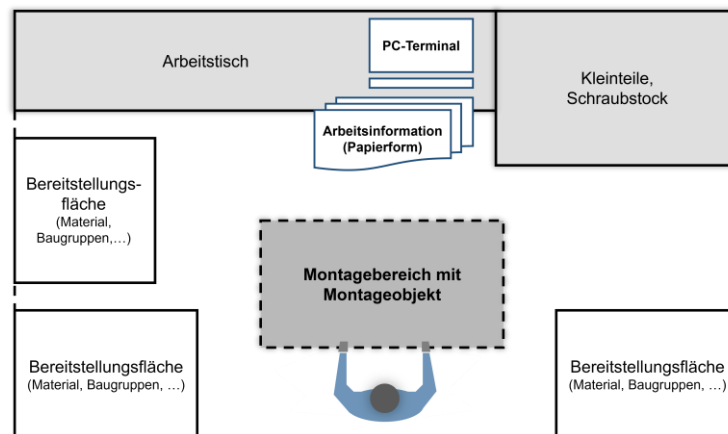


Abbildung 2 | Schematische Darstellung der Montageorganisationsart „Baustellenmontage“, in Anlehnung an Lotter und Wiendahl (2012) und Rupprecht et al. (2020)

In der Baustellenmontage bleibt das Montageobjekt während des kompletten Montageprozesses stationär und alle Komponenten, Materialien, Baugruppen, Werkzeuge, Informationen und Mitarbeiter*innen begeben sich zur Ausführung der Montageaktivitäten zum Bauteil. Die Anzahl der Arbeitsplätze in der industriellen Baustellenmontage beträgt konservativ geschätzt ca. 50 000 der rund 270 000 Industriearbeitsplätze im Bereich der Großproduktfertigung in Österreich (Statistik Austria, 2019). Abbildung 3 zeigt beispielhaft Montagearbeitsplätze für die Herstellung einer Lokomotive sowie einen CFK-Layup-Prozess einer Flugzeugkomponentenmontage.



Abbildung 3 | Links: Montage einer Lokomotive, rechts: CFK-Layup-Prozess einer Flugzeugkomponente (Aviation Media & IT GmbH [aero.de], 2021; Siemens AG, 2021)

Die Charakteristika der Baustellenmontage können wie folgt zusammengefasst werden (Lotter & Wiendahl, 2012; Rupprecht et al., 2020):

- hoher manueller Arbeitsanteil (manuelle Montage) und geringer Automatisierungsgrad
- kleine Losgrößen und hohe Flexibilität

- Mensch als zentraler Handlungsakteur, Zusammenarbeit von mehreren Menschen
- hoher Grad an qualifizierten Mitarbeiter*innen erforderlich
- großer Arbeitsbereich resultierend aus großen Montageobjekten und verschiedenen Bereitstellungsflächen von Komponenten
- Informationen über den Montageprozess und die Abfolge der Tätigkeiten
- Informationen über die einzelnen Standorte der Montageobjekte, Komponenten und deren Bereitstellungsflächen
- Informationen über exakte Montagepositionen am Montageobjekt
- strukturierte Unterweisung und Arbeitsinformationen notwendig
- zentrale Informationsbereitstellung aktuell über Papierdokumente oder PC-Terminals

Speziell bei Montagestationen mit hohem manuellen Arbeitsanteil und einem geringen Industrialisierungsgrad ist ein hohes Ausbildungsniveau der Mitarbeiter*innen erforderlich. Zudem sind klar strukturierte Arbeitsanweisungen bzw. intuitiv verständliche Arbeitsinformationen notwendig, um Qualität zu produzieren, Fehler zu vermeiden und effiziente Prozesse umzusetzen (Gorecky et al., 2014; Lotter & Wiendahl, 2012). Zu den Arbeitsinformationen zählen u. a. Arbeitsanweisungen, Dokumente zu Normen und Standards, Unterstützungs- und Lernanweisungen sowie Informationen über Lokalitäten von Bauteilen, Komponenten und genauen Montagepositionen. Der aktuelle Stand der Technik im Bereich der Informationsbereitstellung in der Baustellenmontage ist einerseits der Einsatz von Papierdokumenten und mündlichen Informationen sowie andererseits die Verwendung von zentralen PC-Terminals oder Bildschirmen über der Montagestation. Diese Formen der Informationsbereitstellung führen zu hohen Warte- und Suchzeiten bei der Informationsbeschaffung und -aufklärung und machen zwischen 20 und 50 % des Aufwands aus, woraus sich ein hohes Potenzial zur Produktivitätssteigerung und weiteren Forschungsarbeiten ergibt (Funk, Kosch & Schmidt, 2016; Jost et al., 2017; Lorenz, 2014).

Dementsprechend resultieren aus den Gegebenheiten der Großgerätemontage folgende Herausforderungen zur aktuellen Informationsbereitstellung bzw. Interaktion mit dem Informationssystem:

- lange Wege zur Informationserhaltung und Informationsklärung, bedingt durch zentrale Terminals und zentrale Datenablage, Rückfragen oder nicht aktuelle Arbeitsdokumente
- hoher Aufwand zur Aktualisierung von Papierdokumenten (z. B. Arbeitsanweisungen)
- fehlende (digitale) Informationsbereitstellung und Interaktionsmöglichkeit direkt am Montageobjekt bzw. am Ort der Montagetätigkeit
- fehlende Informationen von Montageobjekten und Montagepositionen direkt im Arbeitsraum bzw. am Ort des Geschehens
- fehlende Führung durch den Montageprozess bzw. großen Arbeitsraum
- Herausforderung bei der Gestaltung, Integration und Vernetzung von modernen Systemen im bestehenden Arbeitsprozess

Ausgehend von den Spezifika in der Großgeräte- und Baustellenmontage kann zusammengefasst werden, dass der Mensch für die Durchführung der manuellen Tätigkeiten eine zentrale Rolle spielt und die aktuelle Informationsbereitstellung mehrere spezifische Herausforderungen mit sich bringt.

Demnach besteht der Bedarf, die Informationsbereitstellung und Interaktion im Arbeitsraum der Großgerätemontage zu verbessern, um damit den Menschen gezielt im Montageprozess zu unterstützen. Eine Lösungsmöglichkeit wäre die Entwicklung eines kognitiven Assistenzsystems, welches klar strukturierte Arbeitsanweisungen über einen großen Arbeitsbereich dynamisch und direkt am Ort des Geschehens anzeigt, den Menschen gezielt durch den Montageprozess führt und durch eine nutzeradaptive Interaktionsmöglichkeit gesteuert werden kann.

1.2.2 Dynamisches Projektionssystem

Aus den ersten Recherchen zum Thema Informationsbereitstellung wurde ein innovatives Projektionssystem aus der Veranstaltungstechnik ermittelt, welches dynamisch Informationen über einen großen Raum projizieren kann. Das dynamische Projektionssystem verwendet dazu einen hochauflösenden Projektor und ein digital steuerbares und bewegbares Spiegelsystem (Mirror-Head) zur Ablenkung des Projektionsstrahles als Kernelement (Dynamic Projection Institut GmbH, 2020a). Dazu wurde ein Patent mit der Patentnummer „DE202013010365U1“¹ und dem Titel „Vorrichtung zum Projizieren eines Bildes“ recherchiert (Dynamic Projection Institute, Herstellungs- und Vertriebs GmbH, 2015). Das Patent beschreibt die technische Vorrichtung zum Projizieren eines bewegten oder stehenden Bildes, welche mit einem motorisierten Ablenkmittel ausgestattet ist und dementsprechend das Licht des Projektors in eine einstellbare Richtung umlenkt. Das Patent beschreibt die Halterung des Projektors, die motorisierte Ablenkungseinrichtung und die zugehörigen Signalverarbeitungseinrichtungen (Abbildung 4).

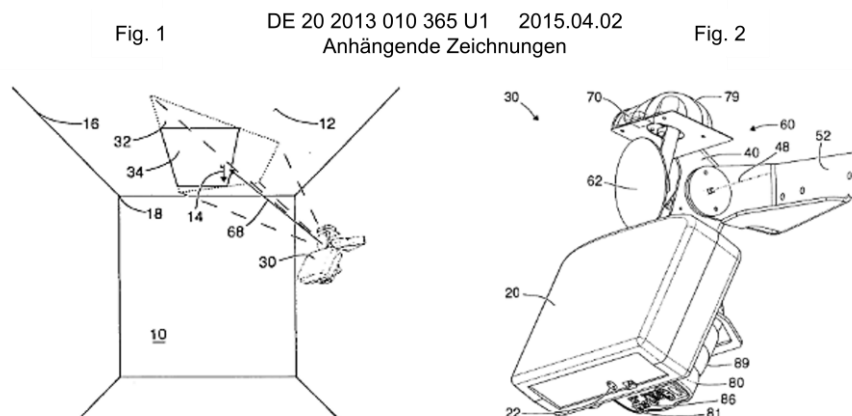


Abbildung 4 | Patentauszug des Projektionssystems mit der Patentnummer DE202013010365U1 (Dynamic Projection Institut GmbH, 2020a)

¹ <https://worldwide.espacenet.com/patent/search/family/052673459/publication/DE202013010365U1?q=DE202013010365U1>.

Der Spiegel besteht aus einem hochpolierten und beschichteten Material, welcher mittels Stepper-Motoren bewegt werden kann. Die Motoren werden digital angesteuert und können dementsprechend eine Pan- und Tilt-Bewegung umsetzen (Abbildung 5).

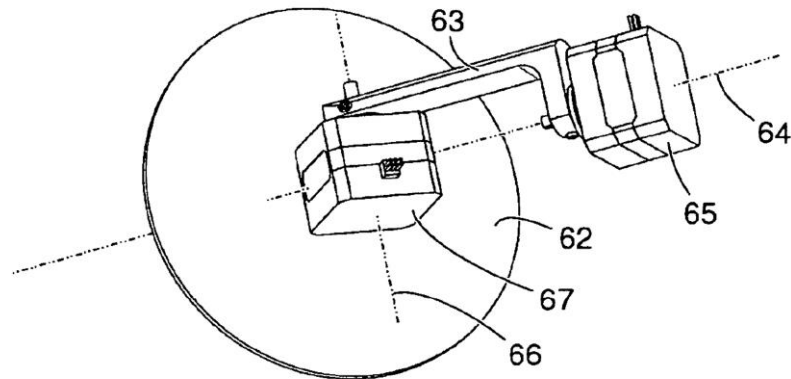


Abbildung 5 | Patentauszug des Projektionssystems mit der Patentnummer DE202013010365U1: Spiegel und Schrittmotoren (Dynamic Projection Institute, Herstellungs- und Vertriebs GmbH, 2015)

Zusammengefasst besteht das Konzept aus einem digital steuerbaren sowie bewegbaren Spiegelkopf (Mirror-Head), welcher den Projektionsstrahl des Projektors in alle Richtungen des Raumes lenken kann. Professionelle Projektoren bieten bereits digitale Schnittstellen und sind mit einer motorbetriebenen Linse ausgestattet, um damit dynamisch den Zoom und Fokus anzupassen (Dynamic Projection Institut GmbH, 2020a). Diese Projektoren sind heutzutage schon unter 10 000 Euro erhältlich und mit wartungsarmen, laserbasierten Leuchtmitteln ausgestattet, welche eine Lebensdauer von über 20 000 Stunden gewährleisten. Eine Helligkeit von 6200 ANSI Lumen, eine Auflösung von Full-HD oder 4K und ein Kontrast von 10 000:1 ermöglicht den Einsatz der Projektoren bei Tageslicht und eine Darstellung von präzisen und scharfen Abbildungen. Diese Spezifikationen ermöglichen somit die Anwendung im industriellen Umfeld zur Anzeige von Informationen direkt am Ort der Montagetätigkeit.

Der Einsatz des Projektionssystems mit Spiegelkopf als dynamisches Spatial Augmented Reality System stellt – adaptiert für die Anforderungen des industriellen Umfelds – eine potenzielle Lösung dar, um Arbeitsinformationen über große Arbeitsbereiche dynamisch und gezielt anzuzeigen. Demnach dient das dynamische Projektionssystem aus der Veranstaltungstechnik als Ausgangslage für die Gestaltung des Spatial Augmented Reality Assistenzsystems und wird im Rahmen des iterativen Designprozesses für die Anforderungen der Informationsbereitstellung in der Großgeräte- und Baustellenmontage angepasst.

1.2.3 Nutzeradaptive Mensch-Maschine-Interaktion mit natürlichen Gesten

Der nachhaltige Einsatz von kognitiven Assistenzsystemen, ausgerichtet am Menschen, dient zur Verbesserung der Arbeitsbedingungen und schließlich zur Steigerung der Ergonomie und Verbesserung der Prozessleistung und stellt demnach eine potenzielle Lösung für manuelle Arbeitstätigkeiten in der Produktion und Montage dar (Reinhart, 2017). Für die optimale

Integration des Menschen in die Interaktion mit dem Assistenzsystem ist es entscheidend, dass darauf geachtet wird, dass der Mensch bei der Ausführung der Tätigkeiten auf sein natürliches Verhalten bzw. seine natürlichen Bewegungen zurückgreifen kann. Demnach besteht die Forderung nach einer intuitiven und natürlichen Auslegung der Mensch-Maschine-Interaktion (z. B. Körper-, Gesten- oder Posenerkennung über Kamerasysteme) sowie die Notwendigkeit, cyberphysische Systeme mit anthropomorphen Dimensionen auszustatten (Garcia et al., 2019; Gorecky et al., 2014). Diese Konzepte werden als „Natural User-Interfaces“ oder „Natürliche Mensch-Maschine-Interaktion“ beschrieben und bedienen sich vor allem der Steuerung mittels Gesten und Mimik, welche mithilfe von Kamerasystemen umgesetzt werden (Preim & Dachzelt, 2015; Schlick et al., 2018). Gorecky et al. (2014) erwähnen, dass eine intuitive und natürliche Interaktion im cyberphysischen Raum ähnlich wie mit realen Objekten stattfinden soll. Augmented Reality eignet sich demnach besonders für die Anwendung von natürlicher Interaktion, weil hier mit dem Einsatz des Körpers die reale Welt mit der digitalen Welt verbunden wird.

Darüber hinaus haben sich mit dem Aufkommen der schnellen Datenanalyse und der Verwendung von Hochleistungskameras die Möglichkeiten der visuellen Gestenerkennung stark verbessert (Gorecky et al., 2014). Garcia et al. (2019) erwähnen ebenso, dass die interaktive Komponente von cyberphysischen Systemen das Maximum erreicht, wenn sie mit einer „Natürlichen Mensch-Maschine-Schnittstelle (Natural Human-Machine Interface – NHMI)“ gestaltet sind.

Der Begriff „Geste“ beschreibt die Bewegung oder Pose des Körpers bzw. von Körperteilen und dient dabei der Übermittlung signifikanter Informationen z. B. an den Computer (DIN EN ISO 9241-960:2018; Preim & Dachzelt, 2015). Vor allem stellt der Einsatz von Gesten, ausgeführt mit den Händen, eine besondere Bedeutung für die Anwendung in der Mensch-Maschine-Interaktion dar (Preim & Dachzelt, 2015).

Dabei besteht die Herausforderung, die nutzeradaptive Interaktionsmöglichkeit für den Kontext des großen Arbeitsraumes in der Großgeräte- und Baustellenmontage zu gewährleisten und damit das Projektionssystem mit natürlichen Bewegungen im ganzen Arbeitsraum zu steuern.

Aufgrund der oben genannten Konzepte besteht die Notwendigkeit, das Spatial Augmented Reality System mit einer nutzeradaptiven Interaktionsmöglichkeit auszustatten, welche die Merkmale und Bewegungen (Posen, Gesten etc.) des Menschen erkennt und eine natürliche Interaktion mit dem Assistenzsystem ermöglicht.

Zusammengefasst liegen die Herausforderungen der Großgeräte- und Baustellenmontage einerseits in der Informationsbereitstellung und Interaktion über einen großen Arbeitsbereich, andererseits darin, dass aktuelle industrielle Spatial Augmented Reality Assistenzsysteme mit keiner dynamischen Komponente ausgestattet sind und demnach nur unzureichend in dieser Montageform eingesetzt werden können. Eine weitere Herausforderung ist die Gestaltung einer nutzeradaptiven Interaktion mit natürlichen Gesten für die intuitive Anwendung im industriellen Umfeld und direkt am Ort des Geschehens im Arbeitsbereich. Die erwähnten Herausforderungen und potenziellen Lösungsansätze stellen die Relevanz des Themas dar und sind somit die Grundlage der vorliegenden Forschungsarbeit.

1.3 Forschungsbedarf und Zielsetzung

Aus dem Kontext der industriellen Großgeräte- und Baustellenmontage und den Herausforderungen der Informationsbereitstellung sowie der Relevanz der Themen Spatial Augmented Reality, dynamische In-situ-Projektion und nutzeradaptive Interaktion ergeben sich die folgenden Forschungsfelder der Arbeit (Abbildung 6).

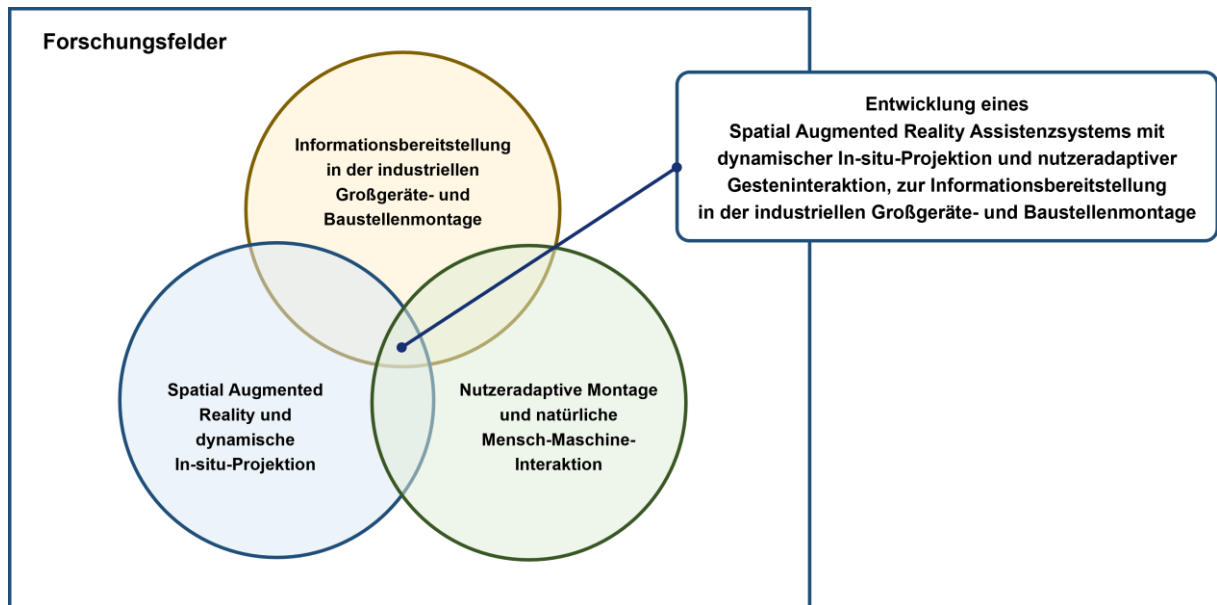


Abbildung 6 | Forschungsfelder der vorliegenden Arbeit (Eigendarstellung)

Der Anwendungsbereich sowie der Kontext der Arbeit stellen das Forschungsfeld „**Informationsbereitstellung industrielle Großgeräte- und Baustellenmontage**“ dar, für dessen Anforderungen ein Assistenzsystem entwickelt bzw. adaptiert wird. Das zweite Forschungsfeld ist „**Spatial Augmented Reality und dynamische In-situ-Projektion**“. Hier wird speziell der Einsatz von Spatial Augmented Reality als Technologie fokussiert, der ein dynamisches Projektionssystem aus der Veranstaltungsbranche implementiert und auf Gestaltungsaspekte der dynamischen In-situ-Projektion und Informationsführung durch den Montageprozess im großen Arbeitsraum eingeht. Der dritte Forschungsbereich stellt Konzepte der **nutzeradaptiven Montage** dar und betrachtet vor allem die Gestaltung der **natürlichen Mensch-Maschine-Interaktion** mit Posen und Gesten zur Steuerung des Informationssystems.

Resultierend aus der Ermittlung der Herausforderungen und Relevanz des Themas sowie aus der Eingrenzung der Forschungsfelder besteht Forschungsbedarf hinsichtlich der Entwicklung eines Spatial Augmented Reality Assistenzsystems, welches Arbeitsinformationen über einen großen Arbeitsraum dynamisch in situ projiziert, den Menschen gezielt durch den Montageprozess führt und verbunden mit einer nutzeradaptiven Interaktion einen erheblichen Mehrwert für die Arbeitstätigkeiten in der Großgeräte- und Baustellenmontage darstellt. Abgeleitet aus dem Forschungsbedarf ergeben sich folgende Forschungsfragen:

Zentrale Forschungsfrage:

Welchen Mehrwert bietet ein Spatial Augmented Reality Assistenzsystem mit dynamischer In-situ-Projektion und nutzeradaptiver Gesteninteraktion zur Verbesserung der Informationsbereitstellung in der industriellen Großgeräte- und Baustellenmontage?

Unterfragen:

- 1) Wie kann das Spatial Augmented Reality Assistenzsystem für die Anforderungen der industriellen Großgeräte- und Baustellenmontage **gestaltet** und **demonstriert** werden?
- 2) Wie kann eine **dynamische In-situ-Projektion und Informationsführung** im Montageprozess umgesetzt werden?
- 3) Wie kann eine **nutzeradaptive Mensch-Maschine-Interaktion mit natürlichen Gesten** gestaltet werden?
- 4) Welchen **Mehrwert** bietet das Assistenzsystem für **Mitarbeiter*innen** (Usability, kognitive Arbeitsbelastung) und **Unternehmen** (Prozesszeit) im Vergleich zu bestehenden Informationssystemen in der Großgeräte- und Baustellenmontage?

Die Zielsetzung der Dissertation ist, durch Gestaltung eines Spatial Augmented Reality Assistenzsystems mit dynamischer In-situ-Projektion und nutzeradaptiver Gesteninteraktion einen konkreten wissenschaftlichen und praxisrelevanten Beitrag zur Verbesserung der Informationsbereitstellung und Interaktionsgestaltung im Kontext der industriellen Großgeräte- und Baustellenmontage zu schaffen. Davon abgeleitet ist das konkrete Ziel, ein kognitives Assistenzsystem zu designen, zu demonstrieren und mit anerkannten Methoden zu evaluieren, um anschließend den Mehrwert für Mitarbeiter*innen und Unternehmen zu ermitteln.

Die Fokussierung sowie die Nicht-Ziele der Arbeit werden, ausgehend vom definierten Forschungsbereich und den Forschungsfragen, in Tabelle 1 dargestellt.

Tabelle 1 | Fokus und Zielsetzung der Forschungsarbeit

Forschungsbereich	Fokus
Informationsbereitstellung in der industriellen Großgeräte- und Baustellenmontage	- Großgerätemontage organisiert als Baustellenmontage in der Industrie - Montagestationen mit hohem manuellem Arbeitsanteil und geringem Automatisierungsgrad - Verbesserung der Ergonomie und Prozesseffizienz - Dynamische In-situ-Informationsbereitstellung über einen großen Arbeitsbereich
	Kein Fokus/Nicht-Ziele
	- Montageformen mit hohem automatisierten Arbeitsanteil - Fließlinien und Kleingerätemontage - Maintenance-Prozesse - Großgeräte- und Baustellenmontage außerhalb des industriellen Umfelds
Spatial Augmented Reality und dynamische In-situ-Projektion	Fokus
	- Einsatz von dynamischen projektorbasierten Informationssystemen - Einsatz von echtzeitfähigen Kamerasystemen und Bilderkennungsalgorithmen - In-situ-Projektion - Digitale Vernetzung von Informations- und Interaktionssystemen
	Kein Fokus/Nicht-Ziele
	- Nicht digital unterstützte Informations- und Interaktionssysteme - Informationsbereitstellung mit PC-Terminals, HMD, Wearables, Tablets

Nutzeradaptive Montage und natürliche Mensch-Maschine- Interaktion	Fokus
	▪ Gestaltung eines Montagesystems mit Fokus auf die Nutzer (= nutzeradaptiv) ▪ Nutzeradaptive Gestaltung des Spatial Augmented Reality Assistenzsystems ▪ Fokus auf natürliche Gesten und Interaktion mit Kamerasystemen ▪ Echtzeitfähige Interaktionsmöglichkeit mit visueller Bilderkennung ▪ Fokus auf praxisrelevante Systeme für die Montage ▪ Erkennung von Menschen und dessen natürliche Bewegungen, Gesten und Posen
	Kein Fokus/Nicht-Ziele
	▪ Akustische oder touchscreenbasierte Interaktion ▪ Einsatz von Wearables, Smartphones oder PC-Terminals zur Interaktion ▪ Gestaltung bezüglich des Arbeitsumfelds (z. B. Einstellung nach Tagesverlauf)

1.4 Forschungsmethodik – Design Science Research

Als Forschungsmethodologie wird der Ansatz „Design Science“ verwendet, welcher in der angewandten Forschung wie in den Ingenieurwissenschaften und der Wirtschaftsinformatik breite Anwendung findet und darauf ausgerichtet ist, Fragestellungen aus der wirtschaftlichen Praxis mit nützlichen technologischen Entwicklungen zu lösen (Alan R. Hevner et al., 2004; Alan R. Hevner, 2007; March & Smith, 1995). Im Speziellen wird auf die **Design-Science-Research-Methode** eingegangen, welche auf den Arbeiten von Alan R. Hevner et al. (2004) und Alan R. Hevner (2007) aufbaut und eine iterative Vorgehensweise mit mehreren Designzyklen ermöglicht.

In den Ingenieurwissenschaften und der Wirtschaftsinformatik werden im Gegensatz zu den Sozial- und Wirtschaftswissenschaften die Forschungs- und Untersuchungsobjekte (Artefakte) erst während des Forschungsprozesses gestaltet. Diese gestalteten Objekte sowie demonstrierten Ergebnisse werden anschließend kritisch bewertet und evaluiert sowie in einem iterativen Prozess verbessert. Speziell in der anwendungsorientierten Forschung ist es entscheidend, dass ein positiver Nutzen und Mehrwert außerhalb des wissenschaftlichen Kontexts, z. B. für Mitarbeiter*innen oder Unternehmen aus der Industrie, entsteht (van Aken, 2004, 19f). Es werden demnach Lösungen für Anwendungsprobleme entworfen und pragmatische Zielformulierungen entwickelt. March und Smith (1995) haben ein Vorgehensmodell entwickelt, welches auf dem gestaltungszentrierten Design-Science-Ansatz basiert und für anwendungsorientierte Forschung geeignet ist. Dabei wird als bedeutsames Ziel der Design-Science-Research-Methode die Nützlichkeit der Forschungsergebnisse betont (Alan R. Hevner et al., 2004; A. R. Hevner & March, 2003; March & Smith, 1995; van Aken, 2004). Design besteht sowohl aus dem Design-Prozess (Abfolge von Aktivitäten zur Erstellung eines Artefaktes) als auch aus den gestalteten Artefakten selbst (Walls et al., 1992). Die entstandenen Artefakte unterziehen sich einer Evaluierung, um schließlich wiederum das Problem besser zu verstehen und die Qualität des Design-Prozesses und des Artefakts zu erhöhen. Diese zweiseitige Betrachtungsweise unterstützt das Problem- und Lösungsverständnis, weil ständig ein Perspektivenwechsel durchgeführt werden muss (Alan R. Hevner et al., 2004). Die Design Science Methode, nach den Grundzügen von March und Smith (1995), besteht aus den Kernaktivitäten „Design“ und „Evaluieren“ sowie den aus dem naturwissenschaftlichen Kontext stammenden Aktivitäten „theoretisieren“ und

„begründen“. Unter „Design“ wird die Gestaltung und Entwicklung eines Artefaktes verstanden, welches nach einem definierten Ziel ausgerichtet sein soll. Die Ergebnisse der Phase „Design“ können Konstrukte, Modelle, Methoden und Implementierungen sein, welche anschließend in der Phase „Evaluierung“ auf den Mehrwert, Nutzen sowie auf die Zielerreichung untersucht werden (March & Smith, 1995). Auf diesen Grundzügen stellen Alan R. Hevner et al. (2004) ein konzeptionelles Framework vor, welches die zwei Forschungsparadigmen der Informationssystemforschung „Behavioral Science“ (deutsch „Verhaltenswissenschaften“) und „Design Science“ (deutsch „Gestaltende Wissenschaften“) miteinander kombiniert. „Behavioral Science“ kennzeichnet die Forschung, welche organisatorische Abläufe und das resultierende menschliche Verhalten erklären und vorhersagen soll. Design Science versucht, menschliche und organisatorische Möglichkeiten („business needs“) durch Schaffung innovativer Artefakte zu verbessern und einen Nutzen daraus zu schaffen. Das Framework baut auf sieben Guidelines, welche folgend dargestellt werden (Alan R. Hevner et al., 2004):

1. **Design als ein Artefakt:** Design Science fordert ein gestaltetes Artefakt in Form eines Konstrukts, Modells, Methode oder Implementierung.
2. **Problemrelevanz:** Praxisrelevante Problemstellungen und „business needs“ gehen den technologiegestützten Lösungen voraus.
3. **Design-Evaluation:** Geeignete Evaluierungsmethoden dienen zur Bewertung der Nützlichkeit, Effizienz und Qualität des Artefakts.
4. **Beiträge zur Forschung:** Erfordernis von klaren Beiträgen zur Forschung mit der Umsetzung der Design-Artefakte und Design-Methoden.
5. **Forschungsstärke:** Design Science fordert die Verwendung von formalen Methoden hinsichtlich der Gestaltung und der Evaluation des Artefakts.
6. **Design als Suchprozess:** Unter Einhaltung von gegebenen Rahmenbedingungen werden im Problemumfeld zielführende Artefakte gesucht und gestaltet.
7. **Forschungskommunikation:** Die wirkungsvolle Kommunikation der Forschungsergebnisse muss gegenüber managementorientierten und technisch orientierten Interessengruppen stets umgesetzt werden.

Alan R. Hevner (2007) beschreibt und erweitert Design Science Research als Verkörperung dreier eng miteinander verbundenen Zyklen von Aktivitäten (Abbildung 7).

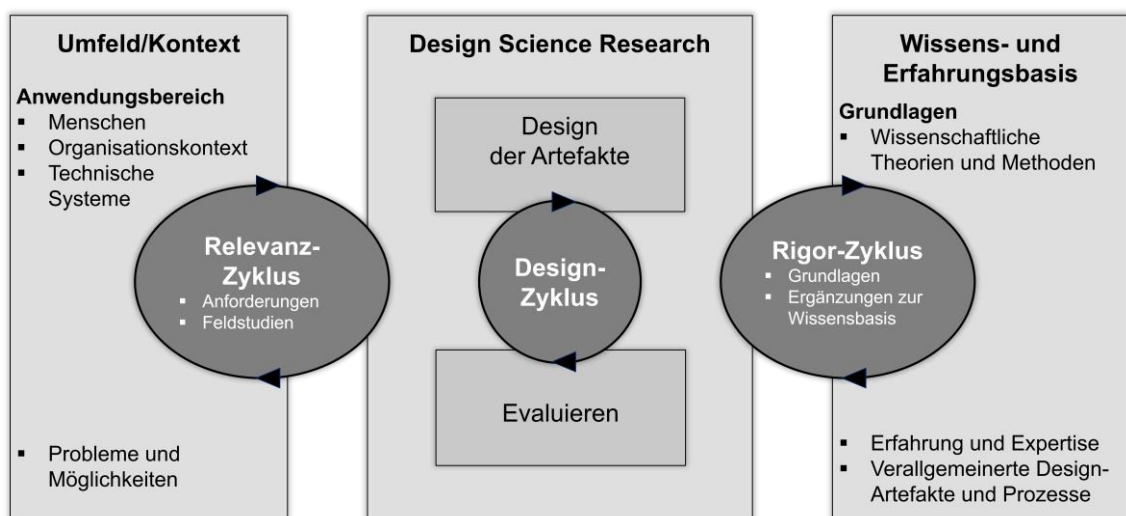


Abbildung 7 | Zyklen der Design-Science-Research-Methode, in Anlehnung an Alan R. Hevner (2007)

Die Zyklen bestehen aus dem **Relevanz-Zyklus**, welcher Anforderungen aus dem anwendungsorientierten kontextuellen Umfeld in die Forschung einbringt, dem **Rigor-Zyklus**, welcher das Wissen von Theorie und bestehenden Methoden herleitet, mit der Erfahrung und dem Fachwissen zusammenführt sowie neues Wissen aus der laufenden Forschung hinzufügt und schließlich aus dem zentralen **Design-Zyklus**, welcher die Entwicklung und das Design sowie die Evaluierung der Artefakte umsetzt.

Die Zyklen folgen einem ständig wiederholenden Prozess und verknüpfen Theorie und Praxis, woraus sich schließlich die Artefakte ableiten, welche wiederum stetig durch neue Erkenntnisse verbessert werden. Die verbesserten Artefakte liefern dann rückführend wieder einen neuen Beitrag zu den theoretischen und praktischen Erkenntnissen, was den kompletten Design-Prozess-Zyklus wieder schließt (Alan R. Hevner, 2007).

Folgend werden die Design-Research-Methode nach Alan R. Hevner (2007) sowie die Guidelines von Alan R. Hevner et al. (2004) auf das Forschungsvorhaben angewandt (Tabelle 2).

Tabelle 2 | Umsetzung der Guidelines der Design-Research-Methode, in Anlehnung an Alan R. Hevner et al. (2004)

Guidelines	Umsetzung der Guidelines in der Dissertation
▪ Design als ein Artefakt: Design Science fordert ein gestaltetes Artefakt in Form eines Konstrukts, Modells, einer Methode oder Implementierung.	Das Spatial Augmented Reality Assistenzsystem mit dynamischer In-situ-Projektion und nutzeradaptiver Interaktion mit natürlichen Gesten wird in Form von Konzepten, Grafiken und Renderings schematisch dargestellt und anschließend als Konstrukt in Form eines Demonstrators aufgebaut.
▪ Problemrelevanz: Praxisrelevante Problemstellungen und „business needs“ gehen den technologiegestützten Lösungen voraus.	Das praxisrelevante Problem sowie die Herausforderungen werden anhand der Analyse des Kontexts und der Systematischen Literaturrecherche (SLR) nach Ramdhani et al. (2014) ermittelt und der Forschungsbedarf abgeleitet. Dabei stellen der mangelnde Einsatz von Assistenzsystemen zur dynamischen In-situ-Informationsbereitstellung sowie die fehlende nutzeradaptive Interaktion über große Arbeitsräume in der Großgeräte- und Baustellenmontage die Hauptmerkmale dar.
▪ Design-Evaluation: Geeignete Evaluierungsmethoden dienen zur Bewertung der Nützlichkeit, Effizienz und Qualität des Artefakts.	Folgende Evaluierungsmethoden werden verwendet: ▪ Nutzerakzeptanz: „Technology Acceptance Model“ (TAM) nach Davis et al. (1989) ▪ Usability: „System Usability Scale“ (SUS) nach Brooke (1996) ▪ Prozessoptimierung: „Task Completion Time“ (TCT) nach dem „General Assembly Task Model“ (GATM) nach Funk, Kosch et al. (2015) ▪ Kognitive Arbeitsbelastung: NASA-RTLX-Methode nach Hart und Staveland (1988) und Hart (2006)
▪ Beiträge zur Forschung: Erfordernis von klaren Beiträgen zur Forschung mit der Umsetzung der Design-Artefakte und Design-Methoden	Die Erkenntnisse aus dem Design-Research-Prozess, der Demonstration sowie der Evaluierung der Artefakte, dienen als Grundlage für die Forschung und der Entwicklung des Spatial Augmented Reality Assistenzsystems in der industriellen Großgeräte- und Baustellenmontage.
▪ Forschungsstärke: Design Science fordert die Verwendung von formalen Methoden hinsichtlich der Gestaltung und der Evaluation des Artefakts.	Der Forderung wird in Form der Durchführung einer Systematischen Literaturrecherche, der Umsetzung und Demonstration von prototypischen Artefakten im industrienahen Forschungsumfeld sowie durch die Verwendung von anerkannten Evaluierungsmethoden entsprochen.

▪ Design als Suchprozess: Unter Einhaltung von gegebenen Rahmenbedingungen werden im Problemumfeld zielführende Artefakte gesucht und gestaltet.	Durch Nutzung der Rahmenbedingungen der TU Wien-Pilotfabrik und des Netzwerks zu Industriepartnern, wurde das Problemumfeld analysiert und nach zielführenden Artefakten und dem potenziellen Einsatz von kognitiven Assistenzsystemen im Bereich der industriellen Großgeräte- und Baustellenmontage gesucht.
▪ Forschungskommunikation: Die wirkungsvolle Kommunikation der Forschungsergebnisse muss gegenüber managementorientierten und technisch orientierten Interessengruppen stets umgesetzt werden.	Die Forschungskommunikation wird über Publikationen von Papers (vgl. Mayrhofer, Rupprecht & Schlund, 2019; Rupprecht et al., 2020; Rupprecht et al., 2021; Rupprecht & Schlund, 2021; Schlund et al., 2018) bei fach einschlägigen Konferenzen und Zeitschriften wie z. B. der CIRP Conference on Manufacturing Systems (CIRP CMS), International Conference on Production Research Manufacturing Innovation: Cyber Physical Manufacturing (ICPR) und der „Zeitschrift für Arbeitswissenschaft (ZfA)“ umgesetzt und durch Demonstrationen in der TU Wien-Pilotfabrik der praxisorientierten Community vermittelt.

Die Design-Research-Methode wurde auf die Rahmenbedingungen und Anforderungen des Forschungskontexts angepasst und wird folgend dargestellt (Abbildung 8).

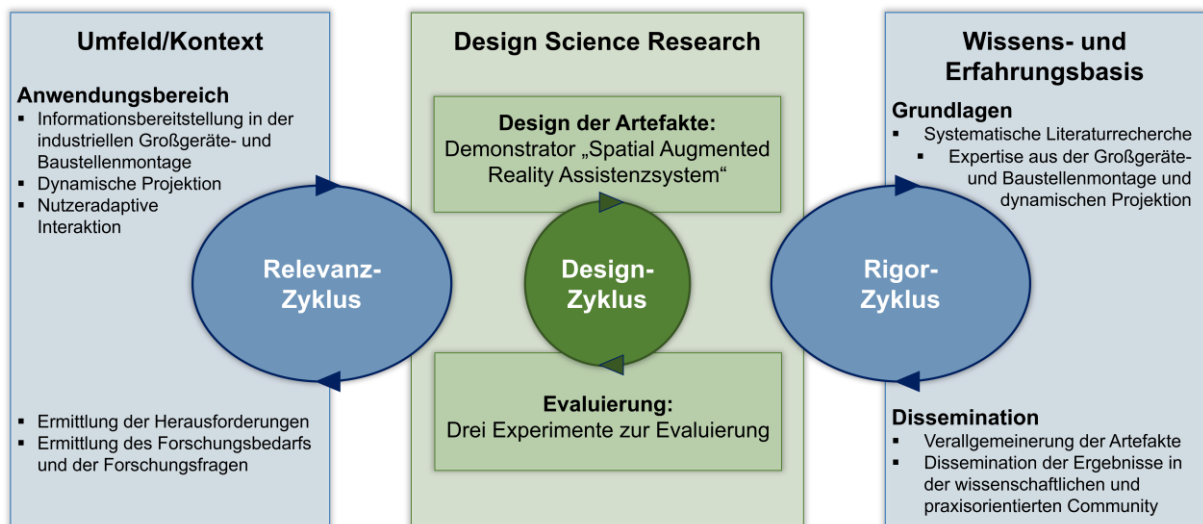


Abbildung 8 | Design-Science-Research-Methode der Forschungsarbeit, in Anlehnung an Alan R. Hevner (2007)

Der Bereich „Umfeld und Kontext“ umfasst die definierten Forschungsfelder, deren praxisrelevanten Herausforderungen und den Forschungsbedarf. Der Kontext fließt über den iterativen Relevanz-Zyklus in die Gestaltung der Artefakte mit ein. Der Bereich „Wissens- und Erfahrungsbasis“ unterstreicht die wissenschaftliche Relevanz des Forschungsvorhabens und bezieht die Expertise im Bereich der Großgeräte- und Baustellenmontage sowie der dynamischen Projektion mit ein. Mit der Systematischen Literaturrecherche werden Grundlagen und aktuelle Forschungen ermittelt, welche durch den Rigor-Zyklus in die Gestaltung des Assistenzsystems integriert werden. Design Science Research ist der zentrale Bereich der Methode, welcher die Gestaltung und Evaluierung des Designartefakts beinhaltet. Dabei wird in einem iterativen Design-Zyklus das Spatial Augmented Reality Assistenzsystem gestaltet und mit geeigneten Evaluierungsmethoden bewertet. Die Erkenntnisse werden verallgemeinert und der Forschungscommunity über die Demonstration des Systems in der Pilotfabrik sowie durch Veröffentlichung von wissenschaftlichen Artikeln zur Verfügung gestellt.

1.5 Aufbau und Struktur der Arbeit

Der Aufbau und die Struktur der Arbeit wird in Abbildung 9 dargestellt und gibt einen Überblick über die Inhalte der einzelnen Kapitel.

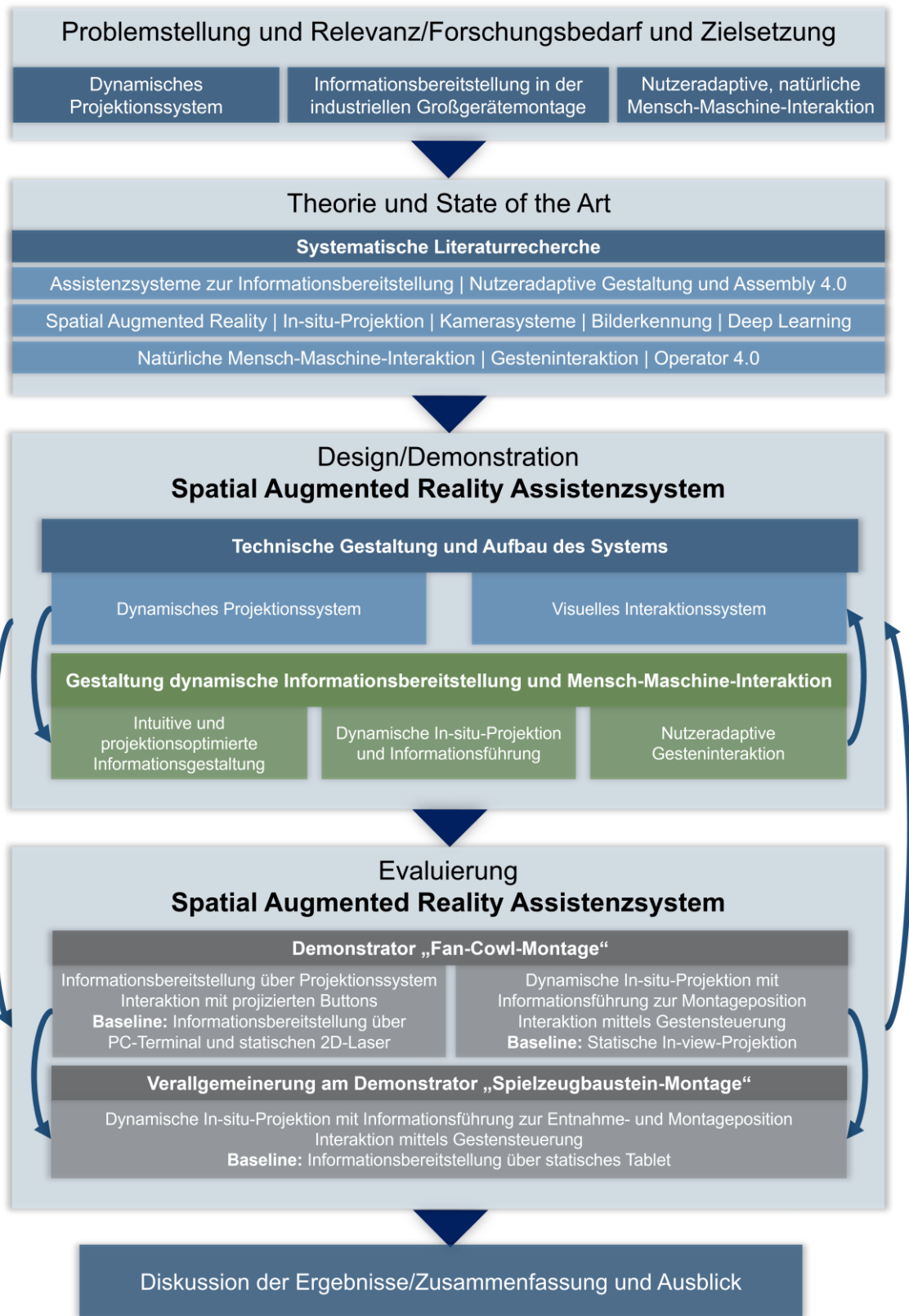


Abbildung 9 | Aufbau und Struktur der Forschungsarbeit (Eigendarstellung)

Die Forschungsarbeit beginnt in Kapitel 1 mit der Darstellung der Problemstellung und Relevanz des Themas. Des Weiteren werden die Herausforderungen der Themenbereiche und des Kontexts ermittelt und die Forschungsfelder sowie der Forschungsbedarfs abgeleitet. Anschließend wird die Zielsetzung inklusive der Forschungsfragen dargestellt und die Definition des Forschungsdesigns sowie der Aufbau der Arbeit präsentiert. In Kapitel 2 werden die Vorgehensweise, die Ergebnisse der Systematischen Literaturrecherche sowie relevante Grundlagen zum Stand der Wissenschaft dargestellt. In Kapitel 3 und 4 werden die Gestaltung und Demonstration des Spatial Augmented Reality Assistenzsystems dargestellt und dessen technischer Aufbau sowie die Umsetzung der dynamischen In-situ-Informationsbereitstellung und nutzeradaptiven Interaktion präsentiert. In Kapitel 5 werden die Vorgehensweise sowie die Ergebnisse der Evaluierung mittels drei Experimenten dargestellt. In Kapitel 6 werden die Ergebnisse diskutiert und die Erkenntnisse aus dem Design der Artefakte sowie der Evaluierung abgeleitet und die Limitationen des Systems dargestellt. In Kapitel 7 wird die Dissemination der Forschungsergebnisse dargestellt und die Forschungsarbeit abschließend zusammengefasst.

2. Grundlagen und Stand der Wissenschaft

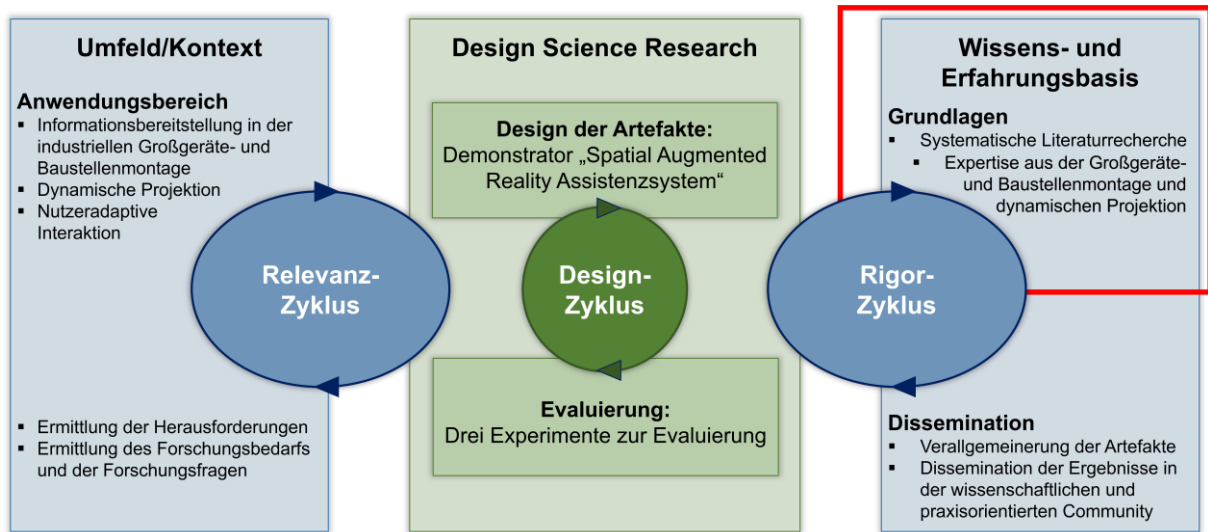


Abbildung 10 | Wissens- und Erfahrungsbasis – Grundlagen, in Anlehnung an Alan R. Hevner et al. (2004)

Das Element „Wissens- und Erfahrungsbasis“ (vgl. Abbildung 10) der Design-Science-Research-Methode nach Alan R. Hevner et al. (2004) bringt über den Rigor-Zyklus relevante Grundlagen sowie den Stand der Wissenschaft in die Forschungsarbeit ein.

In den folgenden Kapiteln werden die Vorgehensweise und Ergebnisse der Systematischen Literaturrecherche präsentiert, wissenschaftliche und praxisrelevante State-of-the-Art-Konzepte, Methoden und Technologien dargestellt und die wissenschaftliche Relevanz des Forschungsvorhabens ermittelt.

2.1 Systematische Literaturrecherche

Um einen konkreten und objektiven Überblick über den Stand der relevanten Forschungsarbeiten in einem definierten wissenschaftlichen Umfeld zu bekommen, bedarf es einer systematischen und strukturierten Vorgehensweise der Suche, Selektion und Evaluierung von wissenschaftlichen Publikationen (Okoli & Schabram, 2010), die in Form einer **Systematischen Literaturrecherche** (Ramdhani et al., 2014) in den wissenschaftlichen Datenbanken Scopus, Science Direct sowie Springer Link durchgeführt wird. Die strukturierte Vorgehensweise, schrittweise Analyse und Reduktion der Suchergebnisse führen zum Erhalt der relevanten wissenschaftlichen Literatur. Ramdhani et al. (2014) definieren dazu die Vorgehensweise anhand von fünf Schritten, die in dieser Arbeit Verwendung gefunden haben (Abbildung 11).

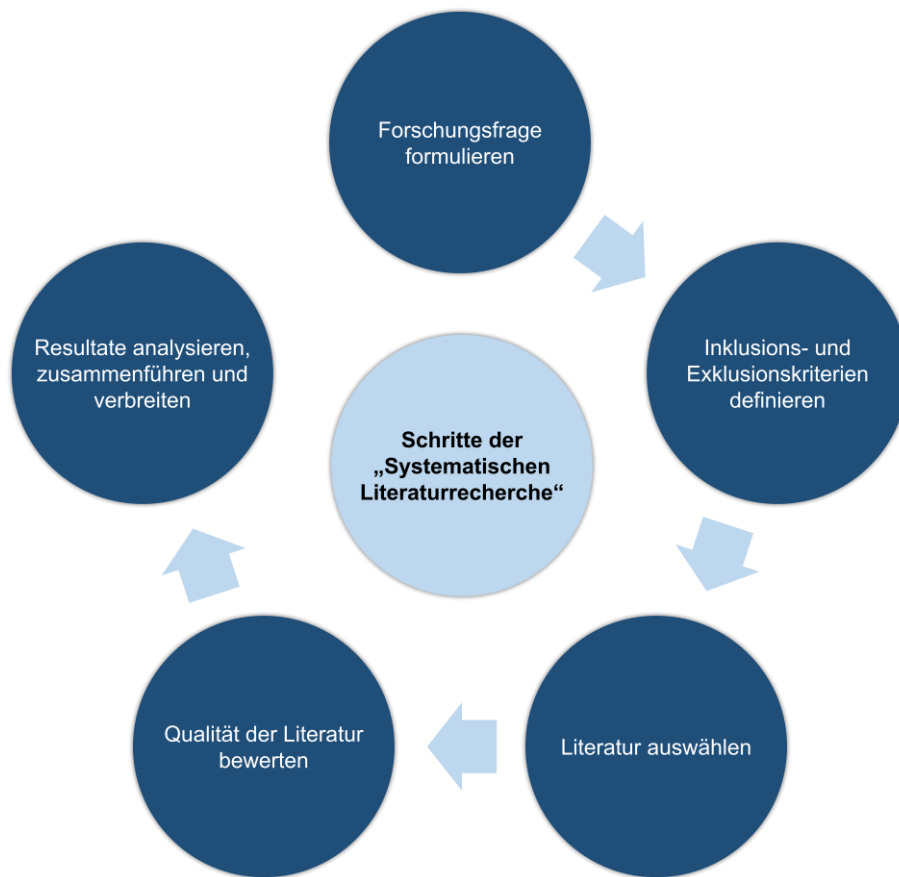


Abbildung 11 | Kriterien der Systematischen Literaturrecherche, in Anlehnung an Ramdhani et al. (2014)

Ausgehend von den definierten Forschungsfragen (vgl. Kapitel 1.3) wurden für die spezifischen Suchanfragen die Suchtools der jeweiligen Literaturdatenbanken verwendet. Für die Darstellung und Filterung der Ergebnisse wurde MS Excel und für die Integration zur Verwendung in der Dissertation das Literaturverwaltungstool CITAVI verwendet. Die detaillierte Vorgehensweise wurde nach Definition der Schlüsselwörter in drei Schritten durchgeführt (Abbildung 12).

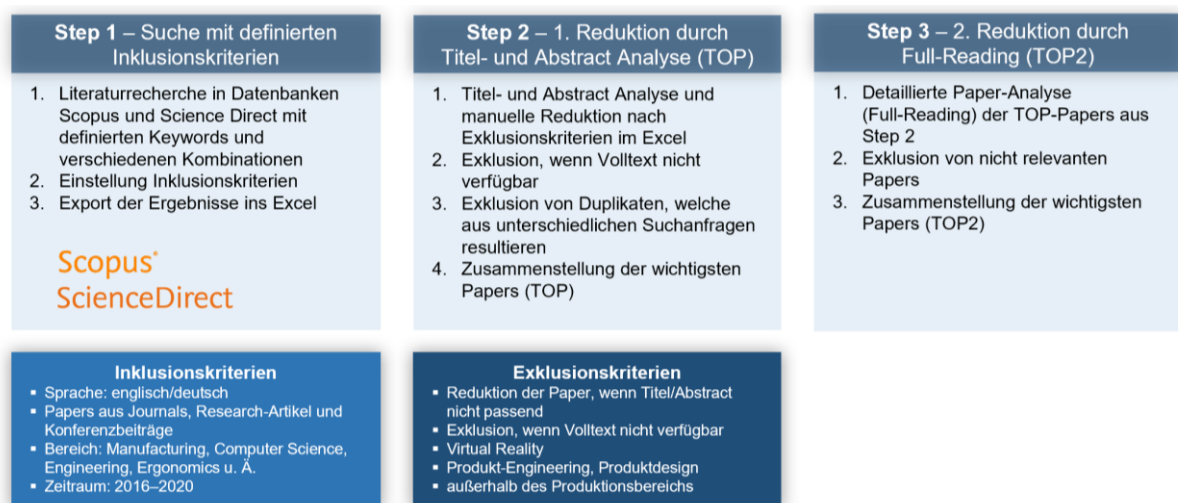


Abbildung 12 | Detaillierte Vorgehensweise mit Inklusions- und Exklusionskriterien (Eigendarstellung)

Um eine strukturierte Suche durchzuführen, wurden im Bereich Produktion und Montage konkrete Schlüsselwörter der menschenzentrierten Arbeitsplatzgestaltung, der nutzeradaptiven Einstellung von Arbeitsplätzen sowie im Bereich Projektion definiert und mit logischen Konnektoren (AND/ODER) zu konkreten Suchbegriffen kombiniert. Die Suchbegriffe wurden dabei in englischer Sprache definiert.

Als Eingrenzung der Suche (Inklusionskriterien) werden nur deutsch- und englischsprachige Papers akzeptiert, welche zwischen 2016 und 2021 veröffentlicht und als Journals, Research-Artikel sowie Konferenzbeiträge in den Forschungsfeldern Engineering und Manufacturing, Computer Science oder Ergonomics deklariert wurden. Die Ergebnisse (Titel, Datenbankname, Suchanfragennummer) aus den verschiedenen Suchanfragen wurden für die weitere Analyse (Step 2) in MS Excel exportiert, geordnet und gefiltert. Mit Step 2 wurden eine Titel- und Abstract Analyse nach definierten Exklusionskriterien durchgeführt und nicht relevante Papers entfernt. Anschließend wurden Duplikate, welche aus den verschiedenen Suchanfragen resultierten, eliminiert und die Volltexte der relevanten Papers (TOP) downgeloadet. In Step 3 wurde zur weiteren Reduktion der Volltext näher betrachtet und bei Relevanz mit dem Tag „TOP2“ versehen. Aus den Literaturverzeichnissen identifizierte weitere relevante Literatur wurde ebenso ins MS Excel unter der Kategorie „WEITERE“ integriert und mit dem Tag „TOP2“ versehen. Nach Analyse der relevanten Papers kristallisierten sich die Begriffe „Operator 4.0“, „Assembly 4.0“, „in-situ projection/instructions“ sowie „projection-based AR system“ heraus, nach welchen weitere Suchanfragen durchgeführt wurden. Die Kombination der Schlüsselwörter und die durchgeführten Suchanfragen werden in Tabelle 3 präsentiert.

Tabelle 3 | Kombination der Schlüsselwörter zu konkreten Suchbegriffen

Suchanfragen	Kombination der Schlüsselwörter
F1	(production OR manufacturing OR assembly) AND (customizable OR individualizable OR personalizable) AND human centered workplace design
F2.1	(production OR manufacturing OR assembly) AND context adaptive configurable workplace design
F2.2	(production OR manufacturing OR assembly) AND user adaptive configurable workplace design
F3	(production OR manufacturing OR assembly) AND projection AND workplace design
F4	(production OR manufacturing OR assembly) AND Spatial Augmented Reality AND workplace design
Weitere Suchanfragen	(production OR manufacturing OR assembly) AND "in-situ instructions" (production OR manufacturing OR assembly) AND "in-situ projection" (production OR manufacturing OR assembly) AND "projection-based AR system" Operator 4.0 Assembly 4.0

Als abschließender Punkt wurden die ermittelten Papers aus der SLR sowie den Literaturverzeichnissen in MS Excel nach ihrer Qualität bewertet und in die Kategorien A, B und C (A = relevanteste) eingeteilt sowie den Tags „Operator 4.0“, „Assembly 4.0“ und „Spatial Augmented Reality“ für die schnelle Auffindung zugewiesen.

Durch die Durchführung der SLR wurde ein guter Überblick über den Stand der Wissenschaft geschaffen. Die Anzahl der gefundenen Papers aus der Systematischen Literaturrecherche sowie die Anzahl der als „relevant“ kategorisierten Artikel wird je nach Suchanfrage in Tabelle 4 dargestellt.

Tabelle 4 | Ergebnisse der Systematischen Literaturrecherche

Suchanfragen-nummer	Schlüsselwörter production OR manufacturing OR assembly) AND ...	Gefundene Paper	Relevante Paper (TOP2)
F1	(customizable OR individualizable OR personalizable) AND human centered workplace design	60	4
F2.1, F2.2	context/user adaptive configurable workplace design	136	13
F3	projection AND workplace design	82	3
F4	Spatial Augmented Reality Assistenzsystem AND workplace design	84	9
Weitere Suchanfragen	Recherche nach „in-situ instructions“, „in-situ projection“, „projection-based AR system“, „Operator 4.0“, „Assembly 4.0“ und relevante Papers aus den Literaturverzeichnissen	65	41
Gesamt		427	70

Insgesamt wurden über 427 Paper aus den verschiedenen Kombinationen der Schlüsselwörter in den Datenbanken sowie aus den jeweiligen Literaturverzeichnissen recherchiert und analysiert. Nach der systematischen Reduktion konnten insgesamt 70 Paper als relevant eingestuft werden. Tabelle 5 gibt einen Überblick über die Zuordnung der recherchierten Papers zu den Begriffen „Spatial Augmented Reality“, „Operator 4.0“ und „Assembly 4.0“.

Tabelle 5 | State-of-the-Art-Analyse – Kategorisierung der Literatur

SotA-Analyse			Schlüsselbegriffe		
Nr.	Jahr	Autor	Spatial Augmented Reality	Operator 4.0	Assembly 4.0
1	2019	Bambušek et al.	X		
2	2018	Bertram et al.			X
3	2019	Blattgerste et al.	X		
4	2017	Blattgerste et al.	X		
5	2018	Blattgerste et al.	X		
6	2017	Bortolini et al.			X
7	2019	Bortolini et al.			X
8	2017	Bosch et al.	X		
9	2017	Büttner et al.	X		
10	2015	Büttner et al.	X		
11	2017	Büttner et al.	X		
12	2016	Büttner et al.	X		
13	2017	Cohen et al.			X
14	2019	Cohen et al.			X
15	2019	Cohen et al.			X
16	2018	Cserteg et al.			X
17	2019	Dhiman et al.	X		
18	2019	Eckert et al.			X
19	2020	Egger und Masood	X		
20	2017	Elkomy et al.	X		
21	2019	Faccio et al.			X

22	2012	Falck et al.			X
23	2012	Fässberg et al.		X	
24	2020	Fletcher et al.			X
25	2016	Funk	X		
26	2016	Funk et al.	X		
27	2015	Funk et al.	X		
28	2016	Funk et al.	X		
29	2015	Funk et al.	X		
30	2017	Funk et al.	X		
31	2016	Funk et al.			X
32	2015	Funk et al.	X		
33	2016	Funk et al.	X		
34	2011	Gaham et al.		X	
35	2019	Garcia et al.		X	
36	2019	Guo et al.			X
37	2016	Hadorn et al.		X	
38	2008	Hakkarainen et al.			X
39	2018	Halme et al.	X		
40	2018	Jetter et al.			X
41	2020	Kaasinen et al.		X	
42	2019	Klumpp et al.		X	
43	2017	Kosch et al.	X		
44	2018	Kosch et al.	X		
45	2019	Lampen et al.	X		
46	2016	Lušic et al.		X	
47	2020	Manghisi et al.		X	
48	2019	Masood und Egger	X		
49	2018	Materna et al.	X		
50	2020	Mattsson et al.		X	
51	2019	Mayrhofer et al.		X	
52	2018	Mengoni et al.	X		
53	2014	Otto et al.	X		
54	2017	Peruzzini und Pellicciari			X
55	2020	Pieters et al.	X		
56	2019	Raisamo et al.		X	
57	2015	Rodriguez et al.	X		
58	2016	Romero et al.		X	
59	2020	Romero et al.		X	
60	2016	Romero et al.		X	
61	2016	Sand et al.	X		
62	2020	Segura et al.			
63	2019	Sochor et al.	X		
64	2018	Suh und Prophet	X		
65	2020	Taylor et al.		X	
66	2019	Temesvári et al.			X
67	2019	Torres et al.	X		
68	2019	Wang et al.	X		
69	2019	Wichmann et al.			X
70	2020	Zolotová et al.		X	

2.2 Nutzeradaptive Gestaltung und Assembly 4.0

Für die moderne Gestaltung der Großgeräte- und Baustellenmontage bedarf es an Konzepten zur Optimierung des Montageprozesses, zur Vernetzung der Montagestation sowie zur physischen und kognitiven Unterstützung des Menschen mit geeigneten Assistenzsystemen (vgl. Kapitel 1). Demnach ist der Einsatz von moderner Technologie (z. B. Augmented Reality, intelligente Sensorik, Kamerasysteme, smarte Informationssysteme etc.) zur adaptiven Einstellung und Anpassung des Arbeitssystems sowie zur Unterstützung der menschlichen Tätigkeiten mit Assistenzsystemen essenziell, die Auslegung einer effektiven sowie natürlichen Mensch-Maschine-Interaktion gewinnt dabei immer mehr an Bedeutung (Bortolini et al., 2017; Galaske & Anderl, 2016; Garcia et al., 2019; Gorecky et al., 2014; Reinhart, 2017). Demnach bedarf es an nutzeradaptiven Konzepten der Montagegestaltung sowie an effektiver Integration des Menschen in die Montagestation bis zu einer Verschmelzung in einem menschenzentrierten cyberphysischen Montagesystem (Romero et al., 2016).

2.2.1 Menschzentrierte Gestaltung und Operator 4.0

Der Mensch steht mit seinen anthropometrischen Merkmalen sowie individuellen Bedürfnissen im Mittelpunkt der Gestaltung der manuellen Montage. Demnach werden folgend geeignete Konzepte präsentiert, welche den Menschen in das cyberphysische System integrieren und bei der menschenzentrierten Gestaltung von Arbeits- und Assistenzsystemen unterstützen.

„Human Centered Design“ ist eine Herangehensweise, um menschliche Bedürfnisse, Fähigkeiten und Handlungsweisen in den Vordergrund zu stellen und die Gestaltungen und Entwicklungen diesen anzupassen (Norman, 2016). Menschzentrierte Gestaltung bedarf demnach einer ausreichenden Kenntnis, einerseits über den Menschen und dessen Psychologie und Verhaltensweisen und andererseits über die zu gestaltenden Artefakte (Technologien, Arbeitsplätze, Arbeitsprozesse etc.) (Alan R. Hevner et al., 2004; Norman, 2016). Die Norm DIN EN ISO 9241-210 beschreibt „**menschzentrierte Gestaltung**“, als „Herangehensweise bei der Gestaltung und Entwicklung von Systemen, die darauf abzielt, interaktive Systeme gebrauchstauglicher zu machen, indem sie sich auf die Verwendung des Systems konzentriert und Kenntnisse und Techniken aus den Bereichen der Arbeitswissenschaft/Ergonomie und der Gebrauchstauglichkeit anwendet“.

Bei der Gestaltung des Arbeitssystems sind folgende Bestandteile zu betrachten (DIN EN ISO 6385:2016-12):

- Gestaltung der Arbeitsorganisation mit Arbeitsaufgaben und Tätigkeiten
- Gestaltung der Arbeitsumgebung und Arbeitsraum
- Gestaltung der Arbeitsmittel und Schnittstellen
- Gestaltung des Arbeitsplatzes

Bei der Entwicklung des Spatial Augmented Reality Assistenzsystems wird im Speziellen auf die Gestaltung des Arbeitsplatzes und der Arbeitsmittel und Schnittstellen, bei definierten Arbeitsaufgaben im Arbeitsraum der Großgeräte- und Baustellenmontage, eingegangen.

Human-Cyber-Physical System und Operator 4.0

Bei cyberphysischen Produktionssystemen verschmelzen digitale und physische Komponenten und neue Technologien, Mensch-Maschine-Schnittstellen sowie User-Interfaces entstehen, was eine erhebliche Auswirkung auf die Arbeitsprozesse und Arbeitstätigkeiten bedeutet (Mattsson et al., 2020; Romero et al., 2016). Romero et al. (2016) stellen dar, dass Maschinentätigkeiten und menschliche Tätigkeiten nicht mehr getrennt gesehen werden, sondern in einer Art Symbiose verschmelzen sollen (Mensch-Automatisierungs-Symbiose). Dies wird unter dem Begriff „**Human-Cyber-Physical System (H-CPS)**“ zusammengefasst. Dabei wird beschrieben, dass die individuellen Fähigkeiten und Fertigkeiten von Menschen durch moderne Assistenzsysteme unterstützt werden sollen, was unter dem Begriff „**Operator 4.0**“ dargestellt wird. Das Konzept „Operator 4.0“ wurde von Romero et al. (2016) an die industriellen Revolutionen angelehnt und analog dazu die spezifischen Arbeitsweisen und Interaktionen mit den Maschinen definiert (Abbildung 13).

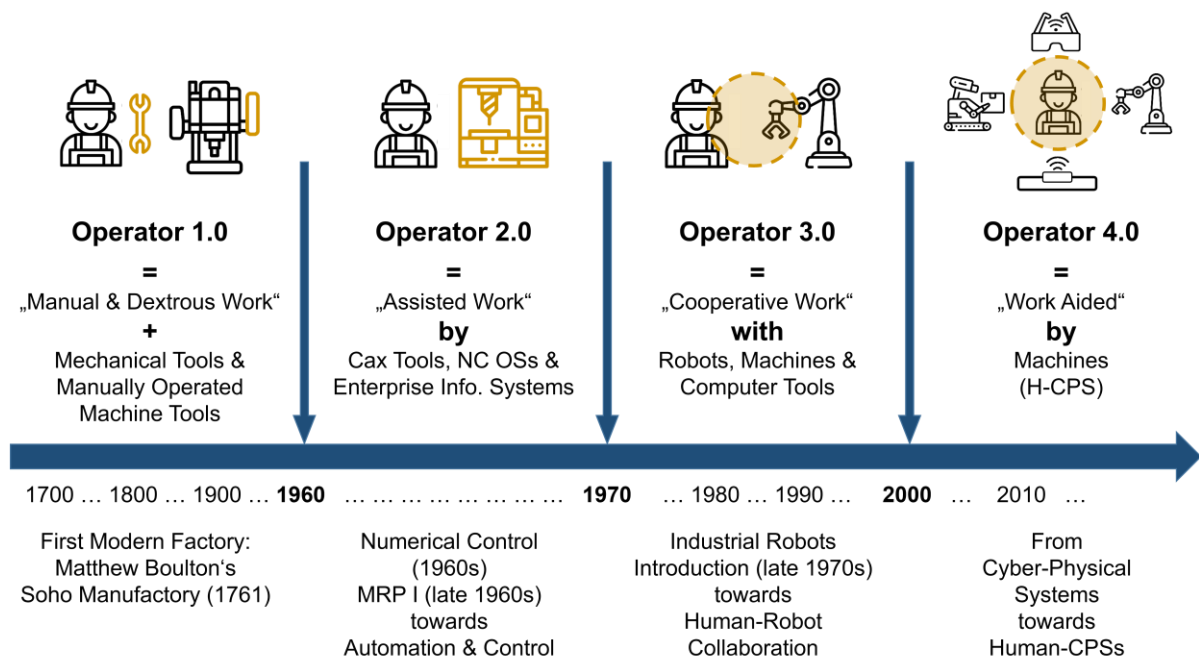


Abbildung 13 | Revolutionen von Operator 1.0 bis Operator 4.0, in Anlehnung an Romero et al. (2016)

Als Operator 1.0 ist jener definiert, der manuelle Arbeiten durchführt und mechanische Werkzeuge und Werkzeugmaschinen als Unterstützung verwendet. Der Operator 2.0 wird bei der Ausführung der Arbeitstätigkeiten von computergestützten Werkzeugen (z. B. CNC-Werkzeugmaschinen) sowie ersten Informationssystemen unterstützt. Der Operator 3.0 bedient den kompletten Produktionsprozess und agiert kooperativ und gemeinsam mit Robotern und Maschinen (Mensch-Maschine-Kooperation, Mensch-Roboter-Kooperation). Schließlich beschreiben Romero et al. (2016) den neuen Begriff des Operator 4.0, welche den „Operator der Zukunft“ repräsentiert. Als Operator 4.0 wird ein kluger, erfahrener Mensch verstanden, welcher einerseits kooperativ mit Robotern agiert und zusätzlich bei Bedarf maschinenunterstützt arbeitet. Unter maschinenunterstütztes Arbeiten wird der Einsatz von intelligenten physischen und kognitiven Assistenzsystemen bzw. die adaptive Gestaltung von

vernetzten Arbeitssystemen verstanden, welche nach den Konzepten der menschlichen cyberphysischen Systeme ausgelegt und mit einer nutzeradaptiven Mensch-Maschine-Interaktion ausgestattet sind (Romero et al., 2016; Zolotová et al., 2020).

Das Konzept der Mensch-Maschine-Symbiose wurde von einigen Autoren schon früher unter dem Begriff „Adaptive Automation (AA)“ beschrieben (Calefato et al., 2008; Peter A. Hancock et al., 2013; P. A. Hancock & Chignell, 1987) und später von Romero et al. (2016) mit der Verbindung des Konzepts „Operator 4.0“ wieder aufgegriffen. Adaptive Automation zielt ebenso darauf ab, die Mensch-Maschine-Symbiose zu optimieren und die Aufgaben zwischen Menschen und Maschinen optimal zu verteilen, sodass sich das Wohlbefinden des Menschen erhöht, die Belastungen reduziert und die Produktionsziele effektiv erreicht werden. Durch eine optimale Verteilung wird flexibel auf die Fähigkeiten des Menschen eingegangen und die Vorteile der Automatisierung genutzt (Romero et al., 2016). Kaasinen et al. (2019) verbindet die Konzepte der adaptiven Gestaltung und Operator 4.0 und beschreibt dies als „Empowering the worker“ (deutsch „Befähigung des Menschen“) im Arbeitssystem.

Ein Human Cyber-Physical System ist ein Gestaltungsparadigma, welches die digitale mit der physischen Welt verknüpft und als zentrales Element den Menschen in dieses System einbringt. Für die Integration des Menschen müssen demnach geeignete Schnittstellen und Technologien eingesetzt werden, um mit dem cyberphysischen System interagieren zu können (Romero et al., 2016). Ein zentrales Ziel ist, die menschlichen Fähigkeiten zur dynamischen Interaktion mit Maschinen in der Cyber- und der physischen Welt mittels intelligenter Mensch-Maschine-Schnittstellen zu verbessern. Diese Mensch-Maschine-Interaktionsschnittstellen müssen für die kognitiven und physische Bedürfnisse der Bediener*innen konzeptioniert werden. Ein weiteres Ziel ist, die menschlichen physischen und kognitiven Fähigkeiten mittels moderner Technologien (Wearables, AR-Systeme) zu erweitern. Bei der Gestaltung von menschenzentrierten, cyberphysischen Systemen sowie der Mensch-Maschine-Interaktion soll ebenso Fokus auf die Usability und Akzeptanz sowie die Einhaltung geltender Normen im Bereich Ergonomie und Sicherheit gelegt werden. Deshalb stehen anthropometrische Merkmale und Individualisierung im Vordergrund, welche mithilfe adaptiver Systeme und eines effektiven Einsatzes von Assistenzsystemen umgesetzt werden können. Daher besteht ein weiterer Zweck adaptiver Systeme darin, den Operator 4.0 zu unterstützen, um eine nachhaltige Entlastung von Stress zu gewährleisten und Kreativität, Innovation und Improvisationsfähigkeit zu fördern (Romero et al., 2016).

Ebenso werden eine „intelligente“ Mensch-Maschine-Schnittstelle und Interaktionstechnologien gefordert, bei der noch Forschungsbedarf nach adaptiven Gestaltungslösungen bzw. sinnvollen menschenzentrierten Lösungen besteht, um die Entwicklung der „Mensch-Automatisierungs-Symbiose“ zu fördern und den Operator 4.0 in der Fabrik der Zukunft zu unterstützen (Romero et al., 2016).

Das Operator 4.0-Paradigma ist demnach ein Versuch, die menschliche Fähigkeiten und Fertigkeiten durch Ergänzung mit den Vorteilen der Maschinen, Digitalisierung und Automatisierung zu ergänzen, anstatt die menschliche Arbeitskraft durch

ausschließliche Automatisierung zu ersetzen (Romero et al., 2020). Ähnlich wie bei Industrie 4.0, wo cyberphysische Systeme die physische und digitale Welt verbinden, wird beim Begriff Operator 4.0 die cyberphysische Welt durch z. B. AR-Systeme mit den Menschen verbunden (Segura et al., 2020).

Die Typologie der Ergänzungen des Operator 4.0 werden folgend dargestellt (Romero David et al.; Zolotová et al., 2020):

- **Super strength operator:** Der Operator verwendet ein Exoskelett.
- **Collaborative operator:** Der Operator verwendet einen kollaborativen Roboter.
- **Virtual operator:** Der Operator benutzt Systeme der virtuellen Realität.
- **Augmented operator:** Der Operator nutzt die erweiterte Realität.
- **Smarter operator:** Der Operator verwendet intelligente persönliche Assistenten.
- **Social operator:** Der Operator nutzt soziale Netzwerke.
- **Analytical operator:** Der Operator verwendet Big-Data-Analysen.
- **Healthy operator:** Der Operator verwendet einen tragbaren Tracker.

Human-in-the-Loop (HITL)

Die Integration von menschlichen Tätigkeiten in cyberphysischen Systemen spielt eine entscheidende Rolle bei den Trends der Digitalisierung und fördert die flexiblen Kompetenzen des Menschen, wobei er diese mit den Vorteilen der Automatisierung verbindet. Vor allem zwei Gestaltungsprinzipien werden betrachtet, welche menschliche Aktivitäten in Produktionssysteme integrieren: „Human-in-the-Loop“ und „Human-in-the-Mesh“ (HITM). Sie stellen die Gestaltung vor unterschiedliche Anforderungen (Fantini et al., 2019). Im Folgenden wird näher auf das Konzept des Human-in-the-Loop eingegangen, das speziell für den ausführenden Operator direkt an der Maschine relevant ist (Fantini et al., 2019).

Human-in-the-Loop ist ein menschenzentriertes Gestaltungskonzept, welches aus definierten Regelkreisen (Loops) der maschinellen und menschlichen Tätigkeiten besteht und ein menschliches Eingreifen bzw. eine beidseitige Interaktion im Arbeitssystem ermöglicht (Romero et al., 2016). Es stellt demnach eine essenzielle Voraussetzung für das Erreichen einer Mensch-Automatisierungs-Symbiose dar (Sirajum Munir et al., 2013). Im HITL-Konzept sind folgende drei Arten der Regelkreise anzutreffen (Gaham et al., 2015; Romero et al., 2016):

1. Die Maschine führt die Arbeitstätigkeiten autonom durch, die der Mensch im Arbeitsprozess überwacht. Dieser hat die Möglichkeit, gegebenenfalls in das System einzugreifen.
2. Die Maschine sowie dessen intelligente Kontrollsysteme überwachen den Menschen in dessen Tätigkeit, greifen gegebenenfalls ein (z. B. bei Fehlhandlungen etc.) und geben anhand von visuellen oder akustischen Signalen Feedback.
3. Die Maschine überwacht den Menschen, jedoch sind menschliche Eingaben für die Steuerung und bei Entscheidungen stets notwendig.

Human-in-the-Loop-Systeme benötigen also einen menschlichen Part, welcher in kritischen Situationen, bei Entscheidungsprozessen sowie bei Fehlern etc. in den Loop eingreift oder nach dem spezifischen Feedback der Maschine gesteuert. Human-in-the-loop-Ansätze

werden einerseits beim maschinellen Lernen eingesetzt, bei dem der Mensch als Feedbackinstanz beim Trainieren der Daten Verbesserungen durchführt (AIM Agile IT Management GmbH), sowie bei cyberphysischen Systemen, bei denen gezielt die Interaktion zwischen Mensch und Maschine gestaltet wird (Gaham et al., 2015; Garcia et al., 2019).

Bei der Anwendung von HITL auf cyberphysische Produktionssysteme (Human-in-the-Loop Cyber-Physical Production Systems) wird im Speziellen die interaktive Dimension der Mensch-Maschine-Interaktion ins Produktionssystem oder Arbeitssystem integriert, die einen intelligenten Entscheidungsrahmen ermöglicht, der den Menschen in die cyberphysischen Loops des Produktionssystems miteinbezieht. Gaham et al. (2015) erwähnen, dass cyberphysische Produktionssysteme mit HITL-Konzepten ein hohes Potenzial für moderne Anwendungen im Produktionsbereich der Zukunft darstellen und neue Gestaltungsmöglichkeiten der Arbeitstätigkeiten und Fertigungssysteme erfordern.

Die dargestellten einzelnen Konzepte und vor allem die Verknüpfung von adaptiven Systemen, Operator 4.0 mit Konzepten der menschenzentrierten Gestaltung von Arbeitsplätzen sowie die Gestaltung der Mensch-Maschine-Interaktion stellen ein hohes Forschungspotenzial dar, welches bei der Entwicklung und Demonstration des Spatial Augmented Reality Assistenzsystems berücksichtigt werden soll.

2.2.2 Adaptive Einstellmöglichkeiten und Assembly 4.0

Reinhart (2017) beschreibt im „Handbuch Industrie 4.0“ in Anlehnung an Oppermann (1994) verschiedene Möglichkeiten der Einstellung und Anpassung von Arbeitssystemen (Abbildung 14).

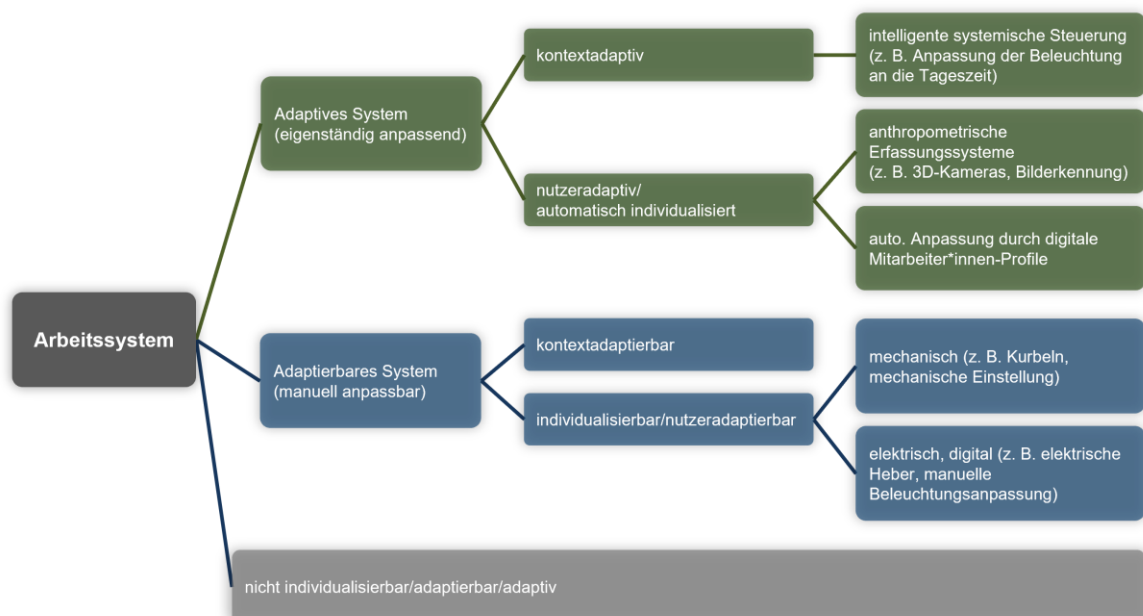


Abbildung 14 | Einstellung von Systemen, in Anlehnung an Oppermann (1994) und Reinhart (2017)

Ein Arbeitssystem kann demnach **nicht individualisierbar/adaptierbar/adaptiv** oder **adaptierbar (manuell anpassbar)** sowie **adaptiv (eigenständig anpassend)** gestaltet werden (Oppermann, 1994; Reinhart, 2017).

Einstellungsvariante: Nicht individualisierbar

Wenn ein Arbeitssystem nicht eingestellt und angepasst werden kann, entspricht dies in der Regel einem standardisierten Arbeitsplatz, welcher nach statistischen anthropometrischen Variablen nach sogenannten „Durchschnittsmenschen“ ausgelegt wird, wobei hier die individuellen Bedürfnisse und Merkmale des Menschen bei der Arbeitsplatzgestaltung nicht berücksichtigt werden können (Bullinger-Hoffmann & Mühlstedt, 2016). Die Maße des Körpers und auch die Bedürfnisse und Empfindungen eines Menschen (auch innerhalb einer Gruppe) sind individuell verschieden und daher eignen sich „Durchschnittsmenschen“ nicht zur Dimensionierung von Arbeitsplätzen (Bullinger, 1994, S. 198–203; Schlick et al., 2018). Es müssen daher Arbeitsplätze gestaltet werden, welche diese Unterschiedlichkeiten berücksichtigen, die Gesundheit nachhaltig nicht schädigen und welche psychisch sowie physisch zumutbar sind und gleichzeitig aber auch die Produktivität positiv beeinflussen. Die schnelle Veränderbarkeit der Gegebenheiten des Arbeitsplatzes muss jedoch gewährleistet werden, um die individuellen und spezifischen Merkmale des Menschen zu berücksichtigen (Lotter & Wiendahl, 2012; Schlick et al., 2018). Erst moderne Technologie, wie IoT-Systeme, intelligente Erkennung von Menschen mit Kamerasystemen und vernetzte Systeme, machen es erst möglich, die spezifischen Merkmale sinnvoll zu berücksichtigen und Arbeitssysteme adaptiv einzustellen. Diese Variante sollte, speziell auch in Kombination mit cyberphysischen Systemen, vermieden werden, weil sie der menschenzentrierten Gestaltung widerspricht.

Einstellungsvariante: Adaptierbares System (manuell anpassbar)

Die adaptierbare Gestaltung eines Arbeitssystems bedeutet, dass durch mechanische, elektrische bzw. digitale Einstellungsmöglichkeiten der Arbeitsplatz nach den individuellen Bedürfnissen oder dem Arbeitsumfeld manuell vom Menschen angepasst werden kann. Als Einstellungsmöglichkeiten können mechanische Stellhebel und Kurbeln, elektrische Stellmotoren sowie digitale Softwarelösungen herangezogen werden. Des Weiteren wird zwischen den Varianten „kontextadaptierbar“ und „nutzeradaptierbar“ differenziert. Die **kontextadaptierbare Einstellung** beschreibt die Anpassung an das Arbeitsumfeld (z. B. manuelle Einstellung der Arbeitshöhe nach den verschiedenen Arbeitstätigkeiten oder Einstellung der Beleuchtung nach Helligkeit der Arbeitsumgebung). Die **nutzeradaptierbare Einstellung** beschreibt, dass hier der Mensch mit seinen individuellen Bedürfnissen und Merkmalen (z. B. Einstellung nach Körpergröße etc.) im Mittelpunkt steht.

Schlund et al. (2018) stellen dazu erste Überlegungen zu verschiedenen Individualisierungsmöglichkeiten im Kontext eines industriellen Montagearbeitsplatzes vor. Individualisierung im Bereich Gestaltung der Mensch-Maschine-Interaktion ist nach DIN EN ISO 9241-171 als „Modifizierung von Interaktion und Informationsdarstellung, um individuellen Fähigkeiten und Bedürfnissen von Benutzern gerecht zu werden“ definiert. Individualisierung bezieht sich demnach auf die Modifizierung, welche erreicht werden kann, ohne dabei die Anwendung umzuprogrammieren, weil die Individualisierungsmöglichkeiten bereits in der Anwendung integriert sind (DIN EN ISO 9241-171). Individualisierbarkeit in Bezug zur Mensch-System-Interaktion ist gegeben, wenn Benutzer diese Interaktion sowie die Darstellung von Informationen ändern können, um diese dadurch an ihre individuellen

Fähigkeiten und Bedürfnisse anzupassen. Die Möglichkeiten der Individualisierung sollen jedoch stets in bestimmten Grenzen möglich sein und die ergonomischen Bedingungen dabei eingehalten werden. Individualisierung im Kontext der Gestaltung kann ein wesentliches Mittel sein, um die Barrieren der Zugänglichkeit des Systems zu senken und dadurch einem großen Benutzer*innenkreis entgegenzukommen (DIN EN ISO 9241-110:2020).

Dementsprechend bedeutet das, dass das eingesetzte Assistenzsystem individuell an die Gegebenheiten, Bedürfnisse und anthropometrischen Merkmale des Menschen angepasst und Maßnahmen zur Umsetzung von Individualisierung gestaltet werden sollen.

Einstellungsvariante: Adaptives System (eigenständig anpassend)

Adaptive Systeme können sich je nach gewünschter Vorkonfiguration eigenständig an spezifische Merkmale anpassen. Dabei wird wiederum zwischen kontextadaptiv und nutzeradaptiv unterschieden. Kontextadaptive Systeme verwenden Sensoren, um die Arbeitsumgebung zu erfassen, spezielle Softwarelösungen, um die Daten zu analysieren und mit der gewünschten Konfiguration zu vergleichen, sowie vernetzte Aktoren, welche die Befehle sinngemäß umsetzen (z. B. Analyse des Tagesverlaufes durch Lichtsensoren, Vergleichsmessung der Helligkeit des Arbeitsbereiches und automatische Anpassung, ohne dass der Mensch hierzu manuelle Einstellungen vortreffen muss). Nutzeradaptive Systeme stellen sich demnach automatisch auf die Merkmale und Bedürfnisse des Menschen ein.

Taxonomie der nutzeradaptiven Gestaltung in der Baustellenmontage

Damit sich der Montagearbeitsplatz nach dem User und dessen Merkmalen nutzeradaptiv und eigenständig einstellt, ist es notwendig, Möglichkeiten und Konzepte aufzuzeigen, wie der Mensch und dessen Merkmale erkannt und identifiziert werden und wie sich das Montagesystem adaptiv anpasst. Rupprecht und Schlund (2021) erweitern dazu die Überlegungen zu „Möglichkeiten der Gestaltung individualisierbarer Montagearbeitsplätze“ (Schlund et al., 2018), um nutzeradaptive Erkennungs- und Identifikationsmöglichkeiten sowie nutzeradaptive Anpassungsmöglichkeiten im Kontext der industriellen Großgeräte- und Baustellenmontage (Tabelle 6).

Tabelle 6 | Taxonomie der nutzeradaptiven Gestaltung, in Anlehnung an Rupprecht und Schlund (2021)

Taxonomie der nutzeradaptiven Gestaltung		
Dimensionen und Einstellmöglichkeiten	Erweiterung: Nutzeradaptive Erkennungs- und Identifikationsmöglichkeiten	Erweiterung: Nutzeradaptive Möglichkeiten zur Anpassung des Montagearbeitsplatzes
Nutzeradaptive Einstellmöglichkeiten WAS kann am Baustellenmontagearbeitsplatz nutzeradaptiv gestaltet werden?	WIE können Menschen adaptiv erkannt und identifiziert werden? WIE können deren individuellen Merkmale und Präferenzen zugewiesen, gespeichert und abgerufen werden?	WOMIT/WODURCH können die Einstellmöglichkeiten adaptiv nach den individuellen Merkmalen und Präferenzen am Arbeitsplatz umgesetzt werden?
EM 1 – Adaptive Arbeitsfläche Arbeitshöhe Arbeitswinkel/-rotation	Möglichkeit 1: Adaptive Identifikation des Menschen und Abruf der zugehörigen vorab gespeicherten individuellen Merkmale aus digitalen Personenprofilen Adaptive Identifikation, um welche Person es sich handelt, durch -) personalisierten RFID-Chip in der Arbeitskleidung und Erkennung durch RFID-Gateways, -) Identifikation der Person durch Bilderkennung und Abgleich mit Mitarbeiter*innen-Datenbank-Gesichtsscanner, Iris-Scanner, Fingerprint-Scanner.	EM 1 – Adaptive Arbeitsfläche Automatische Einstellung der Arbeitsfläche in Höhe, Winkel und Rotation mit angesteuerten -) elektrischen Antrieben und -) elektro-pneumatischen/elektro-hydraulischen Zylindern und Antrieben.
EM 2 – Adaptive Anordnung physischer Gegenstände und Materialbereitstellung Physische Gegenstände am Arbeitsplatz (Werkzeuge, Behälter, Hilfsmittel ...) Anordnung der physischen Materialbereitstellung Anordnung von physischen Bedienelementen		EM 2 – Adaptive Anordnung physischer Gegenstände und Materialbereitstellung Physische Anordnung/Materialbereitstellung -) mit selbstfahrenden Materialwägen, -) durch flexible und kollaborationsfähige Roboter, -) durch fahrerlose Transportsysteme, -) durch automatische Seilzüge/Hängebahnen.

EM 3 – Adaptive Anordnung und Design digitaler Elemente/User-Interface Struktur, Anordnung und Design Bedienoberfläche/Bedienelemente/ User-Interface Softwareauswahl	Abruf und Zuweisung der individuellen Merkmale durch -) manuelle Vorabfassung und -speicherung der individuellen Merkmale und persönlichen Präferenzen in digitalen Personenprofilen und Abruf und Zuweisung nach Identifikation des Menschen. Gespeicherte Merkmale in digitalen Personenprofilen (demonstrativ) durch -) anthropometrische Merkmalen (Körpergröße, Gewicht, BMI, Proportionen ...), -) persönliche Präferenzen (z. B. präferierte Arbeitshöhe, Anordnung von Elementen, Stärke der Assistenz, gewünschte Beleuchtungsstärke etc.).	EM 3 – Adaptive Anordnung digitaler Elemente/User-Interface durch -) dynamische Elemente mit Projektionssystemen/Spatial Augmented Reality Systemen, -) personalisierte User-Interfaces, Designs auf Tablets/Bildschirme, -) personalisierte Elemente beim Einsatz von Datenbrillen/AR, -) personalisierte Wearables (Smartwatches etc.), Apps für Smartphones.
EM 4 – Adaptive Beleuchtungsparameter Beleuchtungsstärke Beleuchtungsart (passiv, aktiv, natürlich, künstlich) Lichtfarbe/Lichttemperatur Anordnung der Beleuchtungsquellen	Möglichkeit 2: Adaptive Erkennung des Menschen sowie dessen individueller Merkmale Adaptive Erfassung, ob es sich um einen Menschen handelt, durch -) Kamerasysteme (Stereokamera, 3D-Kamera) und Bilderkennungsalgorithmen (z. B. YOLO, OpenCV), -) Trackingsysteme (Markertracking). Adaptive Erfassung der zugehörigen individuellen Merkmale durch -) Erkennung des Menschen inklusive von anthropometrischen Merkmalen, Bewegung, Posen und Gesten. -) Keine Erfassung/Speicherung und Abruf von persönlichen Präferenzen möglich. -) Keine Vorabspeicherung von anthropometrischen Merkmalen (Live-Daten).	EM 4 – Adaptive Beleuchtungsparameter -) personalisierte, digitale Lichtsteuerungssysteme, Apps; -) intelligente, smarte LED-Panels, LED-Inserts; -) moderne Tageslichtzufuhrsysteme; -) Sensoren für Wetterverhältnisse, Tag und Nacht; -) adaptive Blaulichtfilter, Graulichfilter.
EM 5 – Adaptive klimatische Parameter Temperatur Luftfeuchtigkeit Belüftungsparameter (Richtung, Stärke, Strömungsart) Reinheit und Fremdpartikelkonzentration	Adaptive Erfassung der zugehörigen individuellen Merkmale durch -) Erkennung des Menschen inklusive von anthropometrischen Merkmalen, Bewegung, Posen und Gesten. -) Keine Erfassung/Speicherung und Abruf von persönlichen Präferenzen möglich. -) Keine Vorabspeicherung von anthropometrischen Merkmalen (Live-Daten).	EM 5 – Adaptive klimatische Parameter -) digitale Belüftungs- und Temperatursteuerung; -) gekoppelte Sensoren am Arbeitsplatz, digitale Personenprofile; -) selbstfahrende mobile Belüftungssysteme, Roboterarme; -) individualisierbare, vernetzte Luftbe- und -entfeuchter, Ventilatoren; -) individualisierbare digital vernetzte Heizungs- und Kühlgeräte.
EM 6 – Adaptive akustische Parameter Geräuschpegel (dB) Frequenzen Musik, akustische Informationen	EM 6 – Adaptive akustische Parameter -) digitale Steuerung von Lärmschutzmaßnahmen, -) selbstfahrende Lärmschutzmaßnahmen, -) bewegende Lärmschutzmaßnahmen mit Robotern, -) Schallsensoren am Arbeitsplatz, -) automatische Noise-Cancelling-Kopfhörer, -) digitale Steuerung von akustischen Informationssystemen, vernetzte Kopfhörer.	EM 6 – Adaptive akustische Parameter -) digitale Steuerung von Lärmschutzmaßnahmen, -) selbstfahrende Lärmschutzmaßnahmen, -) bewegende Lärmschutzmaßnahmen mit Robotern, -) Schallsensoren am Arbeitsplatz, -) automatische Noise-Cancelling-Kopfhörer, -) digitale Steuerung von akustischen Informationssystemen, vernetzte Kopfhörer.
EM 7 – Adaptive (Arbeits-) Informationssysteme / Digitale Assistenzsysteme Informationsinhalt Informationsart (visuell, auditiv, kinästhetisch, olfaktorisch; physisch, digital) Informationsbereitstellungsort (Bildschirm, Wearables, Data Glasses etc.) Informationsumfang/-granularität (Häufigkeit, Dauer, Detailtiefe)	EM 7 – Adaptive (Arbeits-) Informationssysteme / Digitale Assistenzsysteme -) personalisierte Informationssysteme (Spatial AR, AR, Wearables, Bildschirme etc.); -) personalisierte Informationsbereitstellungsorte (direkt am Arbeitsplatz, an zentralen Terminals, auf Wearables etc.); -) personalisierte Informationsinhalte, -granularität, -details etc., je nach Qualifikation, Erfahrung etc.	EM 7 – Adaptive (Arbeits-) Informationssysteme/digitale Assistenzsysteme -) personalisierte Informationssysteme (Spatial AR, AR, Wearables, Bildschirme etc.); -) personalisierte Informationsbereitstellungsorte (direkt am Arbeitsplatz, an zentralen Terminals, auf Wearables etc.); -) personalisierte Informationsinhalte, -granularität, -details etc., je nach Qualifikation, Erfahrung etc.
EM 8 – Adaptive physische Assistenzsysteme Nutzungsdauer der Assistenz (Häufigkeit, Dauer) Verwendung der Assistenz (Ja/Nein) Stärke der Assistenz Art der Assistenzsysteme	EM 8 – Adaptive physische Assistenzsysteme -) personalisierte Nutzungsdauer und Häufigkeit, -) Entscheidungsassistenz für Verwendung Ja/Nein/optional, -) personalisierte Stärke der Assistenz durch automatische Aktoren, -) automatisierte „Just in time“-Assistenzsysteme (richtiger Ort, richtige Zeit, richtiger Prozess).	EM 8 – Adaptiver Einsatz von Assistenzsystemen -) personalisierte Nutzungsdauer und Häufigkeit, -) Entscheidungsassistenz für Verwendung Ja/Nein/optional, -) personalisierte Stärke der Assistenz durch automatische Aktoren, -) automatisierte „Just in time“-Assistenzsysteme (richtiger Ort, richtige Zeit, richtiger Prozess).
EM 9 – Adaptive Mensch-System-Interaktion Art der Interaktion (Eingabe/Ausgabe) Gestaltung der Interaktion Höhe der Systemvorgabe Höhe der Arbeitsteilung zw. Mensch und Maschine	EM 9 – Adaptive Mensch-System-Interaktion -) personalisierte Interaktionsmöglichkeitsart (Eingabe- /Ausgabesystem), -) personalisierte Einstellung der Systemvorgabe, -) personalisierte Einstellung der Arbeitsteilung.	EM 9 – Adaptive Mensch-System-Interaktion -) personalisierte Interaktionsmöglichkeitsart (Eingabe- /Ausgabesystem), -) personalisierte Einstellung der Systemvorgabe, -) personalisierte Einstellung der Arbeitsteilung.
EM 10 – Adaptive Arbeitsorganisation Taktlänge/Taktzeit Arbeitszeit Arbeitsfähigkeit/Arbeitsinhalt Arbeitsteilung allgemein	EM 10 – Adaptive Arbeitsorganisation -) personalisierte Taktlänge durch flexible Fahrerlose Transportsysteme, Rollwägen; -) personalisierte Arbeitszeit durch flexible Schichtmodelle, Gleitzeiten, Steuerung mittels mobiler App; -) randomisierte Arbeitsteilung, „Job Rotation“, „Enlargement“, Einsatz von Assistenzsystemen.	EM 10 – Adaptive Arbeitsorganisation -) personalisierte Taktlänge durch flexible Fahrerlose Transportsysteme, Rollwägen; -) personalisierte Arbeitszeit durch flexible Schichtmodelle, Gleitzeiten, Steuerung mittels mobiler App; -) randomisierte Arbeitsteilung, „Job Rotation“, „Enlargement“, Einsatz von Assistenzsystemen.

Dabei werden die Dimensionen und Einstellmöglichkeiten des Montagearbeitsplatzes sowie die verschiedenen Varianten der Erkennungs- und Identifikationsmöglichkeiten dargestellt. Des Weiteren werden die Umsetzungsmöglichkeiten der nutzeradaptiven Anpassung zu den jeweiligen Dimensionen dargestellt (Rupprecht & Schlund, 2021). Die adaptiven Umsetzungsmöglichkeiten der Dimensionen EM3, EM7 und EM9 sind dabei von besonderer Relevanz und dienen als Input für die Entwicklung des Spatial Augmented Reality Systems.

Assembly 4.0

Bortolini et al. (2017) betonen, dass die effektive Integration von intelligenten Systemen sowie die digitale Vernetzung im Montageprozess (Montagestation, Mensch, Objekte, Werkzeuge u. a.) der Schlüsselfaktor zur Bewältigung der Herausforderungen sind und summiert dies unter dem Begriff „**Assembly System 4.0 – AS40**“ (Abbildung 15).

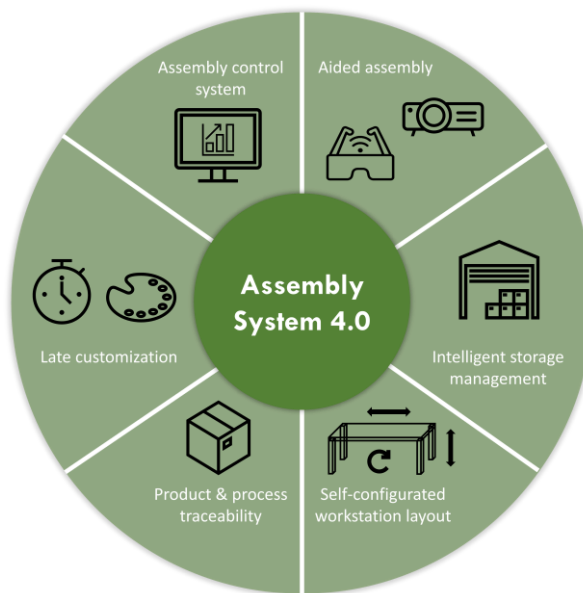


Abbildung 15 | Charakteristik der Assembly 4.0-Montagestation, in Anlehnung an Bortolini et al. (2017)

Als ein essenzielles Merkmal von AS40 wird die **„selbstkonfigurierte Arbeitsstation“** genannt, welche sich in den Abmessungen, unter Berücksichtigung des Montageproduktes und der spezifischen anthropometrischen Merkmale der Mitarbeiter*innen, adaptiv anpasst. Diese selbstkonfigurierte Arbeitsplatzgestaltung hat das Ziel, die Montagetätigkeiten zu optimieren, die Prozesszeit zu verringern sowie die Ergonomie zu verbessern (Bortolini et al., 2017). Bortolini et al. (2019) präsentieren das „Self-Adaptive Smart Assembly System (SASAS)“, welches selbst in der Lage ist, sich in Echtzeit nach den Gegebenheiten des Arbeitsprozesses zu rekonfigurieren. Fraunhofer IPA (2018) stellt ein adaptives Montageassistenz- und Interaktionssystem mit dem Projekttitel „MonSiKo“ vor, bei dem durch moderne 3D-Sensorik ein Roboterarm sowie moderne Kommunikationstechnologie (Spracherkennung, Touchscreen-Eingabe) die Montageschritte überwachen und Fehler erkannt werden. Weitere Aspekte von AS40 sind der Einsatz von intelligent vernetzten Werkzeugen, welche z. B. das richtige Drehmoment einstellen oder dokumentieren, ob eine Schraube korrekt verbaut wurde, sowie ein intelligentes digitales „Assembly control system“ zur Steuerung und Überwachung der Montagestation, die vernetzte Materialbereitstellung, die „Traceability“ von Produkten und Prozessen sowie die Möglichkeit, dass Produkte erst spät im Produktionsprozess individualisiert und an die Bedürfnisse der Kunden angepasst werden (Bortolini et al., 2017).

Ein weiterer essenzieller Aspekt von Assembly 4.0 ist **„Aided Assembly“**, das darauf abzielt, moderne digitale Assistenzsysteme zur Erweiterung der Realität (AR-Systeme) einzusetzen sowie Arbeitsinformationen in neuartiger Form zur Verfügung zu stellen (Bortolini et al., 2017). In der vorliegenden Forschungsarbeit wird der Fokus auf **Aided Assembly** und die Entwicklung eines **Spatial Augmented Reality Assistenzsystems** gelegt, welches eine moderne Informationsbereitstellung inklusive menschenzentrierter Interaktion in der Montage ermöglicht und dementsprechend die Montagestation „erweitert“.

2.2.3 Vision der nutzeradaptiven Gestaltung in der Großgerätemontage

Ausgehend von den Überlegungen der nutzeradaptiven Erkennungs-, Einstell- und Anpassungsmöglichkeiten und Darstellung der Taxonomie wird folgend die **Vision für die nutzeradaptive Gestaltung in der industriellen Großgeräte- und Baustellenmontage** formuliert und als schematisches Rendering (Abbildung 16) dargestellt (Mayrhofer, Rupprecht & Schlund, 2019; Rupprecht & Schlund, 2021; Schlund et al., 2018).

Vision der nutzeradaptiven Gestaltung in der Großgeräte- und Baustellenmontage:

„Ein nutzeradaptiver Montagearbeitsplatz, welcher sich nach den individuellen anthropometrischen Merkmalen sowie persönlichen Präferenzen des Nutzers adaptiv einstellt und den Menschen durch moderne Assistenzsystemen bei dessen Tätigkeiten physisch und kognitiv unterstützt“.

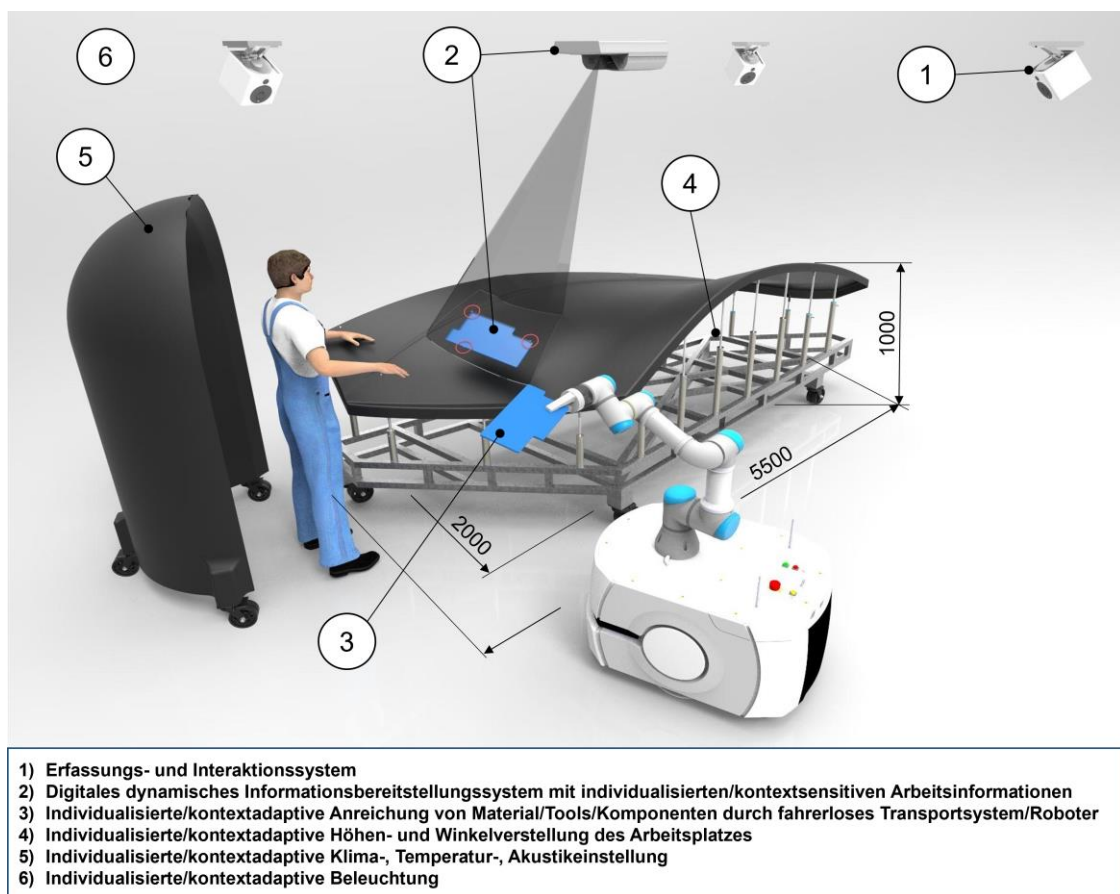


Abbildung 16 | Vision der nutzeradaptiven Gestaltung des Baustellenmontagearbeitsplatzes, in Anlehnung an Mayrhofer, Rupprecht und Schlund (2019)

Die Darstellung umfasst das Erfassungs- und Interaktionssystem (1), welches dazu verwendet wird, um den Arbeitsraum der Baustellenmontage zu analysieren. Dabei werden die Positionen des Menschen sowie weiterer mobiler Plattformen (z. B. fahrerlose Transportsysteme) erfasst und Interaktionen mit dem System ermöglicht. Das digitale dynamische Informationsbereitstellungssystem (2) wird verwendet, um kontextsensitive und individuell aufbereitete Arbeitsdokumente, Trainings- und Lerndaten sowie weitere digitale Dokumente dynamisch auf der Montagestation oder am Boden anzuzeigen. Die Anreicherung von

Materialien oder Komponenten (3) durch z. B. ein fahrerloses Transportsystem inklusive aufgebauten Roboterarmen dient dazu, die richtigen Objekte individuell und je nach Arbeitsschritt just in time zur Verfügung zu stellen. Das Rendering stellt erste Überlegungen dar, wie das Montagesystem individualisiert und adaptiv in Höhe und Winkel (4) angepasst werden kann. Im Weiteren wird eine individualisierte adaptive Klima-, Temperatur- und Akustikeinstellung (5) mit z. B. einer sich mitbewegenden Kabine dargestellt. Abschließend werden Konzepte für die adaptive Anpassung der Beleuchtung (6) präsentiert (Mayrhofer, Rupprecht & Schlund, 2019; Rupprecht & Schlund, 2021; Schlund et al., 2018). Ausgehend von der Vision wird folgend der Fokus auf den Einsatz kognitiver Assistenzsysteme zur Informationsbereitstellung sowie der Mensch-Maschine-Interaktion gelegt, der Grundlage für das Design und die Demonstration des Spatial Augmented Reality Assistenzsystems darstellt (Abbildung 17).

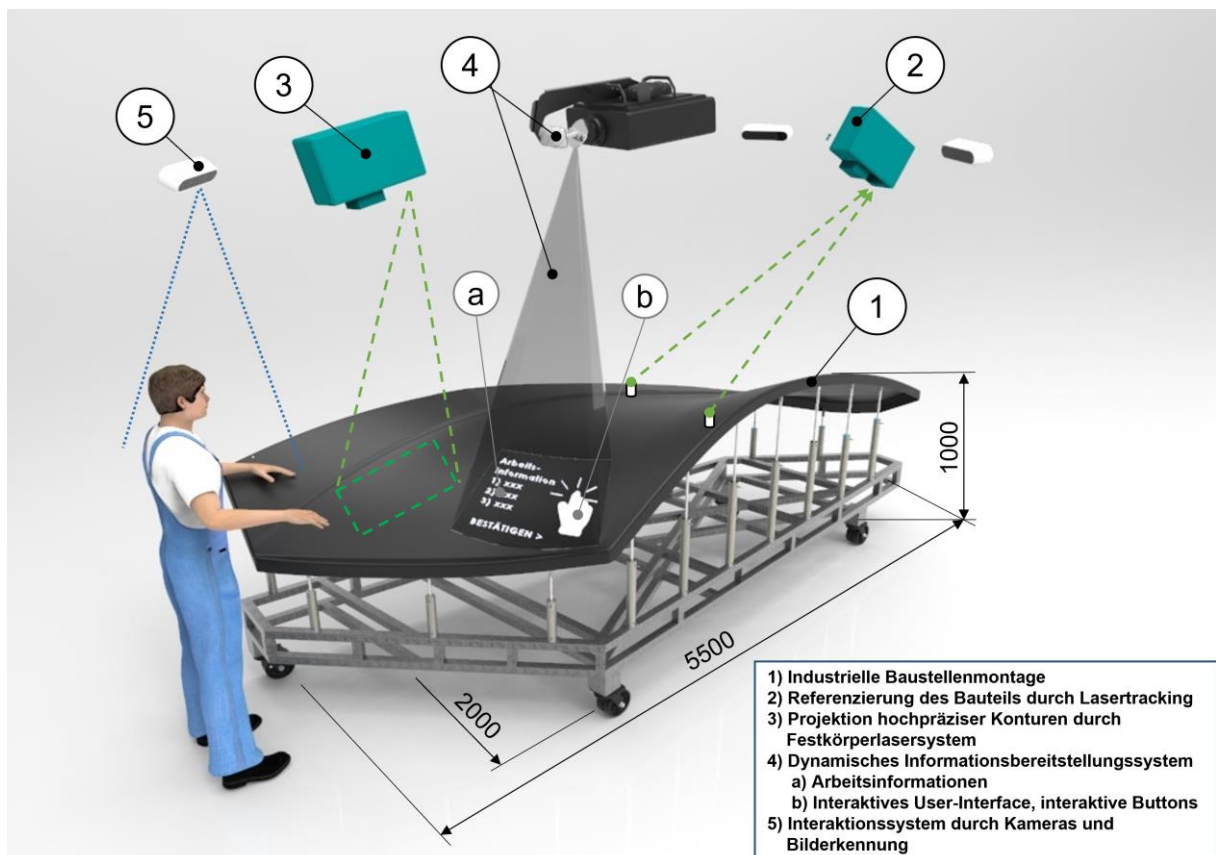


Abbildung 17 | Assistenzsysteme mit Fokus auf Informationsbereitstellung und Mensch-Maschine-Interaktion (Eigendarstellung)

Ein Großgeräte- bzw. Baustellenmontagearbeitsplatz (1) wird mit einem State-of-the-Art-Festkörperlaser-System (2, 3) zur Anzeige von 2D-Konturinformationen dargestellt. Ebenso werden Ansätze des Spatial Augmented Reality Assistenzsystems mit dynamischem Projektionssystem (4) zur Anzeige von Arbeitsinformationen (4a) sowie für die Projektion von interaktiven User-Interfaces bzw. interaktiven Buttons (4b) und Kameras als Interaktionssystem (5) zur Erfassung des Arbeitsbereiches präsentiert.

Die Überlegungen dienen als Vorlage für die Entwicklung des Spatial Augmented Reality Assistenzsystems mit nutzeradaptiver Mensch-Maschine-Interaktion in der industriellen Baustellenmontage. Die Gestaltung des Systems, unter Verwendung eines dynamischen Projektionssystem sowie dem Einsatz von Kameras und Bilderkennung, werden folgend im Detail dargestellt.

2.3 Spatial Augmented Reality zur Informationsbereitstellung

Aided Assembly (vgl. Kapitel 2.2.2) hat den zentralen Aspekt, dem Menschen in der Montage Informationen mit modernen Assistenzsystemen bereitzustellen und demnach die Realität zu erweitern (Bortolini et al., 2017). Des Weiteren werden in Unternehmen und Forschungseinrichtungen bereits Assistenzsysteme (z. B. AR-Systeme etc.) eingesetzt oder in Pilotlinien getestet (Bornmann et al., 2015; Gorecky et al., 2013; Schembera-Kneifel & Keil, 2016). Abbildung 18 zeigt einen Überblick über verschiedene Arten der Informationsbereitstellungsmöglichkeiten für Montagearbeitsplätze und insbesondere für die Großgeräte- und Baustellenmontage (Rupprecht et al., 2020).

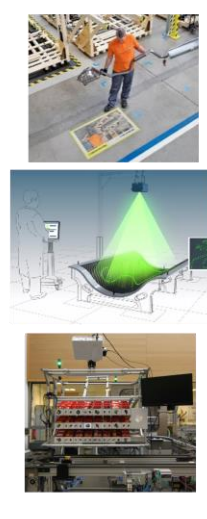
Informationsbereitstellungssysteme					
Papierdokumente, Mappen, Flip Charts, Whiteboards	Stationäre PC-Terminals, Bildschirme, Smartboards	Mobile Geräte/ Wearables (Smartphones, Tablets, Smart Watches)	Head-Mounted Displays (HMD) Augmented Reality, Virtual Reality	Spatial Augmented Reality mit statischem Projektionssystem	Spatial Augmented Reality mit dynamischem Projektionssystem
					

Abbildung 18 | Informationsbereitstellungsmöglichkeiten in der Großgeräte- und Baustellenmontage, in Anlehnung an Rupprecht et al. (2020)

Heutzutage werden großteils noch Papierdokumente in Form von One-Pager oder Dokumentationen und Arbeitsanweisungen in Papiermappen für die Informationsbereitstellung eingesetzt. Des Weiteren werden Flipcharts und Whiteboards für die Arbeits- und Teamorganisation verwendet. Im Bereich der digitalen Bereitstellung sind PC-Terminals, (Überkopf-)Bildschirme oder vereinzelt auch Smartboards in den Arbeitsstationen anzutreffen. Darüber hinausgehend wird in Pilotlinien und Entwicklungsbereichen der Unternehmen sowie in Forschungseinrichtungen und Pilotfabriken der Universitäten mit Wearables, Datenbrillen

und Smartphones sowie intelligenten Augmented Reality-Systemen geforscht (ARENA2036 e.V., 2020; Technische Universität Wien - Pilotfabrik für Industrie 4.0, 2020). Außerdem werden Projektoren zur Anzeige von Arbeitsanweisungen und Schritt-für-Schritt-Anleitungen und Technologien wie RFID zur Erkennung des Montagestatus verwendet (Funk, 2016; Rupprecht et al., 2020; Rupprecht et al., 2021). Kommerzielle Lösungen sind u. a. vom Unternehmen Assembly Solutions GmbH zu finden, das Arbeitsstationen mit statischen Projektoren zur Informationsanzeige anbietet (Assembly Solutions GmbH, 2020). Unternehmen im Bereich Aerospace und Automotiv, wie Boeing (Scott, 2017), Airbus (Wright, 2017) und General Electric (Abraham, M. & Annunziata, M., 2017) forschen bereits an einzelnen Lösungen mittels Augmented Reality unter der Verwendung von Datenbrillen für den Einsatz in deren Montageprozessen, auch für Großgerätemontagen (Fletcher et al., 2020; Medina et al., 2019).

Industrieforschungen wie bei General Electric zeigen, dass durch den Einsatz von Augmented Reality bei Großobjekten eine Produktivitätssteigerung bei der Verkabelung von Windturbinen von 25 % erreichbar ist sowie im Logistikbereich eine Steigerung der Prozesszeit von über 46 % im Vergleich zu Papierlisten ermöglicht werden kann (Abraham, M. & Annunziata, M., 2017). Obwohl Wearables und Datenbrillen für den Einsatz in der Produktion bereits möglich sind und bereits im Mehrmaschinenbetrieb oder bei Service- und Wartungsarbeiten eingesetzt werden (Hořejší, 2015), bestehen speziell beim Einsatz in der Montage Herausforderungen hinsichtlich Handhabbarkeit, eingeschränktem Sichtfeld sowie Bedienung und Tragekomfort (Mach et al., 2018).

Andere Möglichkeiten, um Informationen anzuzeigen, sind Pick-by-Vision-Lösungen und „Motion Controller“, bei denen Gegenstände visuell hervorgehoben werden und erkannt wird, wann ein falsches Teil entnommen wurde. Diese Systeme verwenden die Augmented Reality Technologie in Form eines übergeordneten Projektors für die Projektion von Arbeitsinformationen, Bilder und Videos. Außerdem wird das Konzept der In-situ-Projektion eingesetzt, d. h. die Projektion projiziert genau dort eine visuelle Information, wo ein Teil entnommen werden soll. Durch diese Systeme können ebenso die nächsten Schritte und Tasks projiziert werden, welche anschließend z. B. mittels einer Geste bestätigt werden (Funk, 2016; Funk, Kosch et al., 2015; Funk, Kosch & Schmidt, 2016). Mithilfe von Kamerasystemen können die Positionen der Hände sowie auch die Identifizierung der Teile umgesetzt werden (Bertram et al., 2018).

Aus den bestehenden Forschungen und dargestellten Entwicklungen zu Assistenzsystemen in der Montage konnten folgende Punkte festgestellt werden:

- Das Konzept „Aided Assembly“ mit dem Einsatz von modernen Assistenzsystemen zur Informationsbereitstellung sind zentrale Aspekte in der nutzeradaptiven Montage und Assembly 4.0 (Bortolini et al., 2017).
- Die Forschungen und Entwicklungen beziehen sich vermehrt auf die Montageform der Einzelplatzmontage mit kleinem Arbeitsbereich (Funk, 2016; Funk, Kosch et al., 2015; Funk, Kosch & Schmidt, 2016).

- Augmented Reality Systeme mit statischen Projektoren sowie der Einsatz von Kamerasystemen werden vor allem in der Einzelplatzmontage und Kleingeräte-montage angewendet (Funk, 2016; Funk, Kosch et al., 2015; Funk, Kosch & Schmidt, 2016), in der Großgerätemontage werden dies noch kaum eingesetzt.
- Es werden bereits vermehrt AR-Systeme mit Datenbrillen in der industriellen Großgerätemontage eingesetzt und dabei Produktivitätssteigerungen aufgezeigt (Abraham, M. & Annunziata, M., 2017; Scott, 2017; Wright, 2017).

Ausgehend von den Erkenntnissen der aktuellen Forschung und industriellen Anwendungen wird der Fokus auf die Gestaltung von nutzeradaptiven Konzepten für Montagestationen gelegt und bestehende Ansätze zu Augmented Reality mit Projektoren (Spatial Augmented Reality) herangezogen und für die Anforderungen der Großgeräte- und Baustellenmontage adaptiert. In den folgenden Abschnitten werden Anwendungen von Spatial Augmented Reality präsentiert, Möglichkeiten der Informationsbereitstellung mit SAR erläutert und Konzepte der dynamischen In-situ-Projektion sowie echtzeitfähigen Interaktion mit Computer Vision dargestellt.

2.3.1 Spatial Augmented Reality in der Montage

Augmented Reality stellt eine Technologie dar, welche die Realität des Menschen durch „Erweiterung“ mit digitalen Elementen (z. B. Informationen, Grafiken, Animationen, 3D-Objekte etc.) bereichert und dadurch bei dessen Tätigkeiten (z. B. in der Montage) unterstützt (Azuma, 1997). Als Devices zur Integration der digitalen Elemente werden vor allem Smartphones, Head-Mounted Displays und AR-Brillen verwendet.

Der Begriff „Augmented Reality“ geht auf die Studie „A Survey of Augmented Reality“ von Roland T. Azuma zurück und definiert ein AR-System mit folgender Charakteristik (Azuma, 1997):

- Kombination von realer und virtueller Welt
- Möglichkeit der Interaktionen in Echtzeit
- Möglichkeit der dreidimensionalen Registrierung der virtuellen Elemente mit der realen Welt

Die Kombination der realen und virtuellen Welt sowie die Überlagerung und Erweiterung der Realität mit virtuellen Elementen ist der Schlüsselpunkt des AR-Systems. Dabei können nicht nur die visuellen, sondern auch die akustischen, haptischen, gustatorischen oder olfaktorischen Sinne erweitert werden. Die Interaktion in Echtzeit bedeutet, dass die reale Umgebung in nahezu Echtzeit erfasst wird und die virtuellen Elemente dementsprechend schnell dahingehend angepasst werden. Die dreidimensionale Registrierung der virtuellen Elemente bedeutet, diese so in der realen Welt zu verankern, dass sie unabhängig von der Position des Nutzers betrachtet werden können (Azuma, 1997).

Eine besondere Ausprägung der Augmented Reality ist dabei **Spatial Augmented Reality**, also die „**räumlich erweiterte Realität**“, welche das Tragen von Head-Mounted Displays, Datenbrillen, Wearables, Smartphones oder sonstigen Geräten nicht benötigt (Funk, 2016;

Mengoni et al., 2018). Spatial Augmented Reality greift auf den Einsatz von Projektoren zurück, welche meist über dem Projektionsobjekt (z. B. Montagestation) angebracht sind. Damit können relevante Montageinformationen direkt auf der Montagefläche angezeigt werden. Die Interaktionskomponente in Echtzeit wird mittels Kamerasystemen und vernetzten Sensoren umgesetzt (Funk, 2016; Sochor et al., 2019).

Das erste Spatial Augmented Reality Assistenzsystem, welches Projektoren sowie Kamerasysteme kombinierte, stellt *Wellner's Digital Desk* (Abbildung 19) dar (Wellner, 1993).

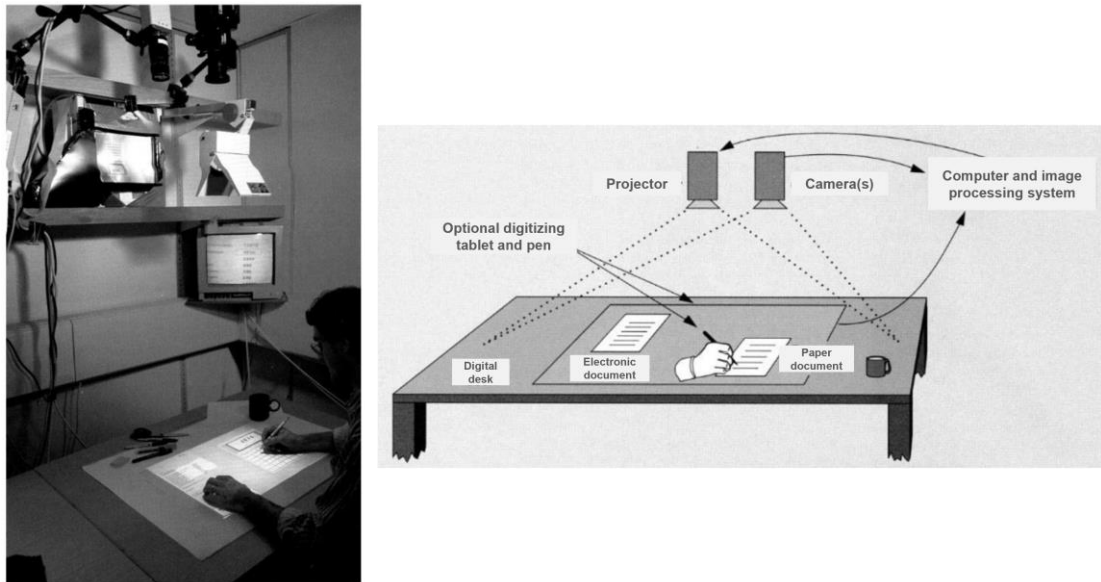


Abbildung 19 | Erstes Spatial Augmented Reality Assistenzsystem *Wellner's Digital Desk*, in Anlehnung an Wellner (1993)

Dabei wird über einen Arbeitstisch eine Kamera und ein Projektor platziert und Bilder über einen Computer eingespielt. Der Mensch bekommt dabei direkt auf seiner Arbeitsfläche ein elektronisches Dokument projiziert, worin er Informationen für die Umsetzung seiner Tätigkeiten findet (Wellner, 1993). Heutzutage werden vor allem in der Einzelplatzmontage mittels Spatial Augmented Reality Assistenzsystemen statische Informationen direkt im Arbeitsbereich angezeigt. Laut dem Fraunhofer-Institut für Gießerei-, Composite- und Verarbeitungstechnik (IGCV) können diese Informationen aus folgenden Kategorien bestehen (Sochor et al., 2019):

- Montagevideos
- textliche Montageanweisungen
- akustische Informationen
- Montagebilder
- Zeitanzeige und Prozessfortschritte
- Anzeige von Prozessschritten

Spatial Augmented Reality bietet demnach eine komfortable und vor allem eine barrierefreie Möglichkeit (kein Tragen von Devices), den Menschen bei dessen Arbeitstätigkeiten mit gezielter Projektion von Informationen zu unterstützen sowie eine direkte Interaktion vor Ort zu ermöglichen (Mengoni et al., 2018).

In folgenden Abschnitten werden der Stand der Forschung von statischen Spatial Augmented Reality Assistenzsystemen sowie neue Entwicklungen unter der Verwendung von dynamischen Projektionssystemen dargestellt. Letzteres ist für den Einsatz in der Großgeräte- und Baustellenmontage zur dynamischen Anzeige von Informationen über einem großen Arbeitsbereich sowie direkt am Montageort von hoher Bedeutung (vgl. Kapitel 1.2.2).

Statisches Spatial Augmented Reality Assistenzsystem

Bei statischen Spatial Augmented Reality Assistenzsystemen befindet sich der Projektor statisch über dem Projektionsbereich, dabei wird die Projektion direkt auf der Projektionsfläche (z. B. Arbeitsfläche der Einzelplatzmontage) angezeigt.

Statische Laserkontur-Projektoren sind Stand der Technik in der Baustellenmontage speziell beim Layup-Prozess von Carbonfaser-Matten (CFK) (Abbildung 20) in der Flugzeugkomponentenmontage oder für die Anzeige von einfärbigen 2D-Konturen für das Zuschneiden von Holzbauteilen in der Holzindustrie (Funk, 2016; SL-Laser GmbH, 2020a). Für die Anzeige der 2D-Informationen (2D-Konturen, Texte etc.) am Bauteil werden spezielle Projektoren mit einem 532 Nanometer grünen diodengepumpten Festkörperlaser eingesetzt (Aerospace Manufacturing Magazine - MIT Publishing, 2018; SL-Laser GmbH, 2020a).

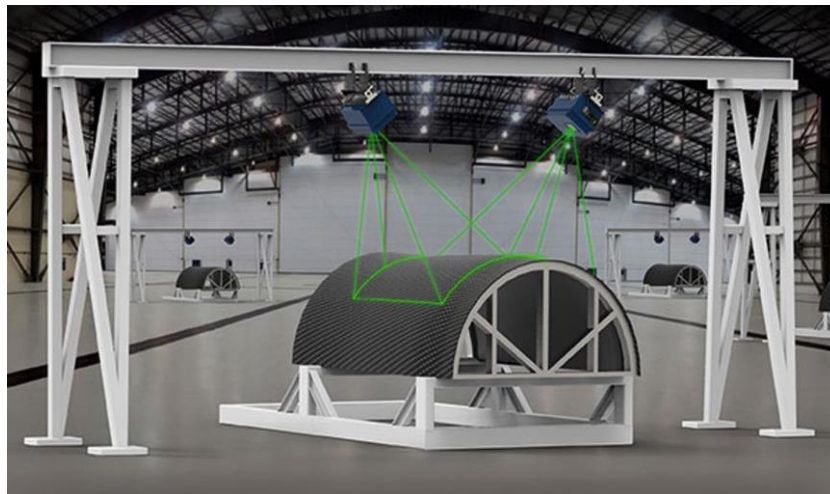


Abbildung 20 | Laserprojektoren für 2D-Informationen (FARO Europe GmbH & Co. KG, 2020b)

Die Interaktion mit dem System, zum Start der Projektion der Konturen, erfolgt über zentrale PC-Terminals oder Fernbedienungscontroller. Anbieter dieser Lasersystemen sind u. a. die SL-Laser GmbH, LAP GmbH und Faro GmbH (FARO Europe GmbH & Co. KG, 2020a; LAP GmbH Laser Applikationen, 2020; SL-Laser GmbH, 2020b).

Für die Anzeige der 2D-Konturen auf verschiedenen Bauteilgeometrien werden die CAD-Daten mittels einer speziellen Software aufbereitet (z. B. mit „Siemens Fibersim“), welche die 3D-Daten in eine 2D-Geometriekoordinaten umwandeln. Nach Referenzierung des Montageobjektes mittels Laser-Targets werden die Konturen anschließend projiziert. Das System hat den Vorteil, dass die Lasertechnologie eine hochpräzise Darstellung der 2D-Konturen ermöglicht, was für den speziellen Einsatz im Flugzeugbau von Bedeutung ist.

Der große Nachteil besteht jedoch darin, dass erweiterte Arbeitsinformationen mit mehrfarbigen Bildern, Videos, Animationen oder sonstige interaktive Darstellungen damit nur unzureichend oder gar nicht dargestellt werden können.

Weitere Entwicklungen im Bereich von statischen AR-Projektionssystemen im Großgerätebereich sind bei ŠKODA AUTO Deutschland GmbH anzutreffen (Abbildung 21), bei dem Mitarbeiter*innen bei der Beladung von Paletten mittels projektorbasierter Information unterstützt werden. Der Vorteil hier ist wieder, dass der Mensch keine AR-Brille, HMD oder ein anderes tragbares Wearable benutzen muss (Schäfer, 2019; ŠKODA AUTO Deutschland GmbH, 2020). Sie unterscheiden sich von den statischen Lasersystemen dadurch, dass hier Bilder, Videos sowie mehrfarbige Informationen angezeigt werden können.

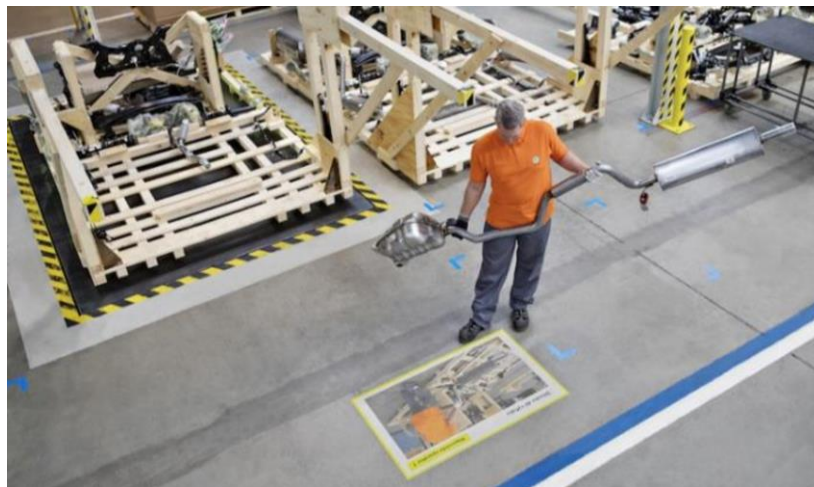


Abbildung 21 | Einsatz von SAR bei ŠKODA AUTO (ŠKODA AUTO Deutschland GmbH, 2020)

Wie die Interaktion als zentraler Teil von Augmented Reality sowie die detaillierte Gestaltung dieses Assistenzsystems umgesetzt wurde, konnte hier nicht ermittelt werden.

In der Kleingeräte- und Einzelplatzmontage wird Spatial Augmented Reality mit statischen Projektoren (Abbildung 22) bereits eingesetzt, um gezielt Arbeitsinformationen direkt auf der Arbeitsfläche anzuzeigen und mittels RGB-Tiefenkameras Interaktionen zu ermöglichen (Bertram et al., 2018; Büttner et al., 2016; Funk, 2016; Sand et al., 2016).

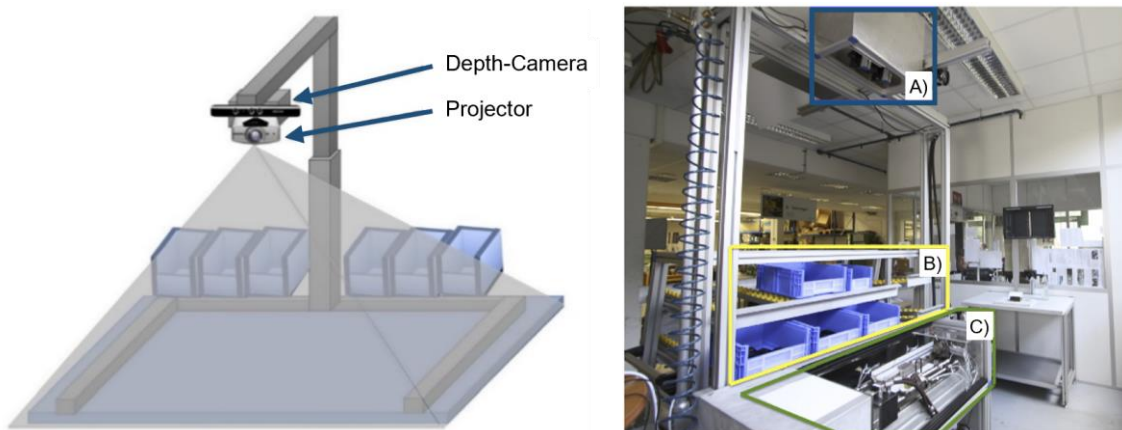


Abbildung 22 | Statische Projektion in der Kleingeräte montage, in Anlehnung an Funk (2016)

Der graue Kegel in Abbildung 22 zeigt den möglichen Projektionsbereich, welcher sich über einen Arbeitsbereich einer Einzelplatzmontage (ca. 1,5 x 1,5 Meter) erstreckt. In Funk (2016) wird ein Projektor der Marke „ASUS K330“ sowie die „Microsoft Kinect v2“ als Tiefenkamera verwendet, welche in 96 Zentimetern Höhe über dem Arbeitsbereich angebracht wurde.

Für die Großgeräte- und Baustellenmontage eignet sich das Konzept der projektorbasierten Spatial Augmented Reality Assistenzsysteme nach Funk (2016) jedoch nur mangelhaft, da der Projektionsbereich sowie die Interaktionsmöglichkeit in dieser Konstellation ausschließlich auf einen kleinen Arbeitsbereich beschränkt sind.

Das Spatial Augmented Reality Assistenzsystem muss dementsprechend für den Einsatz in der Baustellenmontage adaptiert und erweitert werden, um auf großen Arbeitsbereichen effektiv Informationen zu projizieren und eine Interaktion zu ermöglichen.

Dynamisches Spatial Augmented Reality Assistenzsystem

Die Größe der Fläche, auf die projiziert werden kann, ist einerseits vom Abstand des Projektors zur Projektionsoberfläche und andererseits von der Möglichkeit, das projizierte Bild im Raum dynamisch zu bewegen, abhängig. Statische Projektoren, welche über einen Montagearbeitsplatz in ca. 2 bis 3 Metern Entfernung montiert werden, können in einem Bereich von ca. 1,5 x 1,5 Meter (vgl. Abbildung 22) projizieren und sind deswegen für die Großgeräte- und Baustellenmontage nur bedingt anwendbar. Deshalb müssen geeignete Konzepte für die Ausleuchtung von großen Bereichen gefunden oder Möglichkeiten geschaffen werden, wie statische Projektoren dementsprechend adaptiert werden können.

Bei dem Konzept der „Immersive CAVE“ werden mehrere Projektoren verwendet, um auf große Flächen bzw. im 3D-Raum Projektionsinhalte anzuzeigen. Diese Systeme sind u. a. in Unternehmen für Präsentation zwecke sowie in Forschungseinrichtungen wie z. B. bei Fraunhofer IAO anzutreffen (Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, 2020). Dabei werden mehrere Projektoren in horizontaler sowie vertikaler Ausrichtung zusammengeschaltet und die zu projizierenden Inhalte auf die jeweiligen Projektoren synchronisiert. Für die immersive 3D-Projektion eines kompletten Raumes sind bis zu 20 Projektoren sowie geeignete Rechner mit hoher Rechenleistung für die Aufbereitung der Bild- und Videodaten erforderlich (Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO, 2020). Die Projektionen werden jedoch auf „Leinwänden“ angezeigt und eine direkte Projektion z. B. auf ein in der Cave befindliches Bauteil ist hier nicht möglich.

Dynamische Ablenkung des Projektionsstrahls durch ein Spiegelsystem

Ein weiteres Konzept stellt die dynamische Ablenkung des Projektionsstrahls dar, bei dem das erzeugte Bild des Projektors über einen oder mehrere Spiegel in den Raum projiziert wird (vgl. Kapitel 1.2.2).

2.3.2 Informationsbereitstellung mit Spatial Augmented Reality

Intuitiv gestaltete Arbeitsanweisungen ermöglichen eine verbesserte kognitive Verarbeitung von Informationen und reduzieren die mentale Beanspruchung des Menschen am Arbeitsplatz (Schlick et al., 2018). Der Einsatz von Spatial Augmented Reality für die Anzeige von Arbeitsanweisungen in der Montage eignet sich besonders, um Montagefehler zu vermeiden, kognitive Belastungen der Menschen zu reduzieren und einen effizienten Montageprozess zu gewährleisten (Funk, Kosch et al., 2015; Funk, Kosch & Schmidt, 2016; Rupprecht et al., 2020; Tang et al., 2003). Demnach ist für die Vermeidung von Fehlern sowie für die effiziente Ausführung der Montageaufgaben eine **strukturierte Aufbereitung** und eine **intuitive verständliche Visualisierung** der Arbeitsanweisungen notwendig. Das Spatial Augmented Reality Assistenzsystem ermöglicht dabei die gezielte Anzeige der Informationen am jeweiligen Ort des Geschehens in der Baustellenmontage. Dementsprechend ist es entscheidend, eine effektive Bereitstellung und **Positionierung der Information** im Arbeitsraum und am Montageobjekt zu gestalten. Ebenso können herkömmliche Arbeitsanweisungen nicht 1 zu 1 übernommen, sondern sie müssen für die Gegebenheiten der Projektion aufbereitet werden. Die **projektionsoptimierte Aufbereitung der Arbeitsanweisung** stellt einen weiteren Gestaltungsaspekt dar, um eine hochwertige Anzeige mithilfe des Spatial Augmented Reality Assistenzsystems zu gewährleisten (Rupprecht et al., 2020). Das dynamische Spatial Augmented Reality Assistenzsystem bietet dazu die flexible Möglichkeit der Integration und Anzeige von unterschiedlichen und mehrfarbigen Visualisierungsarten (Texte, Piktogramme, Bilder, Animation und Video etc.) und ermöglicht die Bearbeitung, Skalierung und Ver-/Entzerrung der Informationen nach den jeweiligen Projektionspositionen und Oberflächenformen (Rupprecht et al., 2020). Des Weiteren bietet das dynamische Spiegelsystem die Möglichkeit der direkten Projektion am Ort des Geschehens sowie der gezielten Führung durch den Montageprozess (**Informationsführung**), welche mit der dynamischen Anzeige der Arbeitsinformationen an der jeweilig definierten (nächsten) Montageposition umgesetzt werden kann (Rupprecht et al., 2021).

Intuitive Informationsvisualisierung von Arbeitsanweisungen

Die aktuell verwendeten Arbeitsinformationen, welche in Papierdokumenten sowie in PDF-Dokumenten an PC-Terminals zur Verfügung stehen, sind meist als Fließtext strukturiert und beinhalten kaum Abbildungen (Rupprecht et al., 2020). Diese Art der Arbeitsanweisungen sind nicht intuitiv und projektionsoptimiert aufbereitet, was die Verständlichkeit erschwert. Ebenso können Fließtexte aufgrund mangelnder Lesbarkeit, Skalierung und zu hohem Detailgrad nicht ohne Anpassungen für die Anzeige mittels Projektionssystem übernommen werden. Um Arbeitsanweisungen dementsprechend aufzubereiten und darzustellen sowie eine Benutzerfreundlichkeit und optimale Verständlichkeit zu gewährleisten, sind Empfehlungen und Gestaltungsrichtlinien zur Informationsdarstellung anzuwenden und eine standardisierte Strukturierung der Informationen essenziell.

Die Norm DIN EN ISO 9241-112 beschreibt die Grundsätze der Informationsgestaltung und vermittelt, dass Informationen so dargestellt werden sollen, dass sich eine optimale

Gebrauchstauglichkeit einstellt. Sie sollen den Menschen effektiv und effizient bei den Tätigkeiten unterstützen sowie zur Zufriedenheit bei der Umsetzung der Arbeitsinhalte beitragen. Die Norm empfiehlt für die Darstellung von Informationen folgende sechs Gestaltungsprinzipien:

- **Entdeckbarkeit:** Informationen sollen schnell erkennbar sein.
- **Ablenkungsfreiheit:** Informationen sollen nicht von anderen Tätigkeiten bzw. von Gefahrenstellen ablenken.
- **Unterscheidbarkeit:** Informationen sollen klar voneinander unterscheidbar sein.
- **Eindeutige Interpretierbarkeit:** Es soll klar und deutlich beschrieben werden, was die Information aussagt.
- **Kompaktheit:** Die Information soll auf das Wesentliche konzentrieren.
- **Konsistenz (interne und externe):** Informationen sollen innerhalb des Arbeitskontextes sowie im Unternehmen gleich gestaltet werden.

Die Grundsätze dienen zur optimalen Gestaltung der Gebrauchstauglichkeit, sind grundsätzlich voneinander unabhängig und können je nach gewählter Priorität angepasst werden. Demnach soll bei der Gestaltung priorisiert werden, welche Parameter relevanter sind, je nachdem welche Technologie verwendet wird, welcher Arbeitskontext vorherrscht und für welche Benutzergruppe (z. B. unerfahrene oder erfahrene Werker*innen) die Informationen gestaltet werden (DIN EN ISO 9241-112). Die Norm DIN EN ISO 9241-125 beinhaltet spezifische Empfehlungen zur Darstellung visueller Informationen. Folgend werden einige Aspekte zur Strukturierung und Gruppierung von Informationen, welche für die Aufbereitung der Informationen zur Anzeige mittels Projektionssystem sowie zur Gestaltung der Benutzerschnittstelle relevant sind, präsentiert:

- Alle Informationen, die zur Ausführung der Aufgabe notwendig sind, müssen dargestellt werden.
- Visuelle Benutzerschnittstellen sollen klar erkennbar sein und jeweils an derselben Stelle der Arbeitsanweisung platziert werden.
- Farben und Symbole sollen gezielt eingesetzt werden und internationalen wie auch kulturellen Normen entsprechen sowie eine Konsistenz innerhalb des Prozesses aufweisen.
- Informationen sollten so platziert werden, dass sie den Benutzererwartungen entsprechen, schnell gefunden werden und die Aufgabenanforderungen erfüllen.
- Die Dichte der Information soll so gewählt werden, dass die Informationen nicht überladen wirken und sich negativ auf das Verständnis auswirken.

Für die Gestaltung und Aufbereitung der Arbeitsanweisungen speziell für den Einsatz von Spatial Augmented Reality in der Montage werden folgende Gestaltungsrichtlinien beschrieben (Funk, 2016; Funk, Kosch et al., 2015):

- **Strukturierter Aufbau und prägnante Formulierung:** Die einzelnen Montagetaasks müssen strukturiert aufgebaut und verständlich sein und wenn möglich aus (unternehmensspezifischen) Standardformulierungen bestehen.

- **Intuitiv verständlich und wiederholbar:** Die Arbeitsaufgaben müssen intuitiv verständlich beschrieben werden, damit sie jederzeit (von anderen Personen) wiederholt werden können.
- **Repräsentativ für die Montage:** Die Aufgaben müssen die drei Haupttätigkeiten beinhalten, welche für die Montage entscheidend sind:
 - Entnahme/Aufnahme von Bauteilen
 - Platzierung der Bauteile
 - Zusammenbauen/Montage der Bauteile
- **Skalierbar und veränderbar:** Die Anzahl der einzelnen Arbeitsschritte muss skalierbar und änderbar sein, ohne dass die Grundaufgabe dabei entscheidend verändert wird.

Für den strukturierten Aufbau und die prägnante Formulierung können kurze Texte sowie der gezielte Einsatz von Symbolik (ISO/IEC 11581-10, 2010), wie z. B. Piktogramme, Konturen oder Kurzvideos (Funk, Kosch & Schmidt, 2016), verwendet werden. Dabei wird unter einem Piktogramm ein auf das Wesentliche reduziertes Bildzeichen verstanden, das intuitiv erkennbar und einprägsam sowie ohne Erklärung über die Sprach- und Kulturgenzen hinweg verständlich ist. Damit wird visuell ein Sachverhalt beschrieben sowie gezielt auf Empfehlungen, Aktionen oder Informationen hingewiesen (Bühler et al., 2017; Moser, 2012; Stapelkamp, 2007). Funk, Kosch und Schmidt (2016) haben speziell im Anwendungsfall der Einzelplatzmontage und unter Verwendung eines statischen Spatial Augmented Reality Systems verschiedene Visualisierungsarten verglichen und dabei die „Task Completion Time“ (TCT), die Fehlerrate sowie die kognitive Arbeitsbelastung evaluiert (Abbildung 23).

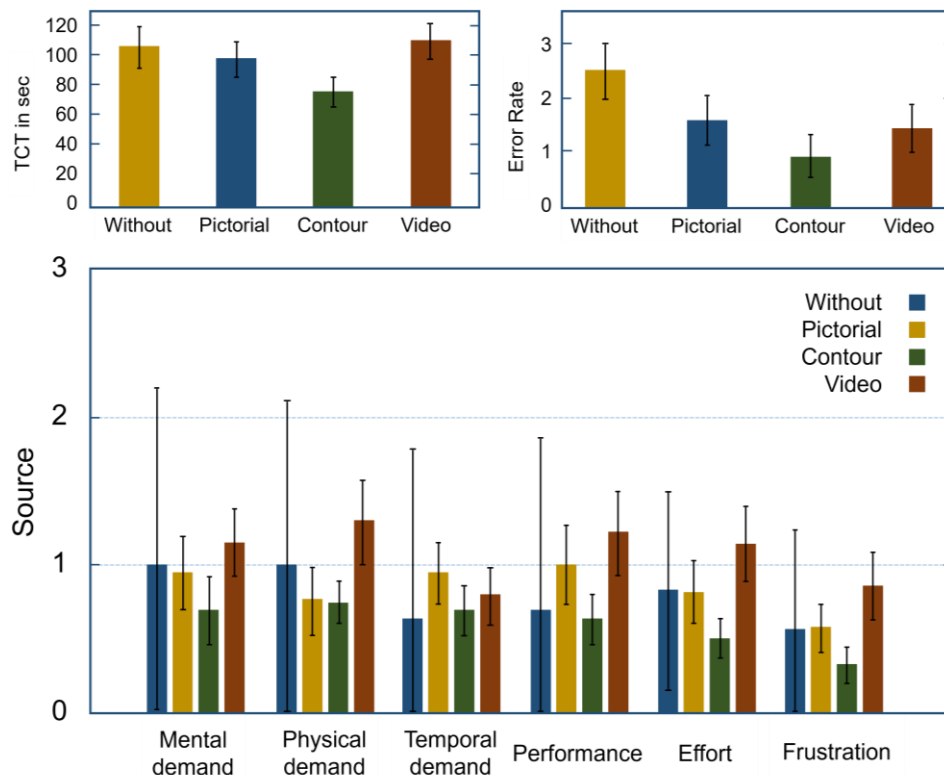


Abbildung 23 | Vergleich von Visualisierungsvarianten, in Anlehnung an Funk, Kosch und Schmidt (2016)

Dabei konnte festgestellt werden, dass die Anzeige von Bildern, Piktogrammen und Konturen im Vergleich zu Kurzvideos einen erheblichen Mehrwert bietet. Piktogramme sowie Konturen schneiden bei der „Task Completion Time“ sowie bei der „Error Rate“ besser ab, was sich positiv auf die Prozesseffizienz auswirkt. Die Resultate des NASA-TLX-Tests zeigen den Mehrwert der kognitiven Arbeitsbelastung und somit für das Wohlbefinden des Menschen (Funk, Kosch & Schmidt, 2016).

Resultierend aus den Erkenntnissen wird in dieser Forschungsarbeit bei der Gestaltung des Spatial Augmented Reality Assistenzsystems sowie der Informationsgestaltung der Einsatz von Piktogrammen, Bildern sowie kurzen textuellen Anweisungen forciert.

Projektionsoptimierte Arbeitsanweisungen

Wie schon bei der Gestaltung von Piktogrammen erwähnt, ist es speziell für die projektionsgerechte Aufbereitung essenziell, die Arbeitsinformationen nach den Gegebenheiten des Projektionssystems sowie der jeweilig vorherrschenden Oberfläche, auf die projiziert werden soll, aufzubereiten. Dabei sind die Art der Projektionsoberfläche (planar, gekrümmt etc.), das Material (Rauheit, Texturen etc.) und die Farbe der Oberfläche sowie die Spezifikationen des Projektionssystems (Auflösung, Kontrast etc.) und die vorherrschende Arbeitsumgebung (Tageslicht etc.) für die Gestaltung der Arbeitsanweisungen entscheidend. Vor allem ist die Farbwahl von Piktogrammen oder textlichen Formulierungen speziell bei der Verwendung von Projektoren essenziell, um hochqualitative Informationen anzuzeigen und die ergonomischen Aspekte der Lesbarkeit zu berücksichtigen. Die Wahl der Farbe ist inhärent mit dem Kontrast und der subjektiven Wahrnehmung der Helligkeit verbunden, dabei beschreibt der Simultankontrast die Wahrnehmung von eingesetzter Farbe (z. B. der Arbeitsinformation) relativ zur Umgebungsfarbe bzw. Farbe der Projektionsoberfläche (Abbildung 24) (Stapelkamp, 2007).



Abbildung 24 | Wahrnehmung von grüner Farbe bei unterschiedlicher Umgebungsfarbe (Simultankontrast), in Anlehnung an Stapelkamp (2007)

Demnach eignet sich die Wahl von Farben, welche in einem definierten Farbkreis gegenüberliegen bzw. sich voneinander abheben, zur Gestaltung des Komplementärkontrasts und zur Optimierung der Lesbarkeit der Informationen. Der Farbkreis von Harald Küppers (vgl. Stapelkamp, 2007) eignet sich besonders für die ästhetische und ergonomische Gestaltung von Informationen. Ebenso ist der Hell-Dunkel-Kontrast bei der Gestaltung zu berücksichtigen (Abbildung 25).



Abbildung 25 | Obere Reihe: Komplementärkontrast, untere Reihe: Hell-Dunkel-Kontrast, in Anlehnung an Stapelkamp (2007)

Für die Gestaltung der Arbeitsinformationen und -anweisungen für das Spatial Augmented Reality Assistenzsystem werden folgende Anforderungen zusammengefasst:

- strukturierter Aufbau der Arbeitsinformation sowie prägnante Formulierung
- kurze prägnante textliche Informationen und keine langen Fließtexte
- Integration der Haupttätigkeiten für die Montage
- Anwendung der geltenden Normen zu Grundsätzen der Informationsdarstellung sowie der Empfehlungen zur visuellen Informationsdarstellung
- Verwendung von Bildern und Piktogrammen
- Analyse der Arbeitsumgebung, des Projektionssystems sowie der Projektionsfläche zur optimalen Gestaltung der Informationen (Farbe, Kontrast, Größe etc.)

2.3.3 Dynamische In-situ-Projektion und Informationsführung

Durch die Möglichkeit des steuerbaren Spiegelkopfes und der resultierenden dynamischen Projektion über einem großen Arbeitsbereich sowie die Gegebenheit, Informationen an unterschiedlichen Positionen in der Montage anzuzeigen, ist es entscheidend, festzustellen, an welcher Position und in welcher Orientierung (in situ) die Arbeitsinformationen projiziert werden sollen. Außerdem ist zu ermitteln, wie die projizierten Informationen den Menschen durch den Montageprozess führen und wie die Interaktion mit dem Informationssystem umgesetzt werden soll. Demnach sind bei der Verwendung von Spatial Augmented Reality in der Großgeräte- und Baustellenmontage die **Informationspositionierung mit In-situ-Projektion** sowie die **Informationsführung** zwei weitere essenzielle Gestaltungsaspekte – neben der intuitiven und projektionsoptimierten Informationsvisualisierung. Dazu werden folgend zwei Designvarianten für die Informationsführung vorgestellt, welche bereits in der Einzelplatzmontage angewandt und evaluiert wurden (Blattgerste et al., 2017; Funk, Kosch & Schmidt, 2016).

Die erste Variante stellt die **statische In-view-Projektion** dar, bei der Arbeitsinformationen in das Sichtfeld des Menschen projiziert werden. Dies kann einerseits mit Head-Mounted Displays, Smartphones oder mit Projektoren geschehen (Blattgerste et al., 2017; Rupprecht et al., 2021). Beim Einsatz von Projektoren werden die Informationen z. B. statisch in der Mitte der Montagestation angezeigt (Rupprecht et al., 2021). Die In-view-Anweisungen zeigen hierzu mittels textueller oder grafischer Informationen die jeweiligen Montageaufgaben an und beschreiben, welche Bauteile zu verbauen bzw. zu entnehmen sind und wo sich die

Montagepositionen befinden. Dementsprechend führt die statische Informationsdarstellung den Menschen nicht aktiv durch Bewegung der Projektion von Montageposition zu Montageposition. Durch Interaktion, z. B. durch Betätigen eines Buttons oder andere Interaktionsmöglichkeiten, wird die Information an der gleichen Position weitergeschaltet (Blattgerste et al., 2017; Rupprecht et al., 2021).

Funk, Kosch und Schmidt (2016) haben dazu analog zu den Papieranweisungen die In-view-Anweisungen als zweidimensionale Ansicht in das Blickfeld des Werkes, mittels eines HMD, projiziert und evaluiert, ob Vorteile gegenüber den Papieranweisungen bestehen. Ebenso wurde analog, wie beim Umblättern der Papieranweisung, auch beim HMD das Weiterschalten dem Nutzer mittels Knopfdruck selbst überlassen.

Daraus resultiert, dass sowohl bei den Papieranweisungen als auch den In-view-Anweisungen über HMD ähnliche Ergebnisse bei der TCT erzielt wurden (Abbildung 26) (Blattgerste et al., 2017).

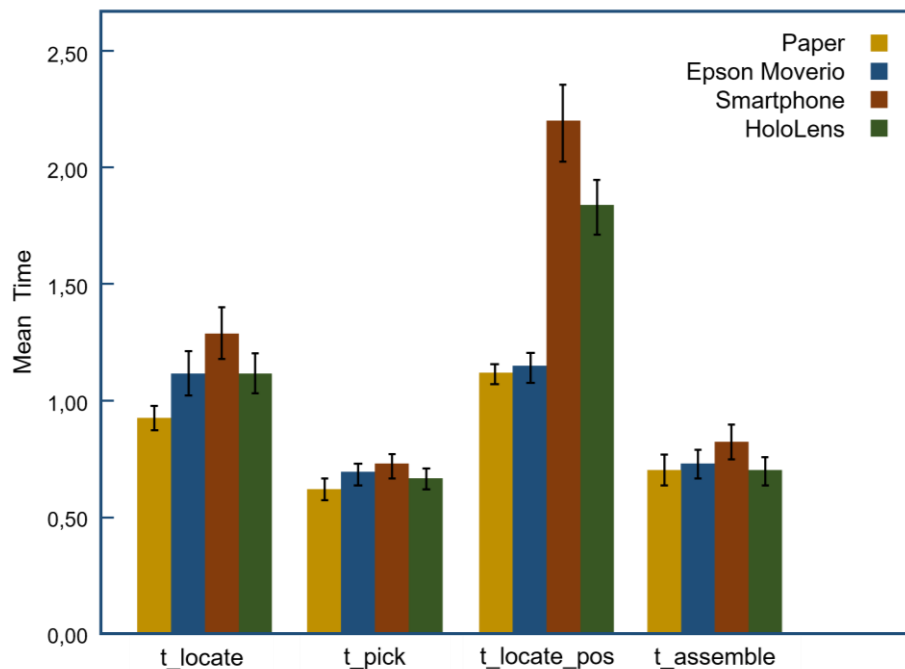


Abbildung 26 | TCT-Vergleich von verschiedenen Systemen, in Anlehnung an Blattgerste et al. (2017)

Ein weiterer relevanter Gestaltungsaspekt in Bezug zur Platzierung der Arbeitsinformationen ist, dass diese so projiziert werden, dass sie den Menschen bei dessen Montagetätigkeiten nicht stören. Demnach soll die Anweisung peripher neben der exakten Montageposition projiziert werden (Blattgerste et al., 2017; Khuong et al., 2014). Ebenso konnte in der Studie festgestellt werden, dass das Tragen eines HMD für den Einsatz über eine ganze Schicht (8 Stunden) nur bedingt geeignet ist, da es zu Kopfweh führen kann (Blattgerste et al., 2017). Demnach kann zusammengefasst werden, dass In-view-Anweisungen bereits mittels HMD getestet wurden und ähnliche Ergebnisse wie bei Papieranweisungen in Bezug auf die Montagzeit liefern und dass auf die Platzierung von Arbeitsanweisungen geachtet werden soll, sodass sie im Montageprozess nicht stören.

Die zweite Variante der Informationsführung stellt die **dynamische In-situ-Projektion** dar, bei der die Arbeitsinformationen direkt **an der richtigen Montageposition und in der richtigen Lage und Orientierung (in situ)** im Arbeitsbereich angezeigt werden und der Mensch mit dem dynamischen Projektionsstrahl zur Montageposition hingeführt wird. Die exakten Montagepositionen werden dabei visuell mit Licht, Konturen oder dynamischen Animationen dargestellt. Demnach wird der Mensch bei dieser Variante Schritt für Schritt durch den Montageprozess geführt (Büttner et al., 2017; Funk, Kosch et al., 2015; Funk, Kosch & Schmidt, 2016; Rupprecht et al., 2021).

Funk, Kosch und Schmidt (2016) haben die Informationsanzeige mit Papier-, Tablet- und In-view-Anweisungen mittels HMD im Vergleich zur In-situ-Projektion, welche mithilfe eines Projektors umgesetzt wurde, evaluiert (Abbildung 27).

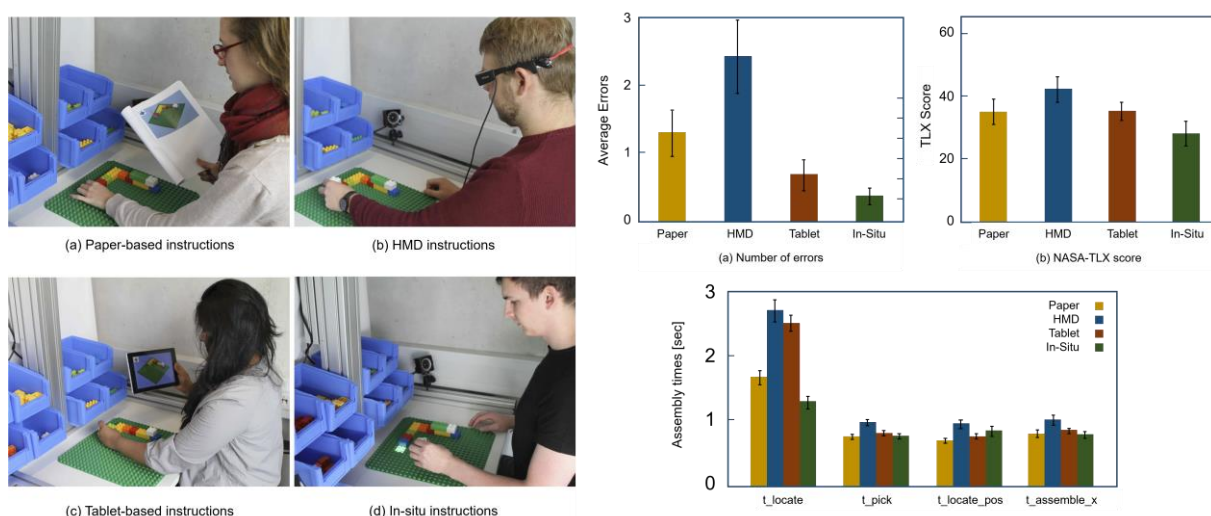


Abbildung 27 | Vergleich und Evaluation der Informationsanzeige mit In-situ-Projektion, Tablet, HMD und Papierarbeitsanweisung, in Anlehnung an Funk (2016) und Funk, Kosch und Schmidt (2016)

Aus dem Vergleich ist ersichtlich, dass die In-situ-Projektion von Arbeitsanweisungen mit statischen Spatial Augmented Reality Assistenzsystemen bei der Fehlerrate, beim NASA-TLX sowie der „Task Completion Time“ erheblich besser abschneidet als die Vergleichssysteme. Durch die Evaluierung in der Einzelplatzmontage kann ermittelt werden, dass In-situ-Anweisungen für die Darstellung von Montageaufgaben hinsichtlich der Montagezeit und der kognitiven Arbeitsbelastung des Benutzers hohes Potenzial zeigen. Der Einsatz von Spatial Augmented Reality mit In-situ-Projektion bietet in der Montage einen erheblichen Mehrwert im Vergleich zu tragbaren Geräten und Papieranweisungen sowie zur Darstellung mit einer In-view-Projektion (Blattgerste et al., 2017; Funk et al., 2017; Funk, Bächler et al., 2015; Funk, Kosch & Schmidt, 2016; Rupprecht et al., 2021). Die In-situ-Projektion wurde auch in einer Langzeitstudie evaluiert, mit dem Ergebnis, dass das Assistenzsystem speziell für neue Mitarbeiter*innen eine gute Unterstützung bietet und ein entspanntes Arbeiten ermöglicht, die Montagezeit jedoch bei trainierten Menschen deutlich höher ist (Funk et al., 2017).

Zusammengefasst kann erwähnt werden, dass Arbeitsanweisungen kompakt und intuitiv verständlich sowie standardisiert gestaltet werden sollen und für die Anforderungen der Projektionssysteme und der Gegebenheiten der Montage optimiert aufbereitet werden müssen. Für die Art der Visualisierung sollen selbsterklärende Piktogramme bzw. Bilder oder Animationen in Kombination mit kurzen Texten verwendet werden. Aus den Forschungen im Bereich der Einzelplatzmontage konnte festgestellt werden, dass die In-situ-Projektion von Arbeitsanweisungen erheblichen Mehrwert im Vergleich zu anderen Informationsdarstellungsarten bringt. Demnach soll durch den Einsatz des dynamischen Projektionssystem die In-situ-Projektion auch in der Großgeräte- und Baustellenmontage angewandt werden, um dabei den Menschen gezielt durch den Montageprozess zu führen und anzuzeigen, wo das nächste Teil aufzunehmen ist und wo es montiert werden soll.

2.3.4 Echtzeitfähige Interaktion mit Computer Vision und Deep Learning

Ausgehend von den Anforderungen, dass ein Spatial Augmented Reality Assistenzsystem nach der Definition von Azuma (1997) eine Interaktion in Echtzeit benötigt und der Mensch ein zentraler Handlungsakteur in der industriellen und manuellen Montage ist, bedarf es moderner Konzepte zur Erkennung des Menschen und zur Auslegung einer prozesseffizienten und natürlichen Mensch-Maschine-Interaktion. Demnach soll die Mensch-Maschine-Interaktion intuitiv und mit dem Einsatz von Körper, Gesten oder Posen umgesetzt und cyberphysische Systeme mit anthropomorphen Dimensionen ausgestattet und verbunden werden (Garcia et al., 2019; Gorecky et al., 2014). Dementsprechend stellt die visuelle Datenverarbeitung und insbesondere die visuelle Posen- und Gestenerkennung ein hohes Potenzial für die Interaktion in cyberphysischen Systemen dar (Posada et al., 2015). Durch „Computer Vision“ bzw. Bild- und Objektkennung sowie durch den Einsatz von „Machine & Deep Learning“ können mithilfe von Kamerasystemen die anthropometrischen Merkmale des Menschen erkannt und für die Interaktionen zur Steuerung des Systems eingesetzt werden (Geiselhart et al., 2016; Y. Wang et al., 2018). Computer Vision hat das Ziel, menschliches Sehen mit Kamerasystemen und geeigneten Algorithmen nachzuahmen, um dabei aus Bildern oder Videos nützliche Informationen zu extrahieren (Klette, 2014; Szeliski, 2011). Für die Erkennung des Menschen sowie die Bewegungserkennung werden vor allem RGB- und/oder Tiefenkameras unterschiedlicher Hersteller (z. B. Microsoft Kinect, Intel RealSense etc.) verwendet (Funk, 2016; Korn et al., 2015; Lun & Zhao, 2015; Rupprecht et al., 2021). Des Weiteren stellen lern- und progressionsbasierte Deep-Learning-Algorithmen wie z. B. „YOLO“ (Redmon et al., 2016) die Möglichkeit dar, eine echtzeitfähige Objekterkennung umzusetzen. Dhiman et al. (2019) verwenden YOLO als Interaktionssystem in Kombination mit Augmented Reality in der industriellen Montage, um manuelle Arbeitstätigkeiten zu analysieren und Montagefehler zu erkennen. Durch die preisgünstige Verfügbarkeit von hochauflösenden Kamerasystemen (RGB-, Stereo- oder 3D-Tiefenkameras) sowie die Verfügbarkeit von Hochleistungsrechnern und den Zugang zu Algorithmen der Bilderkennung wird ein effizienter Einsatz der Bild- und Objekterkennung im industriellen Umfeld heutzutage ermöglicht.

Objekterkennung und Machine Learning

Menschen haben bei der Betrachtung von Bildern die besondere Eigenschaft, Objekte schnell zu lokalisieren und deren Merkmale zu erkennen. Das Ziel der Objekterkennung ist, diese Intelligenz mithilfe von Computer Vision, Kamerasystemen und der Zuhilfenahme von geeigneten Lernalgorithmen maschinell umzusetzen (Joshi, 2020). Objekt- oder Bilderkennung ist ein Verfahren der visuellen Bildverarbeitung zum Lokalisieren und Erkennen verschiedener Eigenschaften und Instanzen eines Objektes in einem Pixel-Bild. Objekterkennungsalgorithmen nutzen typischerweise lernbasierte Methoden, welche auf Machine Learning bzw. Deep Learning basieren (MathWorks, Inc., 2020; Szeliski, 2011).

Machine Learning ist ein Teilgebiet der Künstlichen Intelligenz und wurde erstmalig von Arthur Samuel im Jahre 1959 beschrieben (Samuel, 1959). Machine Learning ist der Prozess, erfahrungsbasiertes Lernen in Informationssysteme und Computer zu integrieren. Dabei werden großen Mengen an (Trainings-)Daten (die zu erlernenden Bilder) den Machine-Learning-Algorithmen zur Verfügung gestellt, um daraus Informationen zu extrahieren und die erkannten Objekte zu „klassifizieren“. Der Prozess, den Inhalt eines Bildes zu erkennen, wird als Klassifikation bezeichnet. Durch das wiederholende Trainieren mit großen Datenmengen „lernt“ der Computer ähnlich wie der Mensch und speichert das Erlernte (die erkannten Objekte bzw. Klassen) als „Erfahrung“ ab, um später, bei Input von neuen Daten darauf zurückzugreifen (Joshi, 2020; MathWorks, Inc., 2020). Die effiziente Erkennung wird durch die Qualität und Quantität der Lerndaten, durch die Algorithmen, welche die Abweichung zwischen den bereits erkannten und neuen Daten ermitteln, und der Funktion, welche danach die Anpassung des Lernprozesses vornimmt, bestimmt (Ayodele, 2010; Joshi, 2020). Beim maschinellen Lernen wird zwischen zwei relevanten Arten zur Vorgehensweise der Klassifizierung von Objekten unterschieden (Ernst et al., 2020; Joshi, 2020; MathWorks, Inc., 2020) (Abbildung 28).

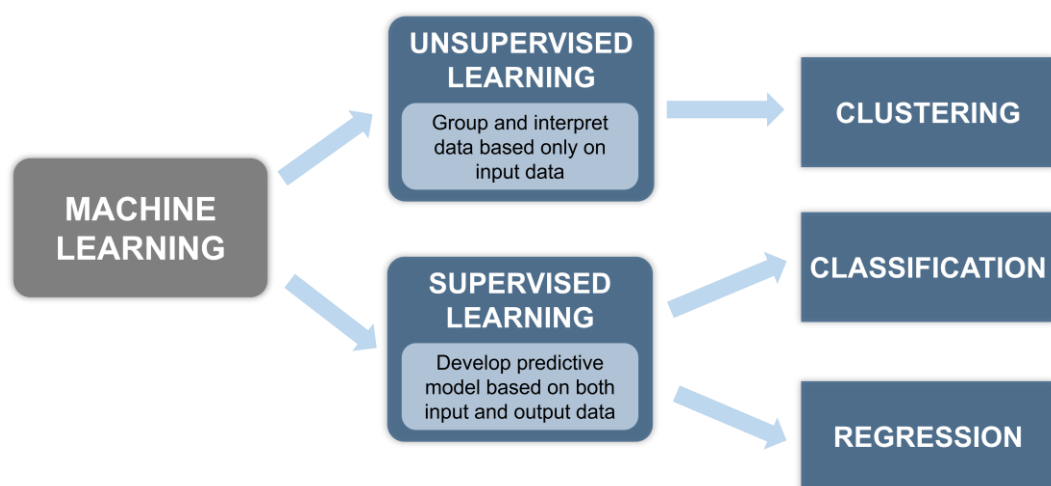


Abbildung 28 | Maschinelles Lernen – Vorgehensweise, in Anlehnung an MathWorks, Inc. (2020)

„**Supervised Learning**“ ist die am meisten eingesetzte Variante und verwendet vom Menschen manuell gelabelte und vorklassifizierte Daten als Trainingsinput (z. B. Label 1: Klasse Hund, Label 2: Klasse Katze). Durch eine hohe Anzahl an Trainingsdaten können zukünftige neue Eingabedaten automatisch vom Algorithmus klassifiziert werden. Das

manuelle Labeln stellt jedoch einen hohen Aufwand dar und kann nur bei bereits bekannten Klassen verwendet werden (Ernst et al., 2020). „**Unsupervised Learning**“ verwendet keine vorklassifizierten Trainingsdaten, sondern versucht anhand der vorliegenden Inputdaten Muster selbständig zu erkennen. Diese Art wird dann eingesetzt, wenn die Klassen vorab noch nicht bekannt sind (Ernst et al., 2020).

Objekterkennung mit Deep Learning und neuronalen Netzwerken

Eine effektive Methode zur visuellen Objekterkennung stellt den Einsatz von Deep Learning dar, um automatisiert Objekte aus bestehenden Bildern zu erkennen (Ernst et al., 2020; He et al., 2016; Voulodimos et al., 2018; Zhao et al., 2019). Deep Learning ist ein Teilbereich von Machine Learning, welches auf neuronale Netzwerke (Convolutional Neural Networks – CNN) basiert (Ernst et al., 2020). Die Methode ist inspiriert vom menschlichen Gehirn und dem Verhalten der einzelnen Neuronen in biologischen Netzwerken, welche durch Signalweitergabe komplexe Entscheidungen durchführen (O'Mahony et al., 2020). Der Begriff „Deep“ bezieht sich im Allgemeinen auf die Anzahl verborgener Schichten („Layer“) des neuronalen Netzes. Herkömmliche neuronale Netze enthalten rund zwei bis drei verborgene Schichten, während tiefe Netze bis zu 150 „Layers“ enthalten können (Ernst et al., 2020; He et al., 2016; MathWorks, Inc., 2020). Folgend wird der Aufbau eines klassischen neuronalen Netzwerks (Abbildung 29), bestehend aus „Input/Output Layer“ sowie den „Hidden Layers“ und den Knotenpunkten, gezeigt (MathWorks, Inc., 2020).

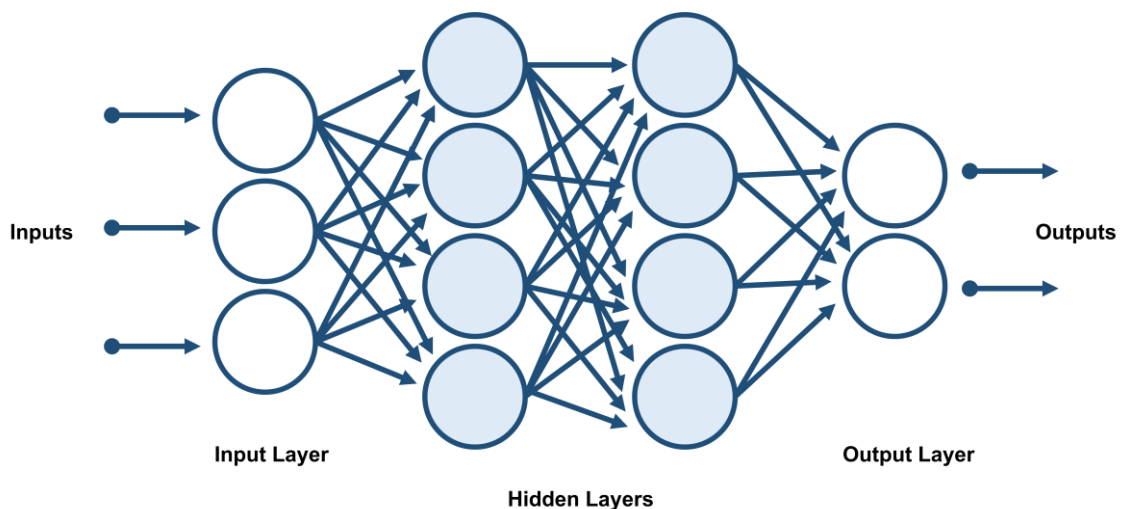


Abbildung 29 | Aufbau eines neuronalen Netzwerkes, in Anlehnung an MathWorks, Inc. (2020)

Deep Learning, speziell in der Objekterkennung von visuellen Daten, wird aufgrund der Verfügbarkeit von klassifizierten Datensätzen und Hochleistungs-Grafikprozessoren erst seit geraumer Zeit großflächig in Forschungen und industriellen Entwicklungen eingesetzt (Voulodimos et al., 2018; Zhao et al., 2019). Speziell für den Bereich des Deep Learning mit neuronalen Netzen bieten sich die Programmbibliotheken „PyTorch“ (<https://pytorch.org>) oder „TensorFlow“ (www.tensorflow.org) an. Deep-Learning-Modelle werden anhand umfangreicher Sätze klassifizierter Daten und neuronaler Netzarchitekturen trainiert, welche die gewünschten Merkmale direkt und automatisiert aus den Daten lernen, ohne dass ein

manueller Eingriff benötigt wird. Abbildung 30 stellt einen Überblick über die Funktionsweise dieses neuronalen Netzwerks dar (MathWorks, Inc., 2020).

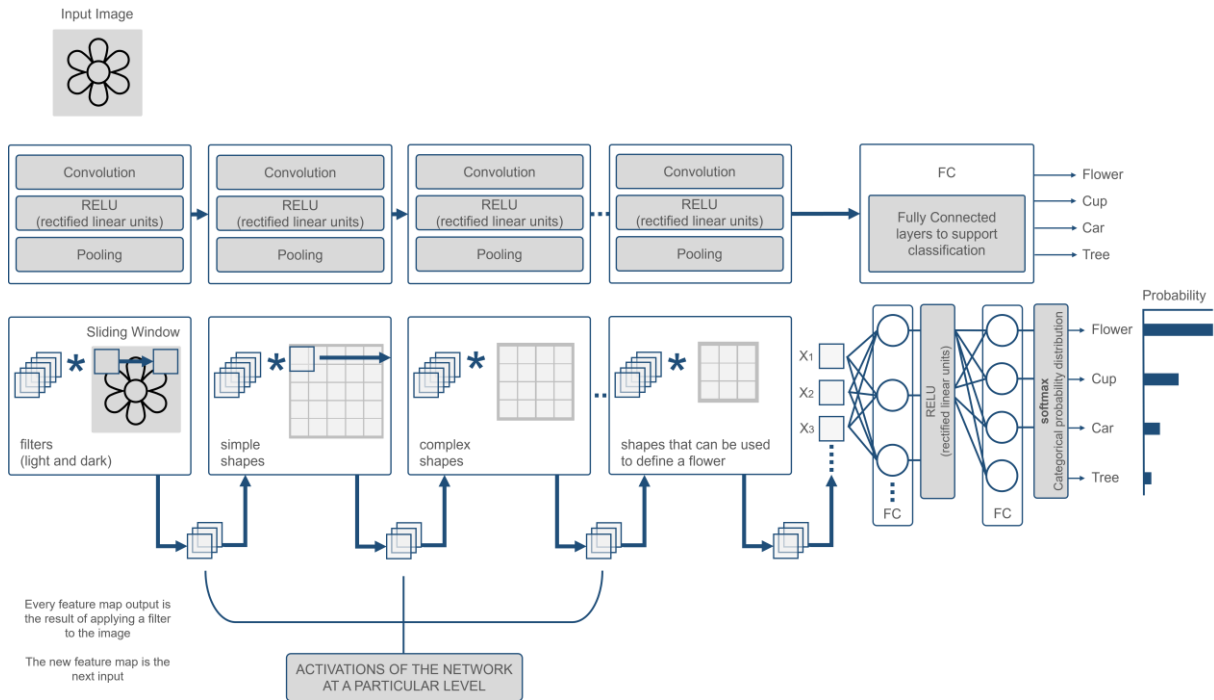


Abbildung 30 | Schematischer Ablauf in einem neuronalen Netzwerk, in Anlehnung an MathWorks, Inc. (2020)

Als Input dient ein Bild, das eine „Blume“ zeigt, die durch das neuronale Netzwerk geschleust wird. Als Output wird eine statistische signifikante Aussage getätigt, was auf diesem Bild zu „sehen“ ist. In diesem Falle wird die Gestalt einer „Blume“ mit der größten Wahrscheinlichkeit also Output angezeigt. Das Netzwerk gehört davor mit einer großen Menge an „Blumenbildern“ trainiert, um danach die richtigen Objekte erkennen zu können. Eine weitere Anwendung eines Deep-Learning-Netzwerks zeigt ein Projekt der Ryerson University von Toronto (www.cs.ryerson.ca/~aharley/vis/conv), bei dem mittels handschriftlicher Eingabe von Ziffern erkannt wird, welche Ziffer geschrieben wurde (Ryerson University, 2020).

Neuronale Netzwerke können selbst trainiert und die Netzwerkarchitektur selbst entworfen werden. Dabei müssen gelabelte Trainingsdaten ebenso zusammengestellt werden, wie die Eigenschaften der einzelnen Schichten manuell angepasst werden, was viel Zeit kostet. Heutzutage kann jedoch auf vortrainierte CNN zurückgegriffen werden, welche dann nach Bedarf nur noch feinabgestimmt werden müssen. Diese vortrainierten Netzwerke greifen bereits auf Tausende bis Millionen von Bildern zurück. Dhiman et al. (2019) analysieren für die Objekterkennung einige Algorithmen für tiefes Lernen wie z. B. AlexNet (Krizhevsky et al., 2017), ResNet (He et al., 2016) und YOLO (Joseph Redmon & Ali Farhadi, 2018; Redmon et al., 2016; Redmon & Farhadi, 2017). Dabei kristallisierte sich der YOLO-Algorithmus heraus, welcher im Vergleich zu den anderen die beste Performance (Genauigkeit/Geschwindigkeit) lieferte (Dhiman et al., 2019; Joseph Redmon & Ali Farhadi, 2018).

YOLO – You only look once

Für die Realisierung der menschenzentrierten Interaktion ist es essenziell, eine Technologie mit bewährter Zuverlässigkeit und Echtzeitverfügbarkeit einzusetzen, um sie im industriellen Umfeld verwenden zu können. Der Deep-Learning-Algorithmus YOLO, basierend auf dem neuronalen Netzwerk-Framework „DarkNet“ (<https://pjreddie.com/darknet>), eignet sich durch die Verwendung der benutzerfreundlichen Variante „AlexeyAB“ (<https://github.com/AlexeyAB/darknet>) besonders für eine unkomplizierte Implementierung in der Praxis. YOLO ist ein regressionsbasierter Algorithmus, der in einem Durchgang („you only look once“) durch ein einstufiges neuronales Netzwerk (Single-Stage-Network) die Objekte aus den visuellen Daten detektiert (Abbildung 31) (Joseph Redmon & Ali Farhadi, 2018; MathWorks, Inc., 2020; Redmon et al., 2016; Redmon & Farhadi, 2017).

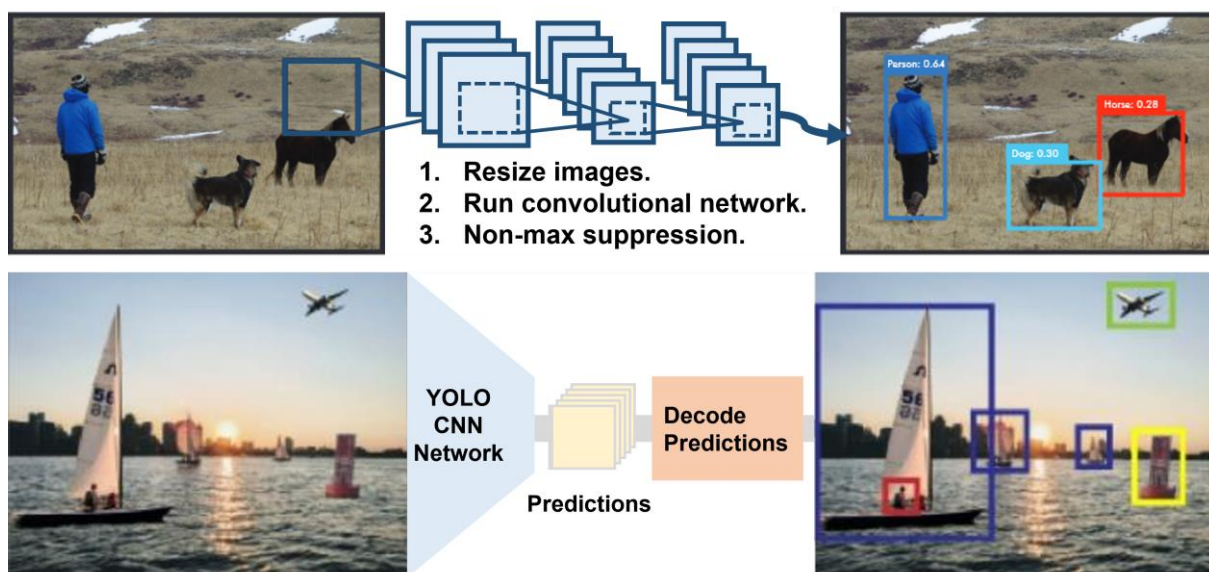


Abbildung 31 | Bilderkennungsprozess mit YOLO, in Anlehnung an MathWorks, Inc. (2020) und Redmon et al. (2016)

Dabei erstellt das CNN Netzwerkprognosen für Regionen über das gesamte Bild, dekodiert diese und klassifiziert sie mit sogenannten Bounding-Boxen, welche die Position und Größe der Objekte charakterisieren (Joseph Redmon & Ali Farhadi, 2018; MathWorks, Inc., 2020; Redmon et al., 2016; Redmon & Farhadi, 2017).

YOLO hat den Vorteil, im Vergleich zu anderen Algorithmen weniger Trainingsdaten zu benötigen, eine höhere Genauigkeit und Erkennungsgeschwindigkeit auszuweisen und über eine bedeutende Open-Source-Community zu verfügen. Der YOLO-Algorithmus wird stetig weiterentwickelt. Joseph Redmon und Ali Farhadi (2018) stellen die Version „YOLOv3“ mit signifikanten Verbesserungen vor. YOLOv3 leistet demnach eine genauere Objekterkennung und eine schnellere Bilderkennungszeit (Abbildung 32) als dessen Vorgänger (Joseph Redmon & Ali Farhadi, 2018).

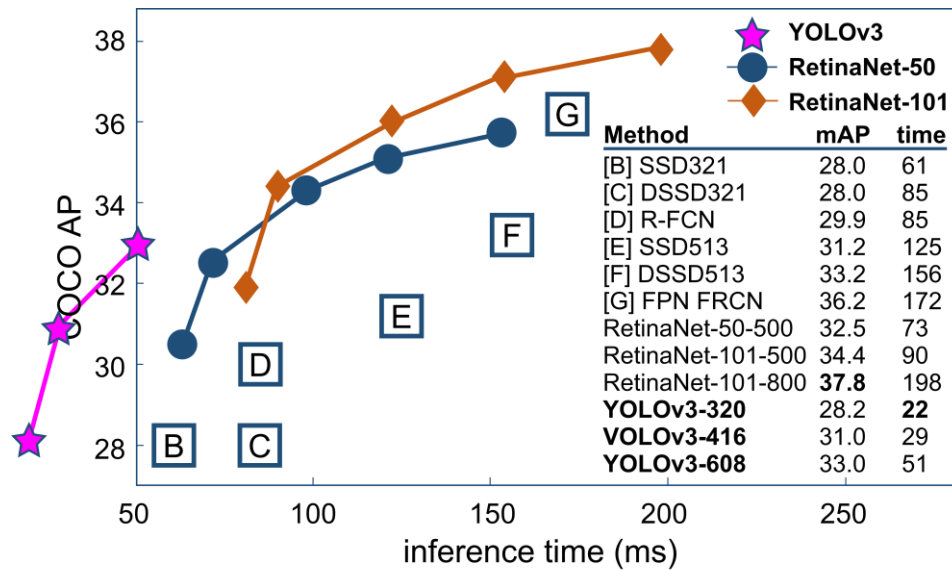


Abbildung 32 | Vergleich von YOLOv3 mit anderen Algorithmen, in Anlehnung an Joseph Redmon und Ali Farhadi (2018)

YOLOv3 erreicht im Vergleich zu anderen Netzwerken eine höheren „Mean Average Precision“ (mAP; Messwert für die Genauigkeit des Objektdetektors) sowie eine schnellere Objekterkennungszeit. Durch die Schnelligkeit und daraus resultierende Echtzeitfähigkeit eignet sich der Algorithmus besonders für Anwendung in der Industrie sowie bei der Verwendung mittels Augmented Reality (Joseph Redmon & Ali Farhadi, 2018). Eine weitere Version ist YOLOv4, welche von Bochkovskiy et al. (2020) mit den modernsten Objektdetektoren verglichen wird (Abbildung 33). YOLOv4 erreicht eine 12-prozentige Verbesserung zu dessen Vorgänger YOLOv3 (Bochkovskiy et al., 2020).

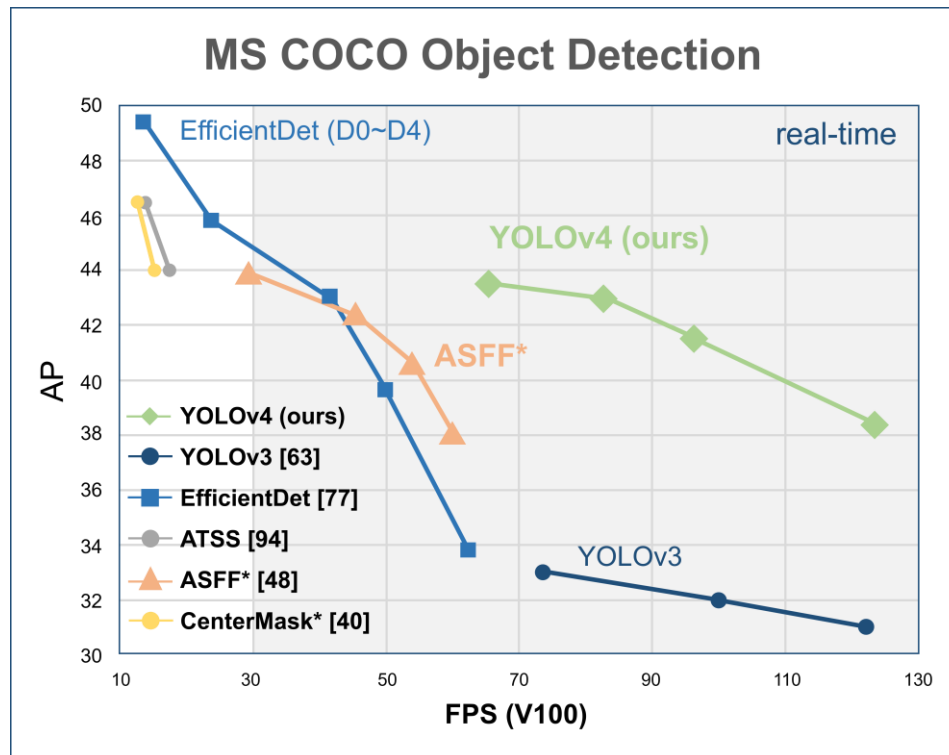


Abbildung 33 | YOLOv4-Algorithmus, in Anlehnung an Bochkovskiy et al. (2020)

In der vorliegenden Forschungsarbeit wird durch den Einsatz von hochauflösenden Kamerasystemen mit RGB-Streams sowie die Anwendung des Objekterkennungsalgorithmus YOLOv3 eine nutzeradaptive Interaktion zur Erkennung des Menschen, dessen Position und Gesten gestaltet. Diese Möglichkeiten der visuellen Interaktion stellen dabei einen essenziellen Gestaltungsaspekt zur Steuerung des dynamischen Projektionssystems dar.

2.4 Nutzeradaptive Mensch-Maschine-Interaktion

Unter **interaktives System** wird eine Kombination aus Hardware und/oder Software verstanden, mit der Benutzer zur Erreichung definierter Ziele interagieren (DIN EN ISO 9241-11:2018-11). Diese Interaktion wird über eine **Benutzungsschnittstelle** durchgeführt, welche Informationen und Steuerelemente zur Verfügung stellt, sodass der Benutzer die jeweilige Arbeitsaufgabe mit dem interaktiven System ausführen kann (DIN EN ISO 9241-110:2020). Demnach ist eine **Benutzer-System-Interaktion** ein Austausch von Informationen zwischen dem Benutzer und dem interaktiven System mittels Benutzerschnittstelle und zur Erfüllung der Arbeitsaufgabe (DIN EN ISO 9241-110:2020). Die ISO/IEC TR 25060:2010 für System und Software-Engineering beschreibt „**User interaction** as exchange of information between a user and an interactive system via the user interface to complete the intended task“.

Der Begriff „Mensch-Maschine-Interaktion“ (MMI) bzw. „Human-Computer Interaction“ (HCI) etablierte sich in den 1980er Jahren mit der ersten Konferenz 1983 in Boston (ACM SIGCHI Conference on Human Factors in Computing Systems) und einer der ersten Veröffentlichungen zum Thema „The Psychology of Human-Computer Interaction“ (Card et al., 1983) sowie den aufstrebenden Entwicklungen von Computern (u. a. Marktstart des „Apple Macintosh“ 1984) mit grafischen Benutzeroberflächen (Graphical User-Interfaces – GUI). Diese neuen Entwicklungen im Bereich Computer und Informationssysteme bedürfen einer verstärkter Forschung im Bereich Nutzergestaltung sowie Mensch-Maschine-Interaktion (MacKenzie, 2013).

Der Fachbereich Mensch-Computer-Interaktion der Gesellschaft für Informatik (GI) definiert die Mensch-Maschine-Interaktion als „Die benutzerorientierte Analyse und Modellierung von Anwendungskontexten, Prinzipien, Methoden und Werkzeuge, für die Gestaltung von interaktiven, vernetzten Systemen und multimodale und multimediale Interaktionstechniken“ (Gesellschaft für Informatik e.V [GI], 2020). Der Fachbereich Mensch-Maschine-Interaktion der TU Wien stellt dar, dass die menschengerechte Arbeitsgestaltung eine zentrale Zielstellung der Mensch-Technik-Interaktion ist und für die Wettbewerbsfähigkeit von Unternehmen einen Vorteil bringt (Institut für Managementwissenschaften, 2020). Eine erste Darstellung der Beziehung zwischen Mensch und Maschine (Computer) wird im folgendem Modell (Abbildung 34) dargestellt (MacKenzie, 2013).

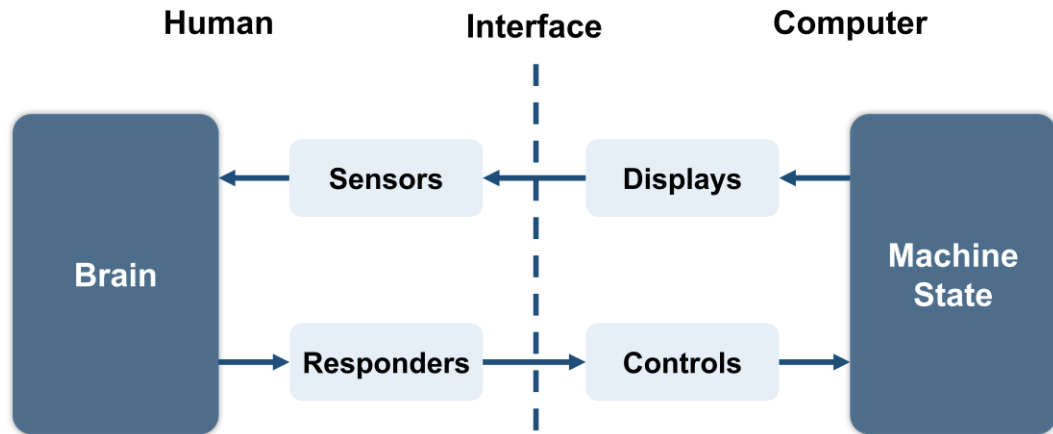


Abbildung 34 | Mensch-Computer-Modell, in Anlehnung an Kantowitz und Sorkin (1983) und MacKenzie (2013)

Dabei wird der Mensch mit dessen Sensoren (z. B. Sehen, Hören, Fühlen etc.) und Responders (z. B. Sprechen, Tastendruck etc.) auf der einen Seite und der Computer mit dessen Ein- und Ausgabegeräten auf der anderen Seite dargestellt. Die strichlierte Linie charakterisiert die Schnittstelle, also jenen Ort, an dem die Interaktion stattfindet, analysiert und optimiert werden kann (MacKenzie, 2013).

Nach Schlick et al. (2018) ist die Mensch-Maschine-Interaktion die Anpassung der technischen Subsysteme (Informationsausgabe/-eingabe) an die Fähigkeiten und Fertigkeiten des Menschen, um arbeitswissenschaftliche Aspekte zu optimieren (Abbildung 35).

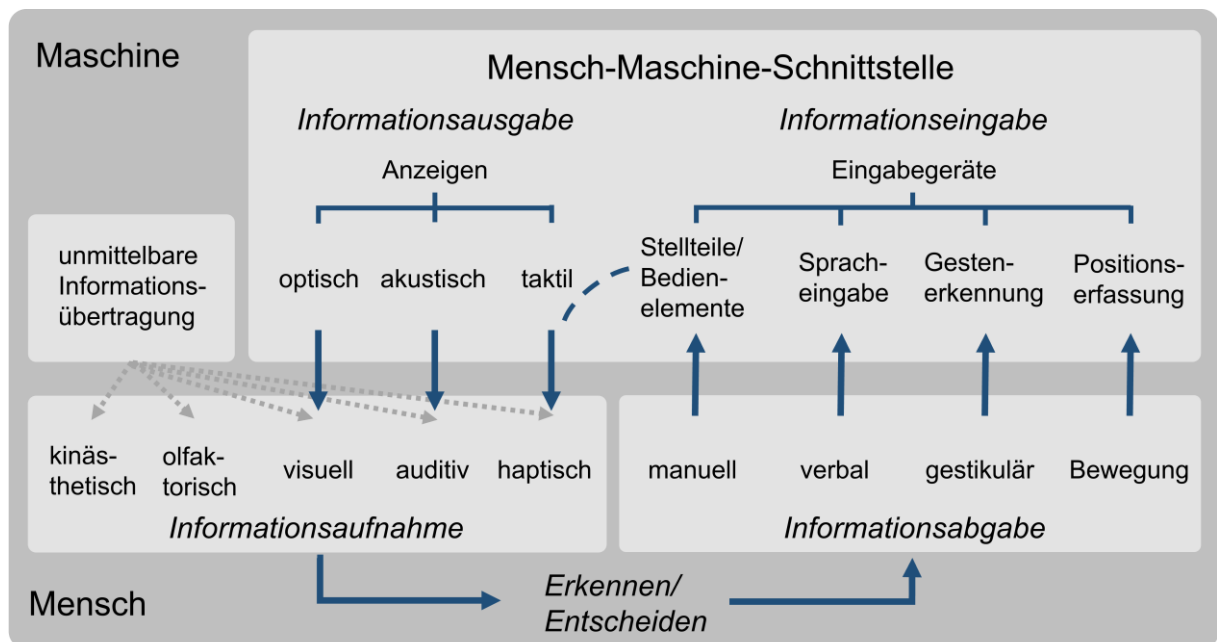


Abbildung 35 | Informationsübertragung in Mensch-Maschine-Systemen, in Anlehnung an Schlick et al. (2018)

Die **Informationsausgabe der Maschine** kann optisch, akustisch oder taktil (haptisch) über verschiedene Anzeigegeräte stattfinden und wird durch unterschiedliche Möglichkeiten (visuell, auditiv, haptisch, kinästhetisch, olfaktorisch) vom Menschen auf- und wahrgenommen.

Nach Erkennen und Interpretation der eingehenden Information sowie Entscheidung über die weiteren Handlungen gibt der Mensch über einen oder mehrere Kanäle (manuell, verbal, gestikulär, Bewegung) die Befehle wieder an die Maschine ab. Die **Informationseingabe** kann mittels verschiedener Eingabegeräte (Bedienteile, Spracheingabe, Gestenerkennung, Positionserkennung etc.) stattfinden, sodass die Information von der Maschine verarbeitet werden kann (Schlick et al., 2018).

Die Gestaltung der Subsysteme und Interaktionsmöglichkeiten, je nach spezifischem Aufgabenkontext, hat großen Einfluss auf die effiziente Umsetzung der Mensch-Maschine-Interaktion und beeinflusst den Informationsverarbeitungsprozess des Menschen in all seinen Phasen (Schlick et al., 2018).

Für die detaillierte Gestaltung der Benutzerschnittstelle sowie die Mensch-Maschine-Interaktion stehen relevante Normen und Gestaltungsrichtlinien zur Verfügung, welche in Abbildung 36 dargestellt werden (Fletcher et al., 2020).

European Machinery Directive	Automation & Robotics	Ergonomics & Human Factors	Digital Systems
EN ISO 12100:2010	EN ISO 10218-1:2011	EN ISO 9241-100:2010	ISO/IEC TR 29194:2015
ISO 13849-1:2015	ISO/TS 15066:2016	EN ISO 9241-110:2006	ISO/IEC 24779-1:2016
IEC 60204-1:2016	BS 8611:2016	EN ISO 9241-129:2010	EN 62264-1:2013
EN 614-2:2000		EN ISO 9241-161:2016	EN 62541
EN ISO 11161:2007		EN ISO 9241-210:2010	EN 62714-1:2014
EN 842:1996 + A1:2008		EN ISO 9241-303:2011	
EN 981:1996 + A1:2008		EN ISO 9241-400:2007	
EN 894-1:1997 + A1:2008		EN ISO 9241-410:2008	
EN 894-2:1997 + A1:2008		EN ISO 9241-920:2009	
EN 894-3:2000 + A1:2008		EN ISO 6385:2016	
EN 894-4:2010		EN 10075-1:2000	
EN 1005-2:2003 + A1:2008		EN ISO 14915-1:2002	
EN 1005-3:2002 + A1:2008			
EN 1005-4:2005 + A1:2008			
EN ISO 14738:2008			

Abbildung 36 | Gestaltungsrichtlinien Mensch-Maschine-Interaktion, in Anlehnung an Fletcher et al. (2020)²

Demnach basiert die angemessene Gestaltung der Benutzer-System-Interaktion auf einem konkreten und klaren Verständnis des Nutzungskontexts, der Arbeitsaufgaben, des Benutzers sowie der erwarteten Ergebnisse. Daraus leitet sich bei der Gestaltung auch eine Zuweisung bzw. Aufteilung der Aufgaben ab, also jene, welche vom Menschen, und jene welche durch das System durchgeführt werden sollten (DIN EN ISO 9241-210). Ein Grundsatz der Gestaltung ist, den Fokus darauf zu legen, wie Benutzer die Arbeitsaufgaben mit dem System ausführen sollen, anstatt zu gestalten, wie das System aussehen soll (DIN EN ISO 9241-210).

2.4.1 Natürliche Mensch-Maschine-Interaktion

Für die optimale Integration des Menschen in die Interaktion muss bei der Gestaltung der Interaktion bzw. des User-Interfaces darauf geachtet werden, dass der Mensch bei der

² Die Norm DIN EN ISO 9241-110:2006 wurde im Oktober 2020 auf die DIN EN ISO 9241-110:2020 aktualisiert.

Ausführung der Tätigkeiten auf sein natürliches Verhalten bzw. seine natürlichen Bewegungen zurückgreifen kann. Diese Konzepte der „Natural User Interfaces“ oder der „Natürlichen Mensch-Maschine-Interaktion“ dienen vor allem der Steuerung mittels Gesten, Mimik oder Sprache, welche mittels Kamerasysteme oder per Berührung umgesetzt werden (Preim & Dachzelt, 2015; Schlick et al., 2018). Gorecky et al. (2014) erwähnen, dass eine natürliche Interaktion im cyberphysischen Raum ähnlich wie mit realen Objekten stattfinden soll. Ebenso wird dargestellt, dass natürliche Interaktion über Sprache und Gesten am intuitivsten zu erlernen und umzusetzen ist. Augmented Reality sowie Sprachassistenten eignen sich demnach besonders, weil hier mit dem Einsatz des Körpers und aller Sinnesorgane die reale mit der digitalen Welt verbunden wird. Darüber hinaus haben sich mit dem Aufkommen der schnellen Datenanalyse und der Verwendung von Hochleistungskameras die Möglichkeiten der visuellen Gestenerkennung stark verbessert (Gorecky et al., 2014). Garcia et al. (2019) erwähnen, dass die interaktive Komponente von cyberphysischen Systemen das Maximum erreicht, wenn sie anhand einer „Natürlichen Mensch-Maschine-Schnittstelle“ (Natural Human-Machine Interface) gestaltet sind. Bei der Entwicklung des Spatial Augmented Reality Assistenzsystems wird einerseits auf die visuelle Informationsausgabe des Systems fokussiert sowie eine geeignete Informationsführung untersucht. Andererseits wird der Schwerpunkt auf die menschliche Informationseingabe über Gesten sowie natürliche Bewegungen gesetzt, welche mit echtzeitfähigen Objekterkennungsalgorithmen umgesetzt werden.

2.4.2 Nutzeradaptive Gesteninteraktion

Der Begriff „Geste“ bedeutet eine spontane oder bewusst eingesetzte Bewegung des Körpers, meist der Hände oder des Kopfes, welche die verbale Ausdrucksweise unterstützt oder mit deren Hilfe etwas Bestimmtes ausgedrückt werden kann (Bibliographisches Institut GmbH, 2021; Kendon, 2004; Preim & Dachzelt, 2015). McNeill (1992) versteht unter Gestik eine Anzahl an Bewegungen durch die Hände und Arme, welche der Kommunikation dienen. Ebenso stellt die Kombination aus Sprache und Gestik die natürliche Interaktion dar (Krueger, 1983, 1993). Die Norm DIN EN ISO 9241-960:2018 beschreibt den Begriff als „Bewegung oder Pose des Körpers bzw. von Körperteilen“. Eine Geste ist demnach die Bewegung von Fingern, Händen, Armen oder weiteren Körperteilen aufgrund einer kommunikativen Absicht, um signifikante Informationen an den Computer zu übermitteln (Preim & Dachzelt, 2015). Für die Gesteninteraktion im Bereich der Mensch-Maschine-Interaktion stellt die mit Händen ausgeführte Gesteninteraktion eine besondere Bedeutung dar (Preim & Dachzelt, 2015).

Dabei können die Handgesten verschiedene Funktionen einnehmen (Cadoz, 1994):

- **Epistemische Funktion:** Die Hand dient als Wahrnehmungsorgan, welches durch Berührung oder Abtasten die Form, Orientierung, Abstand und Gewicht von Objekten erkennen kann.
- **Handlungs- und Betätigungsfunktion:** Die Hand interagiert mit der physischen Umwelt (Objekt verschieben, manipulieren, verformen etc.).

- **Semiotische Funktion:** Über die Gesten werden visuelle oder taktile Informationen an die Umgebung vermittelt. Speziell in der Mensch-Maschine-Interaktion ist diese Funktion für die Übermittlung von Zeichen und Informationen besonders bedeutend, um damit Aktionen auszulösen oder Parameter einer Maschine zu verändern (Cadoz, 1994; La Barré et al., 2009). Dabei können semiotische Gesten in Kombination mit Sprache oder ohne Sprache (autonome Geste) (Kendon, 1986) eingesetzt werden, als intuitiv verständliche Symbole (symbolische Geste) (Quek et al., 2002) oder in einer systematischen Form der Zeichensprache (z. B. Gebärdensprache, Taucheralphabet etc.) auftreten (Preim & Dachzelt, 2015).

Die Norm DIN EN ISO 9241-960:2018 stellt einen Rahmen und Anleitungen zur Gesteninteraktion zur Verfügung (Abbildung 37).

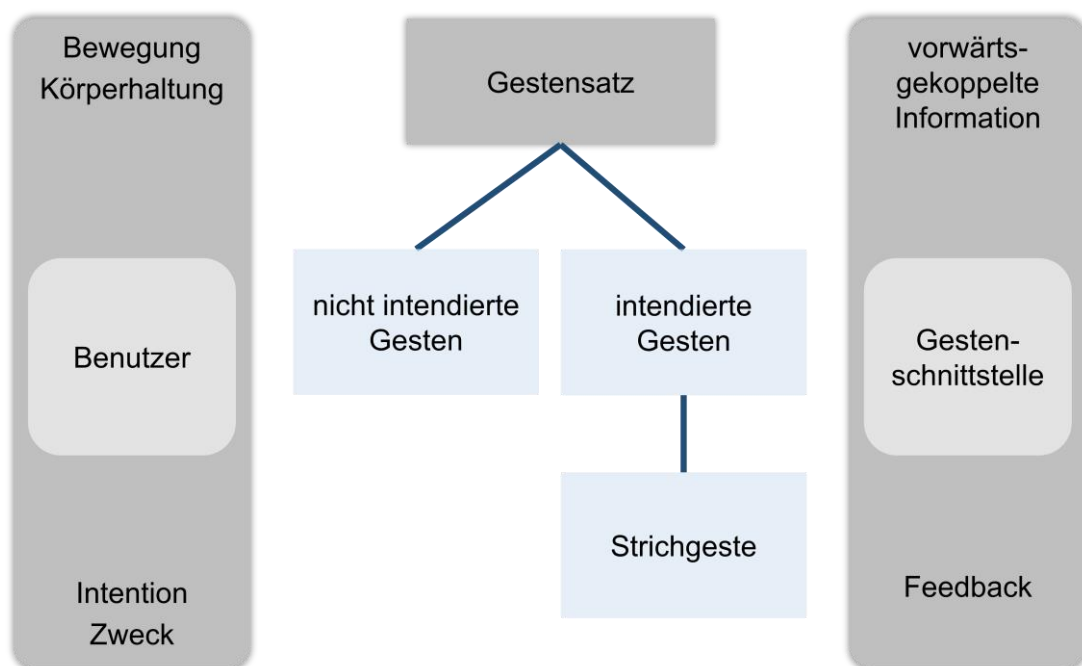


Abbildung 37 | Überblick der Gesteninteraktion, in Anlehnung an die DIN EN ISO 9241-960:2018

Dabei steht ähnlich wie beim Modell zur Mensch-Maschine-Interaktion auf einer Seite der Benutzer mit seinen Bewegungen und der Intention, eine Interaktion über eine Geste auszuführen. Die definierten Gesten, Gestenabfolgen und Gestensätze stellen den zentralen Teil in der Mitte und die (technische) Gestenschnittstelle mit auszulösenden Informationen und Feedback die andere Seite dar (DIN EN ISO 9241-960:2018). Die Norm gibt ebenso Empfehlungen zur Gestaltung und Verwendung von Gesten (DIN EN ISO 9241-960:2018):

- primär Verwendung von bereits vorhandenen, bekannten sowie international genormten und kulturspezifischen Gesten
- durch Benutzer oder die Community vorgeschlagene Gesten
- Gesten, die sich aus dem Nutzungskontext ergeben oder sich in anderen Nutzungskontexten bewährt haben
- für das Gerät oder Technologie typische Gesten

Dabei ist es essenziell, die Konsistenz der Verwendung im Nutzungskontext einzuhalten sowie eine Gestenschnittstelle zu konstruieren, die eine ausreichende Anzahl an Bewegungen unterstützt sowie zwischen intendierten Gesten, solche welche den definierten Zweck erfüllen sollen, und nicht intendierten Gesten (z. B. normale Körpersprache) unterscheiden kann. Eine weitere Gestaltungsrichtlinie ist, dass das System den Nutzer in Kenntnis setzt (Feedback), dass die ausgeführte Geste zum Erfolg geführt hat bzw. anzeigt, dass bzw. wie sich der Zustand des Systems geändert hat (DIN EN ISO 9241-110:2020).

Besonders essenziell für die Gesteninteraktion sind symbolische Gesten (Quek et al., 2002), welche aus kodierten Symbolen oder Zeichen bestehen, die an den Computer übermittelt werden, um einen Befehl auszulösen. Dazu zählen Zeichen und Striche, welche als statische oder dynamische Gesten umgesetzt werden können. Fikkert (2010) hat verschiedene Arten von Gesten und deren natürliche und intuitive Verwendung seitens der Benutzer analysiert. Speziell für das Ab- und Anwählen von Objekten oder z. B. für das Weiterschalten von Informationen stellt sich die Geste „Faust“ und die Geste „offene Hand“ für besonders intuitiv heraus (Fikkert, 2010). In Abbildung 38 wird die Geste „Faust“ gezeigt, welche mittels kamerabasierter Objekterkennung erkannt wurde.

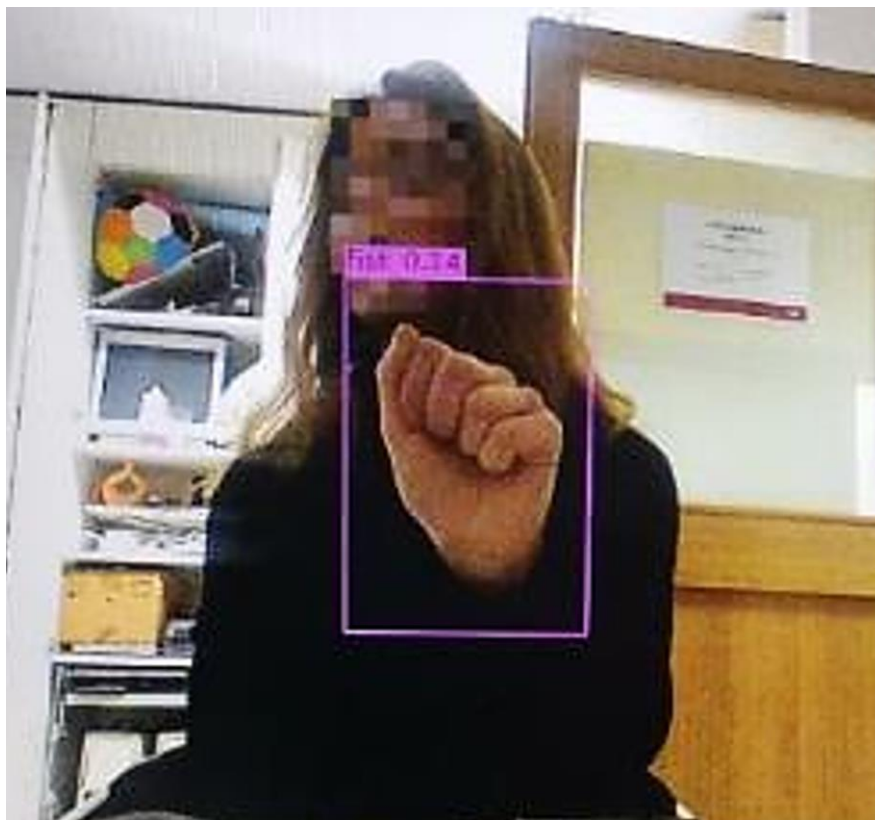


Abbildung 38 | Geste „Faust“ inklusive Objekterkennung (Eigendarstellung)

Gestaltung von Gesten

Bei der Gestaltung von Gesteninteraktionen bzw. der Entwicklung eines Gestensets sollen folgende Eigenschaften diskutiert werden (Preim & Dachse, 2015; Saffer, 2008):

- **Präsenz:** Die Hand soll im Interaktionsraum vorhanden sein, um die Geste für die Interaktion erfassen zu können.

- **Dauer:** Gesten können mit unterschiedlicher Geschwindigkeit ausgeführt werden, was die Dauer beeinflusst. Es muss eine bestimmte Dauer bei der Ausführung der Geste verstreichen, bis die Interaktion ausgelöst wird (um Fehlauflösungen zu vermeiden).
- **Position:** Die Ausführung der Geste kann auf einen definierten zwei- oder dreidimensionalen Raum beschränkt sein. In verschiedenen definierten Zonen können unterschiedliche Aktionen ausgelöst werden.
- **Orientierung:** Es ist essenziell, in welche Richtung bzw. welcher Rotation die Geste ausgeführt wird.
- **Bewegung:** Die Bewegung kann statisch, dynamisch bzw. langsam oder schnell sein, vereinfacht oder komplex sowie raumgreifend oder platzsparend.
- **Größe:** Die Interaktionsgröße oder das Interaktionsvolumen der Geste kann die Interaktionsgröße/-stärke am Computer beeinflussen, d. h. je größer die Geste, umso größer die Interaktion.
- **Objektnutzung:** Es können physische Objekte oder Geräte bei Ausführung der Gestik miteinbezogen werden.
- **Anzahl der Nutzer:** Die Geste kann als Einzeluser oder in Kombination mit einem oder mehreren Usern ausgeführt werden.
- **Bimanualität:** Gesten können unabhängig von linker oder rechter Hand oder mit beiden Händen gleichzeitig ausgeführt werden. Dabei kann eine Hand als Aktivator fungieren und die andere steht für weitere Interaktionen bereit.
- **Konkatenation:** Einzelne Gesten können allein oder in einer Interaktionssequenz aufgeführt werden oder es können bei beidhändiger Verwendung auch parallele Gesten ausgeführt werden.

Weitere Empfehlungen zur Gestaltung von Gestenschnittstellen mit symbolischen Gesten werden im Folgenden präsentiert (Baudel & Beaudouin-Lafon, 1993; Buxton, W., Billingham, M., Guiard, Y., Sellen, A., Zhai, S., 2011; Wachs et al., 2011):

- **Ausnutzung der Muskelspannung:** Der Beginn einer Geste soll durch gezielte Muskelspannung (z. B. der Hand) ausgelöst werden. Das Ende der Geste kann durch Entspannung der Hand signalisiert werden. Dabei sollen aus ergonomischen Gründen lange oder oft wiederholende Anspannungen vermieden werden.
- **Schnelle und reversible Ausführung der Gesten:** Die Ausführung der Geste soll schnell sein und für den Menschen geringen Aufwand bedeuten. Komplexere Gesten können in mehrere Gesten unterteilt werden. Durch Gesten aktivierte Handlungen sollen wieder rückgängig gemacht werden können.
- **Feedback der Gestenaktion:** Der Status der Gestenausführung sollte dem Nutzer stets bekannt sein. Dafür sollte ein adäquates Feedback (z. B. Erfolg oder Misserfolg der Erkennung) nach der Ausführung oder besser währenddessen stattfinden.
- **Intuitive Gesten:** Gesten, welche natürlich in der Ausführung sind, werden von den Nutzern bevorzugt und können schnell erlernt werden (z. B. Wischen). Grundlegende und oft verwendete Funktionen sollen stets mit intuitiven Gesten abgebildet werden. Komplexe Funktionen können durch Gesten mit höherem Aufwand gestaltet werden.

- **Komplexität der Aufgabe:** Gesten sollen für simple Funktionen und Handlungen (z. B. Weiterblättern von Information) eingesetzt werden. Für komplexere Funktionen sollen andere Interaktionsmethoden angedacht werden.
- **Begrenzung der Anzahl von Gesten:** Grundsätzlich sollten so wenig wie möglich Gesten eingesetzt und die Ausführung so unkompliziert wie möglich gehalten werden. Ebenso soll das System nicht ausschließlich durch Gesten gesteuert werden.
- **Individualisierbarkeit von Gesten:** Gesten sollten grundsätzlich individualisierbar sein, d. h. dem Nutzer sollte selbst überlassen werden, mit welcher Geste er die Aktion ausführt, was jedoch einen Widerspruch mit der Konsistenz bedeutet (Norman, 2010).

Prozess der Gestenerkennung: Um Gesten technisch umzusetzen und den Nutzer es zu erlauben, mit seinen Bewegungen Befehle (Gestenbefehl) (DIN EN ISO 9241-960:2018) auszuführen, wird ein „Gestural User Interface“ bzw. eine Gesten-Benutzer-Schnittstelle benötigt (Preim & Dachzelt, 2015). Es gibt mehrere technische Möglichkeiten zur Erkennung der Gesten, folgend wird auf das optische Verfahren eingegangen.

Optische Erkennungsverfahren basieren auf Kamerasystemen bzw. optischen Sensoren, welche den Interaktionsraum zur Erkennung der Gesten erfassen. Dabei werden Methoden der Bildverarbeitung (Computer Vision) (vgl. Kapitel 2.3.4) verwendet, um gestikulierende Körperteile zu identifizieren sowie zu klassifizieren. Die Vorteile optischer Verfahren sind (Wachs et al., 2011):

- Keine Einschränkungen des Nutzers durch z. B. tragbare Devices.
- Sensorische Erfassung ist passiv, lautlos und kontaktlos möglich.
- Kamerasysteme können neben der Gestenerfassung weitere Aufgaben übernehmen.
- Hard- und Software sind kommerziell verfügbar und preiswert.

Die Erkennung sowie Identifikation der Gesten zur Auslösung von Befehlen oder zur Steuerung von Systemen lässt sich in drei Phasen untergliedern (Preim & Dachzelt, 2015):

1. Erfassung, Erkennung und Extraktion relevanter Merkmale (Feature-Erkennung)
2. Verfolgung der Merkmale über eine definierte Zeit (Tracking)
3. Zuordnung einer definierten Bedeutung (Klassifikation)

Demnach bedarf es einer geeigneten Wahl an Soft- und Hardwarekomponenten sowie an Kenntnissen in Computer Vision zur Umsetzung der Gestenerkennung und zur Steuerung von Systemen mit Gestenbefehlen. Die visuelle Gestenerkennung stellt einen potenziellen Mehrwert für die natürliche Interaktion im industriellen Umfeld dar.

Folgend werden die Vor- und Nachteile der Gesteninteraktion (Tabelle 7) dargestellt, um einen Überblick über die Anwendungsmöglichkeiten sowie den Mehrwert für Unternehmen und Mitarbeiter*innen zu geben (Baudel & Beaudouin-Lafon, 1993; Norman & Nielsen, 2010; Preim & Dachzelt, 2015).

Tabelle 7 | Vor- und Nachteile von Gesteninteraktion, in Anlehnung an Preim und Dachzelt (2015)

Vorteile (Baudel & Beaudouin-Lafon, 1993; Preim & Dachzelt, 2015)
Intuitiver Gebrauch Gesten können durch die Vertrautheit/Natürlichkeit problemlos erlernt werden und eignen sich besonders für den Einsatz der Interaktion. Die Erlernbarkeit setzt jedoch ein sorgfältiges Gestendesign voraus.
Gemeinsame Nutzung Anders als bei grafischen Bedienoberflächen, wo grundsätzlich eine sequenzielle Ausführung von Aktionen vorherrscht, erlaubt die Gesteninteraktion die Zusammenarbeit von mehreren Personen sowie die parallele Ausführung von Interaktionen.
Technikfreiheit und Kontaktlosigkeit Gesteninteraktion kommt ohne direkte Eingabegeräte sowie das Tragen von Devices am Körper aus. Deswegen eignet sich der Einsatz besonders in sterilen Umgebungen oder zur kontaktlosen Interaktion.
Schnelligkeit Gesten können bei effektivem Interaktionsdesign und geeigneter Wahl der Erkennungstechnologie äußerst schnelle Interaktionen, auch aus großer Entfernung, ermöglichen.
Nachteile (Norman & Nielsen, 2010; Preim & Dachzelt, 2015)
Fehlende Abgrenzung und Absichtserkennung Bei stetig aktiver Gestenerkennung muss darauf geachtet werden, dass Bewegungen, welche keine Interaktionen ausführen sollen, nicht unbeabsichtigt erkannt werden.
Körperlicher Einsatz und Ermüdung Gesten werden immer mit Körperteilen ausgeführt und erfordern den Einsatz von Muskelaktivitäten. Bei oft wiederholenden Gesten sowie unergonomisch gestalteten Gesten kann es zur Ermüdung kommen. Die schnelle Erkennung der Gesten und ein ergonomisches Gestendesign sind deshalb unabdingbar.
Sichtbarkeit und Entdeckbarkeit Visuelle Elemente zur Ausführung von Aktionen, Menüinhalte oder visuelles Feedback können ohne zusätzliche Ausgabegeräte (Bildschirme, Projektoren) nur bedingt angezeigt werden. Der Nutzer muss vor Bedienung des Systems darüber informiert sein, welche Menüpunkte vorhanden sind bzw. wie und mit welcher Geste die Aktionen ausgeführt werden können.
Indirektheit und fehlendes Feedback Insbesondere haptisches Feedback ist bei der Gesteninteraktion kaum umsetzbar. Demnach müssen alternative Feedbackmöglichkeiten (z. B. visuelles Feedback) gestaltet werden, um die Ausführung der Geste sowie den Status der Aktion anzuzeigen.
Limitierte Anzahl und Komplexität Mit Gesten kann im Vergleich zu anderen Steuerelementen (grafische Menüs) nur eine kleine Anzahl an Aktionen umgesetzt werden. Auch können komplexe Steuerungen über Gesten nur bedingt umgesetzt werden.

Zusammengefasst eignen sich Gesten für die intuitive Umsetzung von natürlichen Interaktionen, speziell für den Einsatz in einem großen Arbeitsraum, wie er in der Großgerätemontage vorherrscht. Für die Entwicklung des Spatial Augmented Reality Assistenzsystems stellt die Interaktion mit natürlichen Gesten einen zentralen Gestaltungsaspekt zur Steuerung des Informationsbereitstellungssystems dar.

2.5 Zusammenfassung der Literaturrecherche

Die wissenschaftliche Relevanz des Forschungsvorhabens wurde mithilfe der Durchführung einer Systematischen Literaturrecherche (vgl. Kapitel 2.1) und der Analyse der Papers ermittelt. Dazu wurden State-of-the-Art-Konzepte, Methoden und Technologien der Forschungsfelder recherchiert, welche als Ausgangslage für das Forschungsvorhaben und die Entwicklung des Assistenzsystems dienen. Tabelle 8 gibt eine Zusammenfassung über die relevanten Elemente der Literaturrecherche.

Tabelle 8 | Zusammenfassung der relevanten Elemente der Literaturrecherche

Forschungsbereich	Kernelemente, Konzepte, Technologien
Informations- bereitstellung in der industriellen Großgeräte- und Baustellenmontage	▪ Hoher manueller Arbeitsanteil (manuelle Montage) und geringer Automatisierungsgrad (Lotter & Schilling, 1994; Lotter & Wiendahl, 2012) ▪ Großer Arbeitsbereich, resultierend durch große Montageobjekte (Lotter & Wiendahl, 2012) ▪ Mensch als zentraler Handlungsakteur (Lotter & Wiendahl, 2012; Reinhart, 2017) ▪ Hoher Grad an qualifizierten Mitarbeiter*innen erforderlich (Gorecky et al., 2014; Lotter & Wiendahl, 2012; Rupprecht et al., 2020) ▪ Strukturierte Arbeitsinformationen notwendig (Rupprecht et al., 2020) ▪ Zentrale Informationsbereitstellung meist in Papierform oder PC-Terminals (Rupprecht et al., 2020)
Spatial Augmented Reality und dynamische In-situ-Projektion	▪ Informationsbereitstellung mit dynamischem Projektionssystem, Adaptives Spatial Augmented Reality Assistenzsystem für die industrielle Baustellenmontage (Rupprecht et al., 2020; Rupprecht et al., 2021) ▪ Dynamisches Projektionssystem mit Spiegelsystem für große Arbeitsbereiche (Dynamic Projection Institut GmbH, 2020a; Dynamic Projection Institute, Herstellungs- und Vertriebs GmbH, 2015; Rupprecht et al., 2020; Rupprecht et al., 2021) ▪ In-situ-Projektion mit statischen Projektoren und einem Mehrwert für Prozesseffizienz und Usability (Funk, 2016; Funk et al., 2017; Funk, Kosch & Schmidt, 2016) ▪ Durchgeführte Forschung zu Spatial Augmented Reality in Kleingeräte- und Einzelplatzmontagen (Büttner et al., 2016; Büttner et al., 2017; Funk, 2016; Funk et al., 2017; Funk, Kosch & Schmidt, 2016; Funk, Kosch, Wolf et al., 2016; Funk, Mayer et al., 2016; Kosch et al., 2018) ▪ Verwendung von Kamerasystemen und Objekterkennung in Kombination mit SAR (Dhiman et al., 2019; Faccio et al., 2019; Funk, 2016; Geiselhart et al., 2016)
Nutzeradaptive Montage und natürliche Mensch-Maschine- Interaktion	▪ Taxonomie der adaptiven Gestaltung in der industriellen Baustellenmontage (Mayrhofer, Rupprecht & Schlund, 2019; Rupprecht & Schlund, 2021; Schlund et al., 2018) ▪ Human-Cyber Physical Production System und Operator 4.0 als Konzepte für die menschenzentrierte Gestaltung in „cyber-physical systems“ (Fantini et al., 2019; Kaasinen et al., 2019; Mattsson et al., 2020; Romero et al., 2016; Romero et al., 2020; Segura et al., 2020; Taylor et al., 2020; Zolotová et al., 2020) ▪ Förderung einer natürlichen Mensch-Maschine-Interaktion (Garcia et al., 2019; Gorecky et al., 2014; Norman, 2010; Preim & Dachzelt, 2015; Schlick et al., 2018) ▪ Eignung der Gesteninteraktion für die natürliche Interaktion im großen Raum (DIN EN ISO 9241-960:2018; Norman, 2010; Norman & Nielsen, 2010; Preim & Dachzelt, 2015) ▪ Assembly 4.0, „self-configured workstation layout“, „aided assembly“ (Bortolini et al., 2017; Bortolini et al., 2019; Bortolini et al., 2020) ▪ Konzept für adaptive Einstellung und Montagegestaltung (acatech, 2019; Kaasinen et al., 2019; Oppermann, 1994; Reinhart, 2017) ▪ Bildererkennung mit Computer Vision und Deep Learning für nutzeradaptive Interaktion (Bochkovskiy et al., 2020; Dhiman et al., 2019; Faccio et al., 2019; He et al., 2016; Joseph Redmon & Ali Farhadi, 2018; O'Mahony et al., 2020; Redmon et al., 2016)

Die ermittelten Kernelemente, Konzepte und Technologien dienen zur Unterstreichung der wissenschaftlichen Relevanz des Forschungsvorhabens und als Input für die Entwicklung des **Spatial Augmented Reality Assistenzsystems**.

3. Spatial Augmented Reality Assistenzsystem mit dynamischer In-situ-Projektion und nutzeradaptiver Gesteninteraktion

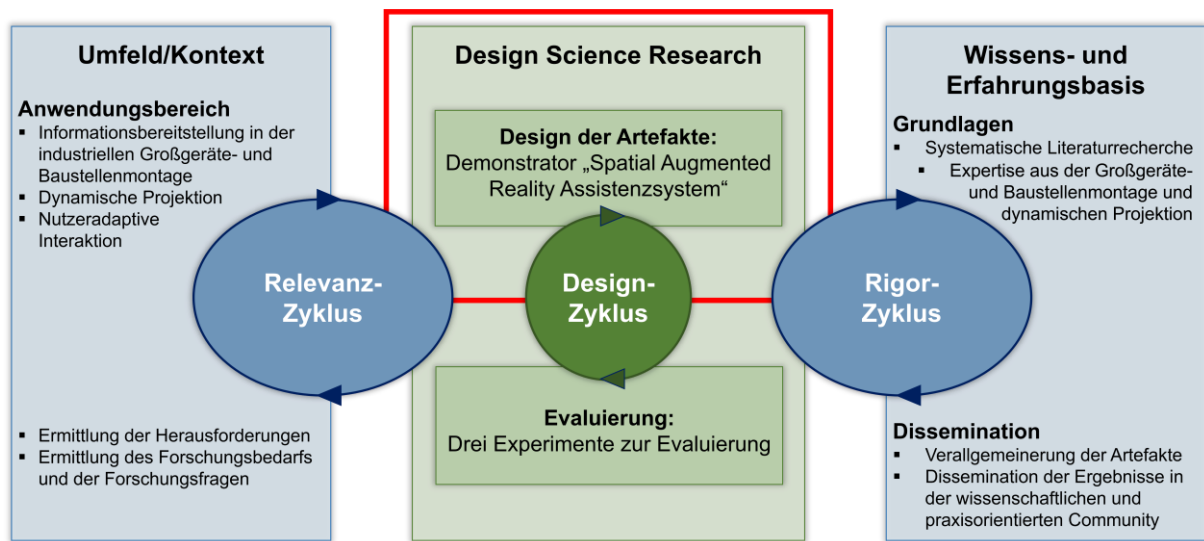


Abbildung 39 | Design der Artefakte, in Anlehnung an Alan R. Hevner et al. (2004)

Das Element „Design der Artefakte“ stellt den Kernpunkt der Design-Science-Research-Methode dar (vgl. Kapitel 1.4 und Abbildung 39), in welchem das Spatial Augmented Reality Assistenzsystem in einem iterativen Designzyklus entwickelt und demonstriert wird (Alan R. Hevner et al., 2004).

In den folgenden Kapiteln werden erste Konzepte des Spatial Augmented Reality Assistenzsystems präsentiert, die technischen Anforderungen, Auswahl der Komponenten und Spezifikationen dargestellt und die Funktionsweise und das Zusammenwirken der Einzelsysteme zu einem Gesamtsystem gezeigt. Daraufgehend wird das System am Demonstrator der Baustellenmontage in der TU Wien-Pilotfabrik präsentiert und die einzelnen Designaspekte der dynamischen In-situ-Projektion sowie der Interaktion dargestellt. Durch die iterative Vorgehensweise werden die Erkenntnisse aus der Demonstration in den Designprozess stetig integriert und die Verbesserungen wiederum in der Demonstration durchgeführt.

3.1 Konzeptidee – Informationsassistenz in der Baustellenmontage

Den detaillierten Konzepten vorausgeschickt, stellt Abbildung 40 das erste Konzept des Spatial Augmented Reality Assistenzsystems in der Großgeräte- und Baustellenmontage dar.

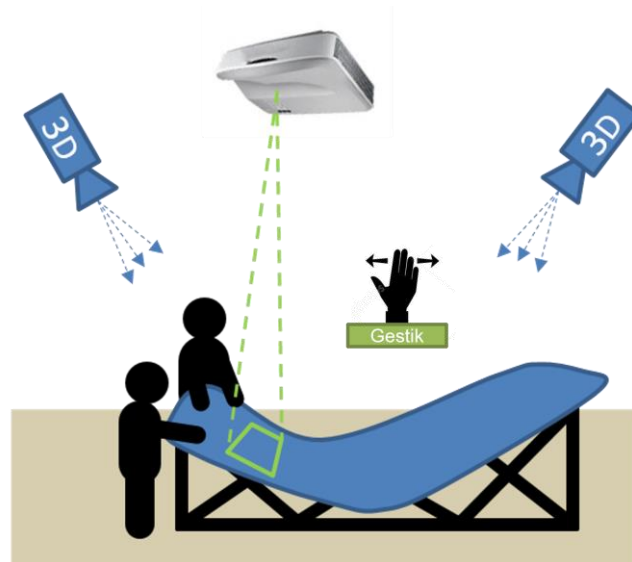


Abbildung 40 | Erste Konzeptskizze eines Assistenzsystems in der Großgeräte- und Baustellenmontage (Eigendarstellung, April 2018)

Die schematische Skizze zeigt eine Montagestation, bei der mithilfe eines Projektors Arbeitsinformationen direkt auf der Arbeitsfläche projiziert werden. Des Weiteren sind 3D-Kameras dargestellt, welche die Mensch-Maschine-Interaktion mittels Gesteninteraktion umsetzen, um damit das Informationssystem zu steuern.

Durch die Systematische Literaturrecherche und State-of-the-Art-Analyse (vgl. Kapitel 2) konnte festgestellt werden, dass in Montagesystemen mit kleinen Arbeitsbereichen (z. B. Einzelplatzmontage) bereits Spatial Augmented Reality Assistenzsysteme zur Anzeige von Arbeitsinformationen eingesetzt werden (Funk, 2016; Funk, Kosch et al., 2015; Funk, Kosch & Schmidt, 2016). Dabei werden vor allem statische Projektoren zur Projektion von Informationen auf die Montagefläche der Einzelplatzmontage sowie Kamerasysteme zur Interaktion mit dem Menschen verwendet. Diese Spatial Augmented Reality Assistenzsysteme eignen sich aufgrund des kleinen Projektions- und Interaktionsbereich nur mangelhaft für den Einsatz in einem großen Arbeitsbereich, wie er in der Großgeräte- und Baustellenmontage vortreffend ist. Ebenso wird beschrieben, dass intuitiv aufbereitete Arbeitsinformationen sich besonders für den ergonomischen und effizienten Einsatz eignen (Schlick et al., 2018) und dass die In-situ-Projektion speziell bei der Einzelplatzmontage im Vergleich zu anderen Informationsbereitstellungsvarianten Vorteile bietet (Funk, Kosch & Schmidt, 2016). Im Zuge der ersten Recherche konnte festgestellt werden, dass bereits in der Veranstaltungstechnik ein geeignetes Patent sowie marktfähige Produkte zur dynamischen Projektion über große Arbeitsbereiche mithilfe von Projektoren und einem digital steuerbaren Spiegelsystem verwendet werden (Dynamic Projection Institute, Herstellungs- und Vertriebs GmbH, 2015) (vgl. Kapitel 1.2.2). Für die Mensch-Maschine-Interaktion konnte außerdem recherchiert werden, dass speziell die natürliche Interaktion (Garcia et al., 2019; Gorecky et al., 2014; Norman, 2010) z. B. mit Gesten (DIN EN ISO 9241-960:2018; Norman, 2010) intuitiv erlern- und anwendbar sei und dass Kamerasysteme mit Verwendung von Computer Vision und Deep Learning (Dhiman et al., 2019; He et al., 2016; Redmon et al., 2016)

vielversprechende Möglichkeiten zur Interaktion in der Montage bringen (Dhiman et al., 2019; Faccio et al., 2019; Funk, 2016).

Ausgehend von der ersten Konzeptidee sowie den Recherchen zum Stand der Wissenschaft und des Systems aus der Veranstaltungstechnik liegt der Fokus der vorliegenden Forschungsarbeit auf der Gestaltung eines Spatial Augmented Reality Assistenzsystems, welches ein dynamisches Projektionssystem einsetzt, um über einen großen Arbeitsbereich direkt an den jeweiligen Montageorten Informationen gezielt anzuzeigen. Darüber hinaus liegt der Fokus auf der natürlichen Mensch-Maschine-Interaktion, welche mittels Geste die Interaktion mit dem Informationssystem ermöglicht. Um die natürliche Mensch-Maschine-Interaktion im großen Arbeitsbereich umzusetzen, werden die Möglichkeiten von modernen Kamerasystemen und der Einsatz von Computer Vision und Deep Learning verwendet.

Demnach soll ein Spatial Augmented Reality Assistenzsystem entwickelt werden, welches in der **industriellen Großgeräte- und Baustellenmontage** effektiv und ergonomisch eingesetzt werden kann und folgende Anforderungen erfüllt:

1. **Dynamische In-situ-Projektion von Arbeitsinformationen über einen großen Arbeitsbereich und direkt an der jeweiligen Arbeitsposition**
2. **Dynamische Führung durch den Arbeitsprozess mittels projizierter Informationen**
3. **Intuitive und projektionsoptimierte Gestaltung der Arbeitsinformationen**
4. **Nutzeradaptive, natürliche Mensch-Maschine-Interaktion (Mensch- und Gestenerkennung) mittels Kamerasystemen, Bilderkennung und Deep Learning**

Die Konkretisierung der Konzeptidee laut den genannten Anforderungen wird in Abbildung 41 anhand eines detaillierteren Renderings dargestellt.

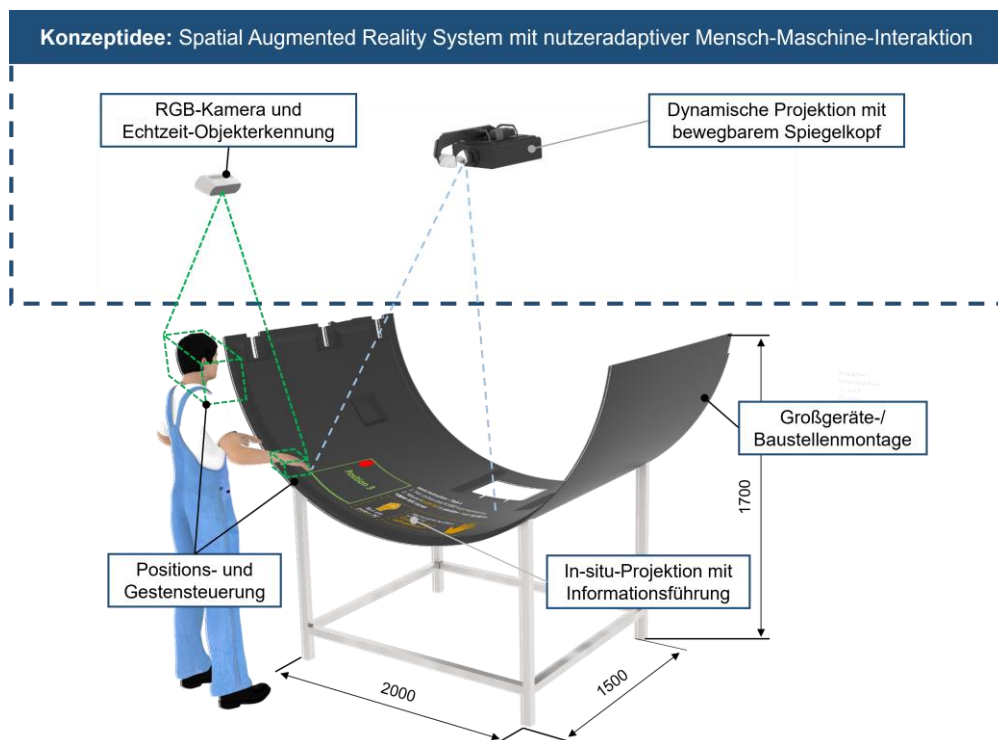


Abbildung 41 | Konzeptidee eines Spatial Augmented Reality Systems mit dynamischer Projektion und nutzeradaptiver Interaktion, in Anlehnung an Rupprecht et al. (2020)

Das Rendering zeigt den Einsatz eines dynamischen Projektionssystems mit bewegbarem Spiegelkopf aus der Veranstaltungstechnik, das zur dynamischen In-situ-Projektion der Arbeitsinformationen in der Großgeräte- und Baustellenmontage eingesetzt wird. Es zeigt den Menschen als Kernelement der manuellen Montage sowie Konzepte zur visuellen Objekterkennung (Mensch, Pose, Geste) mittels Kamera und Echtzeit-Objekterkennung. Die Erkennung des Menschen und dessen Merkmale stellt die Ausgangssituation der nutzeradaptiven Interaktion mit dem Informationssystem dar. Nach Definition der allgemeinen Anforderungen und Präsentation der ersten Konzeptideen ist es für die konkrete Gestaltung der Spatial Augmented Reality erforderlich, die Anforderungen der einzelnen Systeme und Komponenten zu definieren, die Spezifikationen zu ermitteln, geeignete Systeme und Technologien auszuwählen sowie das Zusammenspiel und die Vernetzung aufzuzeigen. Nach Definition der Spezifikationen, werden die einzelnen Gestaltungselemente in iterativen Design-Zyklen am Demonstrator aufgebaut und gestaltet. Der Demonstrator dient einerseits zur Veranschaulichung des entwickelten Systems und andererseits für die experimentelle Evaluierungen mit verschiedenen Anwendergruppen.

Abbildung 42 zeigt schematisch den Aufbau des Design- und Demonstrationssystems sowie die einzelnen Designelemente des Spatial Augmented Reality Assistenzsystems.

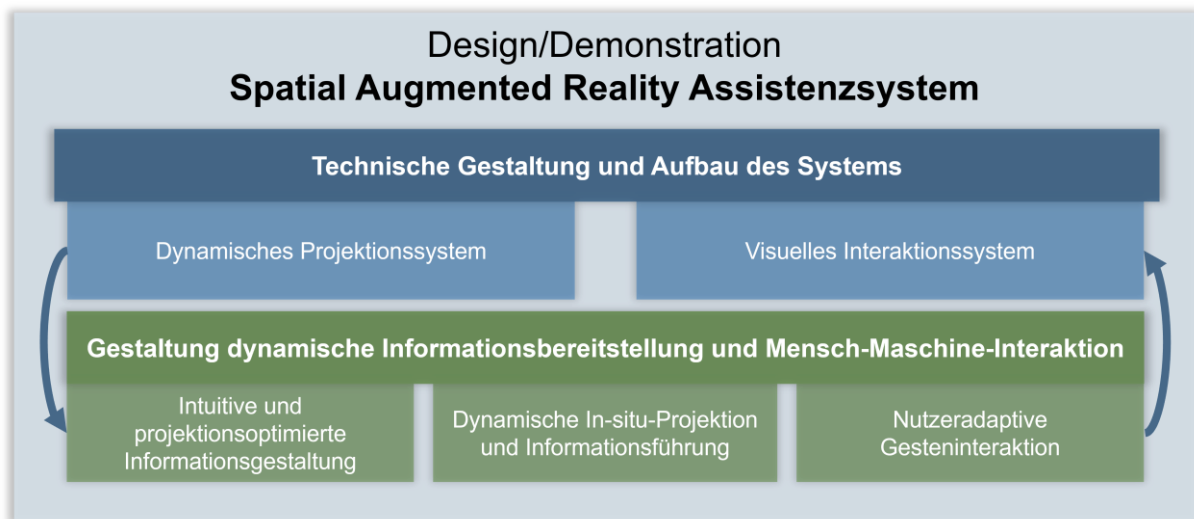


Abbildung 42 | Design- und Demonstrationselemente des Spatial Augmented Reality Assistenzsystems (Eigendarstellung)

Dabei wird in Kapitel 3.2 der technische Aufbau des Systems sowie die Auswahl der einzelnen Spezifikationen des dynamischen Projektionssystems sowie des visuellen Interaktionssystems beschrieben. In Kapitel 3.3 wird anschließend die Gestaltung der dynamischen Informationsbereitstellung mit den einzelnen Designelementen „Intuitive und projektionsoptimierte Informationsgestaltung“, „Dynamische In-situ-Projektion und Informationsführung“ dargestellt und die „Nutzeradaptive Gesteninteraktion“ präsentiert.

3.2 Technische Gestaltung und Aufbau des Spatial Augmented Reality Assistenzsystems

Folgendes Kapitel präsentiert den technischen Aufbau sowie die einzelnen Komponenten des Spatial Augmented Reality Assistenzsystems mit dem dynamischen Projektionssystem als Kernelement der Informationsausgabe und dem visuellen Interaktionssystem zur Informationseingabe. Die Darstellung und das Zusammenspiel der einzelnen Systeme und Komponenten des Systems wird in Abbildung 43 dargestellt.

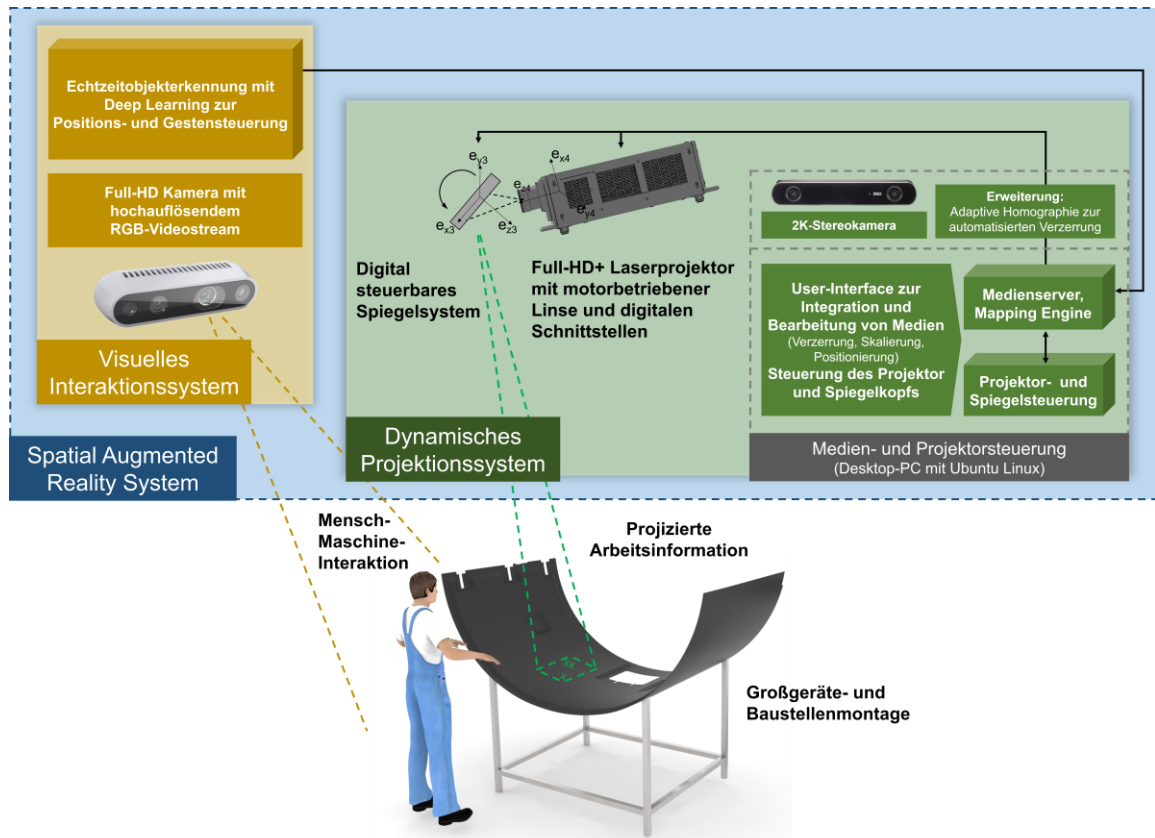


Abbildung 43 | Aufbau des Assistenzsystems, in Anlehnung an Rupprecht et al. (2020)

Das **Spatial Augmented Reality Assistenzsystem** besteht aus folgenden Elementen:

1. Dynamisches Projektionssystem

- Full-HD+ Projektor mit motorbetriebem Linsensystem und digitalen Schnittstellen
- Digital steuerbares Spiegelsystem
- Medien- und Projektorsteuerung (Desktop-PC mit Ubuntu Linux)
 - Authoring- und Medienbearbeitungssoftware inklusive Mapping Engine
 - User-Interface zur Steuerung des Projektor- und Spiegelsystems
 - Adaptive Homographie zur automatisierten Einstellung der Verzerrung

2. Visuelles Interaktionssystem

- Full-HD-Kamera mit hochauflösendem RGB-Videostream
- Echtzeitobjekterkennung mit Deep-Learning-Algorithmus „YOLO“ zur Gesteninteraktion
- Synthetisch erzeugter Gestendatensatz

3.2.1 Dynamisches Projektionssystem

Das dynamische Projektionssystem ist das zentrale Element des Spatial Augmented Reality Assistenzsystems zur Umsetzung der Bereitstellung und Anzeige der Arbeitsinformationen direkt am Arbeitsplatz (vgl. Kapitel 1.2.2). Dabei stellen der Full-HD+ Projektor zur Anzeige der Medien, das Spiegelsystem für die dynamische Bewegung der projizierten Informationen im Raum und die Software und Steuerungseinheiten die Kernkomponenten dar. Der effektive Einsatz des dynamischen Projektionssystems ist demnach das Zusammenspiel der verschiedenen Hard- und Softwarekomponenten, welche eine intuitive Integration von Arbeitsinformation, Erstellung von Projektionsshows sowie die effiziente digitale Ansteuerung von Projektor und Spiegelsystem ermöglichen. Die Anforderungen an die einzelnen Komponenten, der Auswahlprozess sowie die Darstellung der technischen Spezifikation inklusive Abschätzung der Anschaffungskosten werden folgend im Detail beschrieben.

a) Full-HD+ Projektor mit motorbetriebenem Linsensystem und digitalen Schnittstellen

Für die Gestaltung des Spatial Augmented Reality Assistenzsystems in der Baustellenmontage ist es erforderlich, einen leistungsfähigen Projektor auszuwählen, welcher über große Bauräume und Entfernungen sowie bei Tageslicht hochwertige Projektionen in ausreichender Helligkeit liefert, damit Arbeitsinformationen detailgetreu auf der Arbeitsfläche der Montage-station angezeigt werden können. Aufgrund der Tatsache, dass der Projektor primär an der Decke einer Produktionshalle oder einer Vorrichtung in z. B. 4 Metern Höhe montiert wird, ist eine weitere Anforderung, dass das Objektiv und die Linse inklusive dessen Parameter, wie Zoomfaktor, Schärfe und Shutter (Bild aus/ein), von der Ferne aus gesteuert werden können. Demnach soll der Projektor über die notwendigen digitalen Schnittstellen verfügen, um über Softwarekomponenten erzeugte Befehle die Steuerung zu ermöglichen. Für den Einsatz im Produktionsbereich sind ebenso die hohe Lebensdauer sowie die Wartbarkeit der Lichtquelle von Relevanz. Um die zweite Kernkomponente, das Spiegelsystem, optimal verwenden zu können, muss bei der Auswahl des Projektors die Kompatibilität mit dem Spiegelkopfsystem gewährleistet sein. Um diesen Anforderungen gerecht zu werden, wurde ein geeigneter Full-HD+ Projektor der Firma Panasonic (Abbildung 44) mit einem Gewicht von 22,4 Kilogramm und einer Abmessung von 498 x 200 x 581 Millimetern ausgewählt. Die Details der Spezifikationen werden in Tabelle 9 dargestellt.



Abbildung 44 | Projektor Panasonic PT-RZ660BE (Panasonic Connect Europe GmbH)

Tabelle 9 | Spezifikationen des Projektors Panasonic PT-RZ660BE³

Beschreibung	Spezifikationen
Type des Projektors, Farbe	Panasonic PT-RZ660BE, schwarz
Preis	9299,60 Euro netto
Typ der Beleuchtungsquelle	Laserprojektor
Leuchtmittellebensdauer	20 000 h
Projektionstechnik	DLP (Digital Light Processing)
Objektiv	Standard Objektiv T.R.: 1.7 –2.4:1
Helligkeit, Kontrast	6200 ANSI Lumen, 10.000:1
Auflösung, Format	1920 x 1200 WUXGA/Full-HD+, 16:10
Gewicht, Abmessungen	22,4 kg, 498 x 200 x 581 mm (B x H x T)
Besonderheiten der Linse	Lens-shift, motorisch angetriebener Focus und Zoom
Anschlüsse, Schnittstellen	HDMI, DVI-D, D-sub 15pin, 5 x BNC RGBHV, RS232C Din 9 pin, RJ45, SDI IN BNCx1 (3G/HD/SD-SDI), DIGITAL LINK, Lan-Netzwerk kompatibel PJLink

Der ausgewählte Projektor verwendet als Beleuchtungsquelle einen Laser mit einer Leuchtmittellebensdauer von über 20 000 h sowie eine motorisch angetriebene Linse, welche über digitale Schnittstellen gesteuert werden kann. Der Projektor erfüllt demnach die benötigten Anforderungen für den Einsatz in der industriellen Baustellenmontage.

b) Digital steuerbares Spiegelsystem

Als zweite Komponente für die bewegte Projektion wird das dynamische Spiegelkopfsystem dargestellt. Um die Projektion über einen großen Bauraum zu gestalten, können folgende Möglichkeiten angedacht werden:

- Einen Projektor mit hohem Abstand zur Projektionsfläche montieren.
- Mehrere Projektoren über den Arbeitsbereich verteilen und softwarebasiert vernetzen.
- Den Projektionsstrahl umlenken bzw. bewegbar gestalten.

Im iterativen Designprozess der Gestaltung des Projektionssystems wurden Gespräche mit Expert*innen der visuellen Bildverarbeitung durchgeführt, um den optimalen Einsatz für den Montagebereich zu eruieren. Das Ergebnis war, dass für den Anwendungsfall mit einem Arbeitsbereich von 5500 × 2000 × 1000 Millimetern (Winglet-Montage, vgl. Kapitel 4.1) mehrere Projektoren sowie ein Hochleistungsrechner zur Vernetzung und Synchronisation der Projektionen notwendig sind (Abbildung 45), um über das ganze Bauteil projizieren zu können.

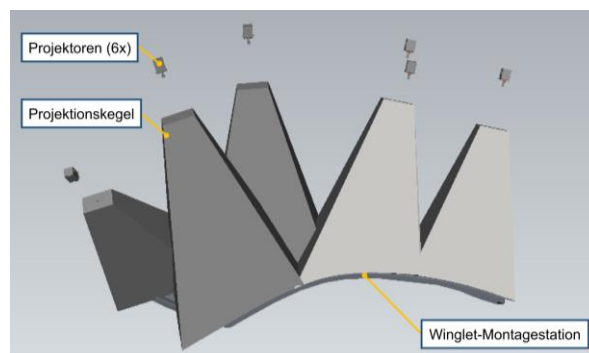


Abbildung 45 | Projektorenanordnung für den Winglet-Demonstrator (Eigendarstellung)

³ https://www.panasonic-center.at/shop/id,3068,Panasonic_PT-RZ660BE_Projektor; Angebot 18-0449 vom 14.11.2018 von Dynamic Projection Institute (Anhang)

Die Ausleuchtung sieht sechs Projektoren vor, welche synchronisiert über der Montagestation angebracht werden. Die Überschlagsrechnung der Investitionskosten für den Einsatz der sechs Projektoren ergaben Anschaffungskosten von rund 70 000 Euro (sechs Projektoren à ca. 10 000 Euro, ein Hochleistungsrechner inklusive Hard- und Softwarekomponenten ca. 10 000 Euro). Diese hohen Anschaffungskosten stellen damit keine kosteneffiziente sowie wirtschaftlich sinnvolle Verwendung für die Baustellenmontage dar. Wie in Kapitel 1.2.2 beschrieben, konnte jedoch ein innovatives Spiegelsystem (Abbildung 46) recherchiert werden, welches auf dem Patent der Ablenkung des Projektionsstrahles basiert (Dynamic Projection Institute, Herstellungs- und Vertriebs GmbH, 2015) und eine potenzielle Möglichkeit für den Einsatz in der Baustellenmontage darstellt.



Abbildung 46 | Mirror-Head zur dynamischen Projektion, in Anlehnung an Dynamic Projection Institut GmbH (2020a)

Die zentrale Komponente des Spiegelsystems ist ein bewegbarer und hochglanzbeschichteter Spiegel, welcher über eine digitale Steuereinheit und einen Schrittmotor in Dreh- und Neigungsrichtung (Pan/Tilt) bewegt werden kann. Die Steuereinheit des Spiegelmotors ist mit einer Ambient-Light-Funktion zur Signalanzeige ausgestattet und enthält integrierte Anschlüsse zur Vernetzung mit dem Projektor bzw. dem Authoringsystem. Der Mirror-Head kann an Wand bzw. Decke montiert werden und enthält eine Aufnahme für den Projektor. Die technischen Spezifikationen werden in Tabelle 10 angezeigt.

Tabelle 10 | Spezifikationen Spiegelsystem, in Anlehnung an Dynamic Projection Institut GmbH (2020c)

Spiegel und Spiegelsteuerung
Beschichteter Spiegel für optimale Reflexion ~98 % Brechungs faktor
Hochauflösende 16-Bit-Pan/Tilt-Spiegelbewegung
Genauigkeit der Bewegung: Schwenken/Neigen aus gleicher Richtung zum definierten Punkt innerhalb ~0,01°
Hochauflösende Schrittmotoren mit wartungsfreiem Direktantrieb
Selbstkalibrierende Spiegelposition für hohe Genauigkeit
Absolute Wiederpositionierungsgenauigkeit kleiner als 0,03°
Ambient-Light-LED mit Hochleistungs-RGB-LEDs
Digitale Schnittstellen und Anschlüsse
DMX-512-Steuerung über 14 DMX-Kanäle
Art-Net™-Anschluss (RJ45)
RS232-Anschluss für Projektor (D-Sub-9-Stecker)

Das System zur Ablenkung des Projektionsstrahls mittels Spiegelsystem stellt im Vergleich zur vorhergehenden Variante eine neuartige und vor allem kostengünstige Möglichkeit für den

Einsatz im Montagebereich dar. Das Ergebnis einer Überschlagrechnung nach dem vorliegenden Angebot⁴ liegt bei rund 15 000 Euro Gesamtinvestitionskosten⁵ (ein Projektor inklusive Deckenhalterung 9500 Euro, ein Spiegelkopfsystem 2500 Euro, Medienserver inklusive Softwarelizenzen 2500 Euro, Bildschirm und Kleinteile 500 Euro).

Für den Anwendungsfall und den ausgewählten Projektor wird das passende Spiegelsystem des Typs Mirror Head – MH11 (Abbildung 47), welches für DLP-Projektoren mittlerer Größe ausgelegt ist, sowie der zugehörige Deckenmontagebügel des Typs MHZ-BW mit VESA-Aufnahme verwendet (Dynamic Projection Institut GmbH, 2020d, 2020e).

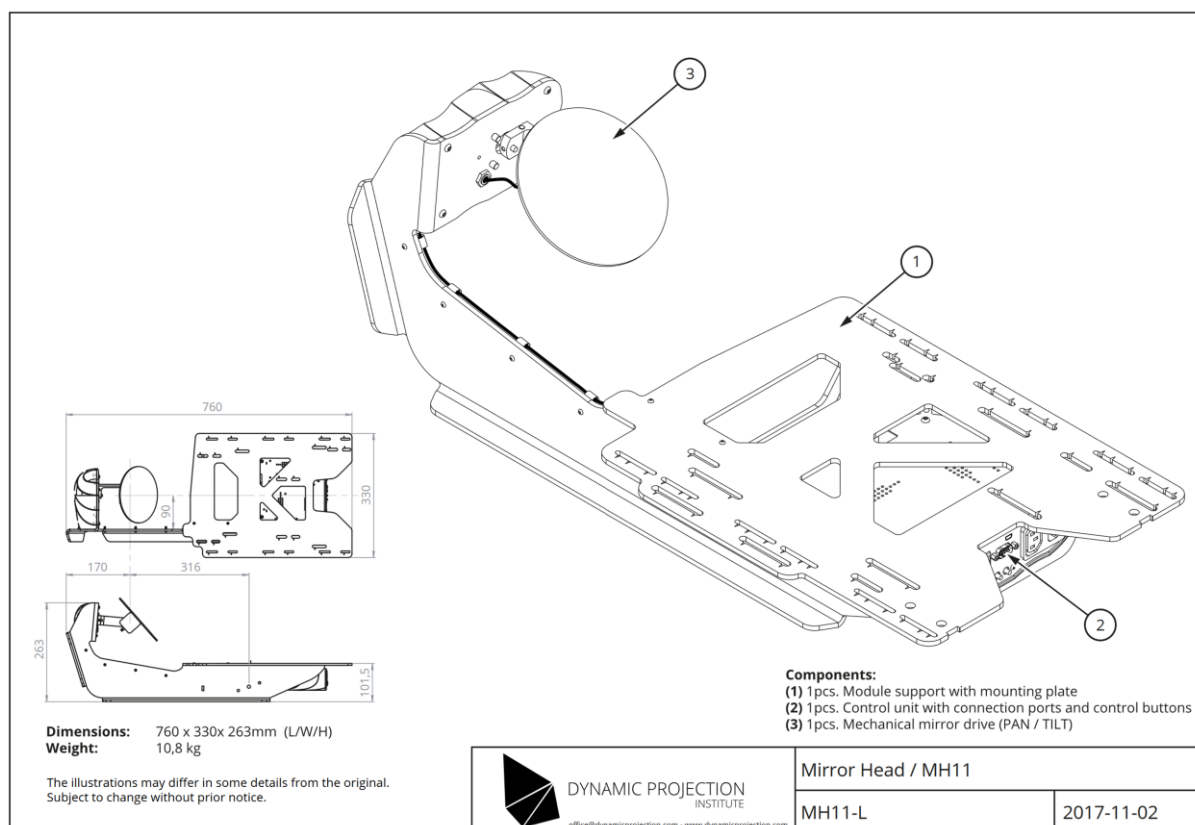


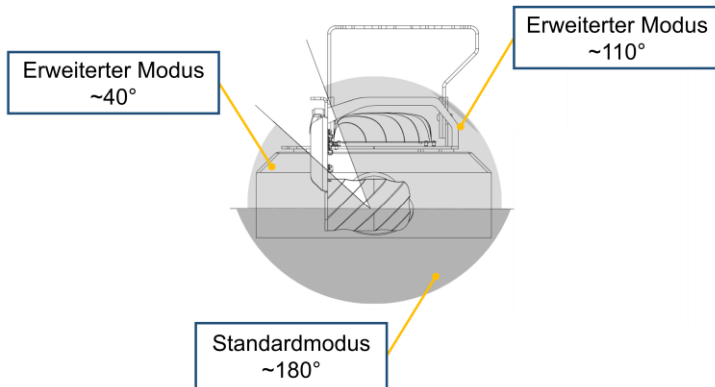
Abbildung 47 | Mirror-Head Typ MH11 (Dynamic Projection Institut GmbH, 2020d)

Der Spiegelkopf inklusive Steuerungseinheit hat ein Gewicht 10,8 Kilogramm und eine Abmessung von 760 x 330 x 263 Millimetern. Um den Projektor an der Decke über der Montagestation zu platzieren, wird ein Deckenmontagebügel mit einem Eigengewicht von 2,8 Kilogramm verwendet. Demnach beträgt das Gesamtgewicht des Spiegelsystems inklusive Projektor 36 Kilogramm. Für die Verwendung im großen Arbeitsbereich der Baustellenmontage kann der ausgewählte Mirror-Head des Typs MH11, die Drehung und Neigung des Spiegels softwarebasiert einstellen und den Winkelbereich des Projektionskegels in zwei Betriebsmodi (Standard/Erweitert) definieren (Abbildung 48).

⁴ Angebot 18-0449 vom 14.11.2018 von Dynamic Projection Institute (Anhang)

⁵ Tabelle 14 gibt eine detaillierte Aufstellung der Kosten des Assistenzsystems.

Frontansicht (Spiegeldrehung)



Seitenansicht (Spiegelneigung)

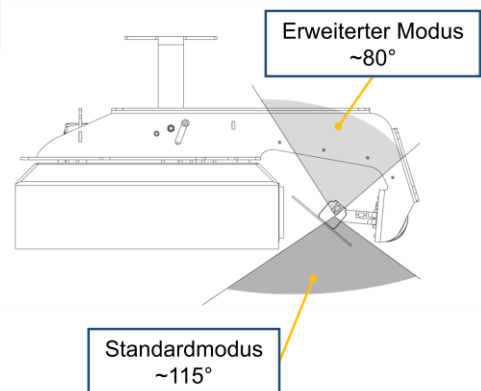


Abbildung 48 | Drehung und Neigung des Spiegelsystems, in Anlehnung an Dynamic Projection Institut GmbH (2020c)

Der Standardmodus verfügt über 180° im Spiegeldreh- und 115° im Spiegelneigungsbereich. Der erweiterte Modus vergrößert die Drehung um 150° und die Neigung um 80°. Um die Drehung und Neigung des Spiegelkopfs und der daraus resultierenden Positionen der Projektion zu steuern, ist es notwendig, die Befehle zentral z. B. über ein Softwaresystem (Medienserver) an den Projektor und Spiegelkopf zu senden. Die einzelnen Komponenten und Verbindungen, welche dafür notwendig sind, werden schematisch in Abbildung 49 dargestellt.

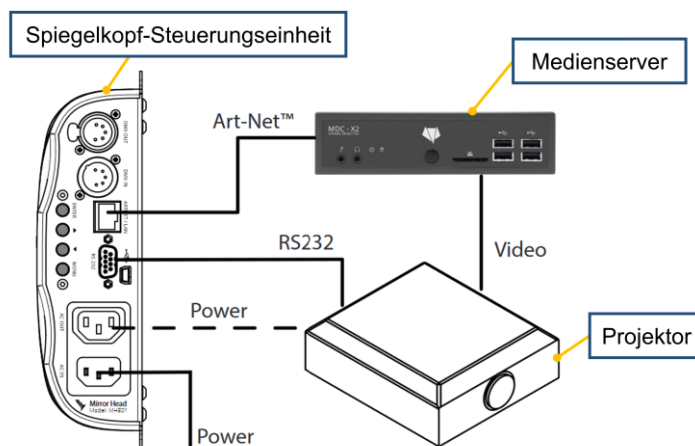


Abbildung 49 | Vernetzung des Spiegelsystems zum Medienserver und Projektor, in Anlehnung an Dynamic Projection Institut GmbH (2020c)

Die Verbindung des Spiegelkopfs mit dem Medienserver erfolgt über das Protokoll Art-Net. Art-Net ist eine vereinfachte Implementierung des DMX512-Protokolls, bei der die Steuerungsinformationen in IP-Paketen übertragen werden. Typischerweise geschieht das über ein lokales Netzwerk wie Ethernet oder WLAN. Das Ambient-Light und die zugehörigen RGB-LEDs des Spiegelkopfes werden ebenso über Art-Net gesteuert. Die Verbindung vom Spiegelkopf zum Projektor erfolgt über eine RS232-Verbindung, die vom Medienserver zum Projektor über HDMI. Der Spiegelkopf bietet ebenso die Möglichkeit der Fernsteuerung des Projektors über DMX-Befehle mittels RS232-Schnittstelle.

c) Medien- und Projektorsteuerung

Die Informationen (Bilder, Videos, Arbeitsanweisungen etc.), welche über den Projektor bzw. über den Spiegelkopf auf das Bauteil in der Montage projiziert werden, werden zentral über eine Authoring- und Medienbearbeitungssoftware (Desktop-PC mit Ubuntu Linux) (Abbildung 50) integriert und mittels Mapping Engine an den Projektor weitergegeben.



Abbildung 50 | Links: User-Interface-Authoring- und Medienbearbeitungssoftware, rechts: PC zur Medien- und Projektorsteuerung (Dynamic Projection Institut GmbH, 2020b)

Die **i) Authoring- und Medienbearbeitungssoftware inklusive Mapping Engine** wird verwendet, um die Arbeitsinformationen, Medien- und Projektionsdaten in das System zu integrieren und eine korrekte und verzerrungsfreie Ausrichtung und Darstellung der Projektion am Montagebauteil umzusetzen. Dabei können die zu projizierenden Informationen je nach vorhandener Projektionsfläche und -position manuell angepasst und bearbeitet werden (Verzerrung, Drehung, Zoom etc.).

Mittels **ii) User-Interface zur Steuerung des Projektor- und Spiegelsystems** können automatisierte Projektionsshows und Prozessabfolgen von Arbeitsinformationen im Arbeitsraum durchgeführt sowie das Einstellen von Projektionspositionen (Pan/Tilt-Stellung des Spiegelkopfes) mittels Speicherung von Pre-Sets umgesetzt werden. Zu jeder Position können der Zoomfaktor, der Fokus sowie weitere Parameter des Projektors eingestellt und zugewiesen werden. Um die erstellten Befehle wie Positionen oder Verzerrungen richtig über den Projektor und Spiegelkopf zu senden, wird die softwarebasierte Mapping Engine verwendet, die im User-Interface grafisch aufbereiteten Bilddaten in zugehörige Softwarebefehle umwandelt und die bearbeiteten und entzerrten Bilddaten den Steuerungseinheiten von Spiegelkopf und Projektor übermittelt (Mapping).

Funktionen der Medienmanagementsoftware im Überblick:

- Integration von Bildern, Videos, Texten und Animationen
- Morphen und Mesh-Korrektur (Verzerrung, Zoom, Drehung) sowie Mapping der zu projizierenden Medien
- Projektionsabfolge und Showerstellung mittels Timeline-Steuerung und Pre-Sets
- DMX Pan/Tilt-Steuerung (Position der Projektion im Raum über Spiegelbewegung)

Das System zur **Medien- und Projektorsteuerung** ist ein für die Anwendungen mit dem Spiegelkopf adaptierter Mini-PC der Marke Shuttle mit der Bezeichnung „MDC-X2/DH170“.

Der PC ist mit einem Intel Core i5-6600 und 3,30 GHz Prozessor sowie einer Speicherkapazität von 120 GB SSD ausgestattet und bietet Schnittstellen wie zwei Display Ports, ein HDMI sowie zwei RJ45 und integriertes WLAN. Der Mini-PC hat ein Gewicht von 1,4 Kilogramm und ist inklusive zeitlich unlimitierter akademischer Lizenz für die Medien- und Steuerungssoftware verfügbar.

Authoringprozess mit manuell eingestellter Verzerrung

Im Folgenden wird die Vorgehensweise (Abbildung 51) zur Integration von Medien, der Erstellung von dynamischen Projektionen sowie der Verwendung der Authoring- und Medienbearbeitungssoftware dargestellt. Dabei wird im Detail demonstriert, wie Arbeitsinformationen integriert und Projektionspositionen definiert werden und wie die manuelle Anpassung der Skalierung und Verzerrung umgesetzt wird. Folgend wird auf die Erstellung der Projektionsabfolge sowie deren Speicherung in Form von Pre-Sets eingegangen.

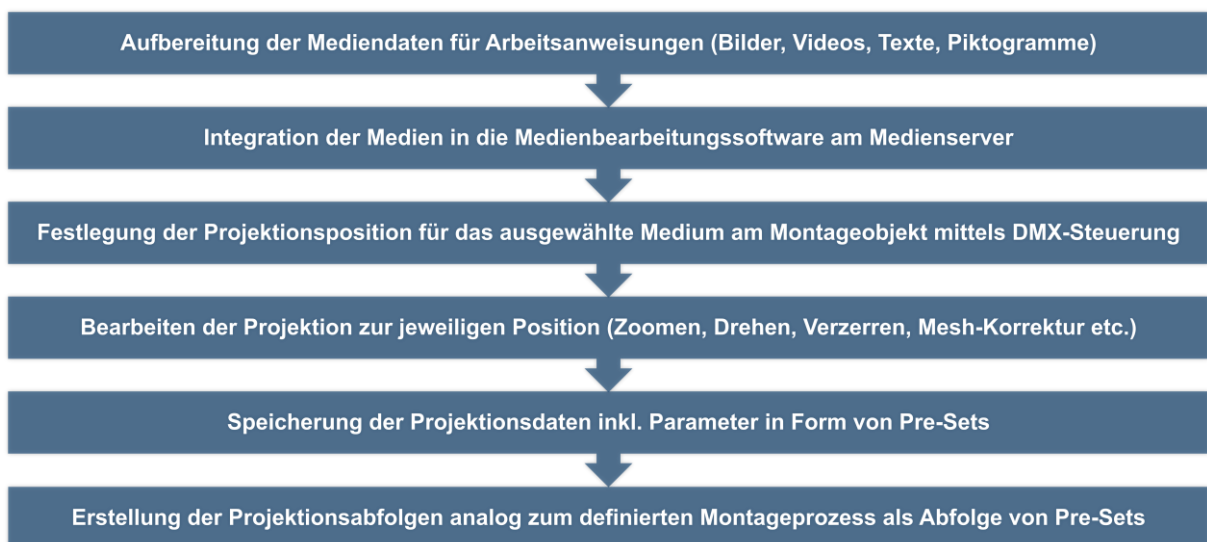


Abbildung 51 | Vorgehensweise zur Erstellung von dynamischen Projektionen (Eigendarstellung)

Bevor die Arbeitsanweisungen ins Projektionssystem integriert werden, besteht die Anforderung, dass die Informationen vorab optimal, intuitiv und projektionsgerecht mit geeigneten Bild- oder Videobearbeitungstools⁶ aufbereitet werden sollen (vgl. Kapitel 3.3.1). Anschließend können die Informationen in verschiedene Bild- oder Videoformate oder als Texte in die Authoring- und Medienbearbeitungssoftware des Projektors integriert werden. Danach werden die Positionen und Abläufe, wie sie im Montageprozess definiert sind, in der Software eintrainiert und die jeweiligen Parameter und Eigenschaften der Projektionen festgelegt. Dabei wird mittels DMX-Steuerung⁷ der Spiegelkopf zur gewünschten Projektionsposition in der Baustellenmontage angefahren und anschließend die Verzerrung

⁶ z. B. Adobe Photoshop, Adobe Illustrator, MS Visio, MS PowerPoint oder ähnliche Programme.

⁷ DMX steht für „Digital Multiplex“ und ist ein Steuerungsprotokoll, welches in der Bühnen- und Veranstaltungstechnik für die Steuerung von Scheinwerfern und bewegbaren Lichtelementen verwendet wird.

und Skalierung, je nach Krümmung des Montageobjektes sowie Position in der Montage, definiert (Abbildung 52).

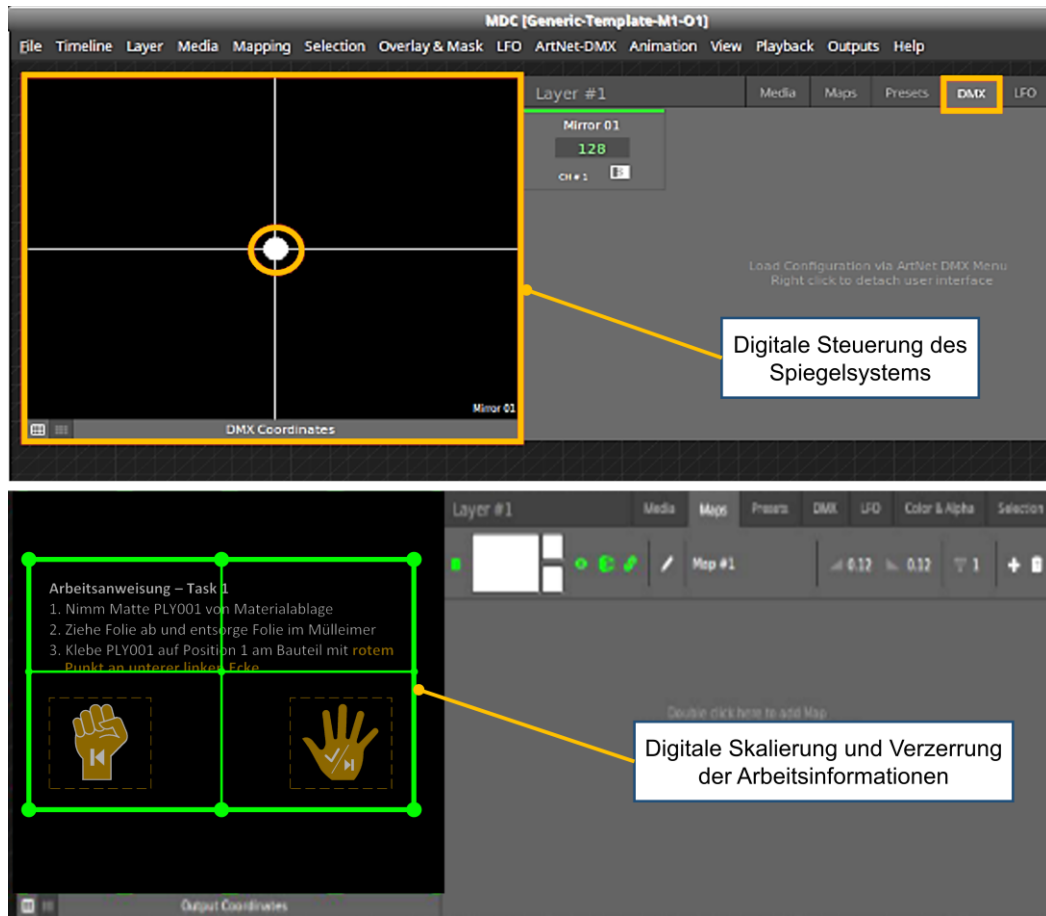


Abbildung 52 | Oben: Digitale Positionssteuerung des Spiegelsystems, unten: digitale Anpassungsmöglichkeit (Skalierung, Drehung, Verzerrung) der integrierten Daten (Eigendarstellung)

Der Spiegel kann mittels Bewegung des weißen Punktes digital zur jeweiligen Montageposition gesteuert werden. Anschließend können die projizierten Daten skaliert, verzerrt und ausgerichtet werden, sodass sie verzerrungsfrei am Bauteil angezeigt werden. Die definierten Parameter zur Projektionsposition, Verzerrung, Orientierung und Skalierung werden anschließend in Pre-Sets abgespeichert und können später mittels Programmbefehlen (z. B. über Gesteninteraktion) abgerufen werden. Durch Verwendung der Authoring- und Medienbearbeitungssoftware sowie der Mapping Engine können Arbeitsinformationen intuitiv integriert und mittels Projektionssystem dynamisch angezeigt werden. Die offenen Schnittstellen und Vernetzungsmöglichkeiten bieten die Möglichkeit, mit programmierten Befehlen, weiteren Geräten (z. B. Wearables, Smartphones, WLAN-Taster etc.) oder mithilfe der Gestensteuerung, die Systeme zu steuern und die Positionen und Arbeitsinformationen abzurufen. Um den Arbeitsaufwand für die Aufbereitung der Arbeitsanweisungen und die manuelle Erstellung der Projektionsabfolge inklusive Einstellung der Verzerrung einschätzen zu können, wird in Tabelle 11 eine Aufschlüsselung der Zeitaufwände präsentiert.

Tabelle 11 | Arbeitsaufwand zur Erstellung einer dynamischen Projektion

Authoringprozess (vgl. Abbildung 51)	Zeitaufwände
Aufbereitung der Mediendaten für Arbeitsanweisungen (Bilder, Videos, Texte, Piktogramme)	60 min, abhängig von der Komplexität der Arbeitsanweisung
Integration der Medien in die Medienbearbeitungssoftware am Medienserver	3 min
Festlegung der Projektionsposition für das ausgewählte Medium am Montageobjekt mittels DMX-Steuerung	5 min
Bearbeiten der Projektion zur jeweiligen Position (Zoomen, Drehen, Verzerren, Mesh-Korrektur etc.)	10 min
Speicherung der Projektionsdaten inkl. Parameter in Form von Pre-Sets	2 min
Erstellung der Projektionsabfolgen analog zum definierten Montageprozess als Abfolge von Pre-Sets	5 min
Gesamtaufwand (ohne Aufbereitung der Mediendaten)	25 min
Gesamtaufwand (mit Aufbereitung der Mediendaten)	85 min

Der Arbeitsaufwand zur Integration von Mediendaten für die Verwendung im dynamischen Projektionssystem beträgt rund 25 min. Der Aufwand zur Erstellung und Aufbereitung der Mediendaten bzw. Arbeitsanweisungen hängt von verschiedenen Parametern (Komplexität der Information, Position und Beschaffenheit der Projektionsfläche, Projektionsumgebung etc.) ab und wird mit durchschnittlich 60 min je Arbeitsinformation abgeschätzt.

iii) Adaptive Homographie zur automatisierten Einstellung der Verzerrung

Über das Medien- und Projektorsteuerungssystem können die zu projizierenden Bilder bzw. Arbeitsinformationen, wie oben beschrieben, manuell je nach Position im Raum und je nach Gegebenheiten der Oberfläche (z. B. Krümmung der Montagestation etc.) angepasst und verzerrt werden. Um den Prozess der Informationsaufbereitung sowie die manuelle Anpassung der Arbeitsinformationen zu verbessern, wurde ein Setting bzw. Algorithmus zur automatisierten Verzerrung („Adaptive Homographie“) implementiert (Hornacek et al., 2021). Zur Aufnahme und Analyse der Projektionsfläche sowie zur Umsetzung der „Adaptiven Homographie“ wurde die Stereokamera „Stereolabs ZED 2“⁸, mit folgenden Spezifikationen (Tabelle 12) eingesetzt (Hornacek et al., 2021).

Tabelle 12 | Spezifikationen der Stereokamera „ZED 2“, in Anlehnung an Stereolabs Inc. (2021)

Beschreibung	Spezifikationen
Abmessungen	175 x 30 x 33 mm
Preis	449 Dollar ⁹ entspricht rund 398 Euro (Wechselkurs vom 21.12.2021)
Video-Output	Auflösung 2,2K (4416 x 1242) mit 15 fps ¹⁰ Auflösung 1080p (3840 x 1080) mit 30 fps
Tiefenkamera-Spezifikationen	Neural-Stereo-Depth-Sensing-Technology Field of View (FOV): 110° (H) x 70° (V) x 120° (D) Depth Range: bis 20 m, Depth Frame Rate: bis zu 100 fps

⁸ Die Stereokamera ZED-2 wurde ausschließlich für die Umsetzung der „Adaptiven Homographie“ verwendet. Für die Gesteninteraktion wurde das RGB-Modul der Intel RealSense Kamera (vgl. Kapitel 3.2.2) verwendet.

⁹ <https://www.stereolabs.com>, abgerufen am 10.12.2021

¹⁰ Anmerkung: fps = frames per second (Bilder pro Sekunde)

Die ZED-2-Stereokamera bietet mit bis zu 2,2K eine hohe Auflösung und mit einem FOV von 110° (H) x 70° (V) x 120° (D) einen großen Blickwinkel. Die Kamera wurde in einem definierten Abstand zum Projektor montiert und der Projektor, die Kamera sowie die Projektionsfläche zueinander kalibriert (Abbildung 53) (Hornacek et al., 2021).

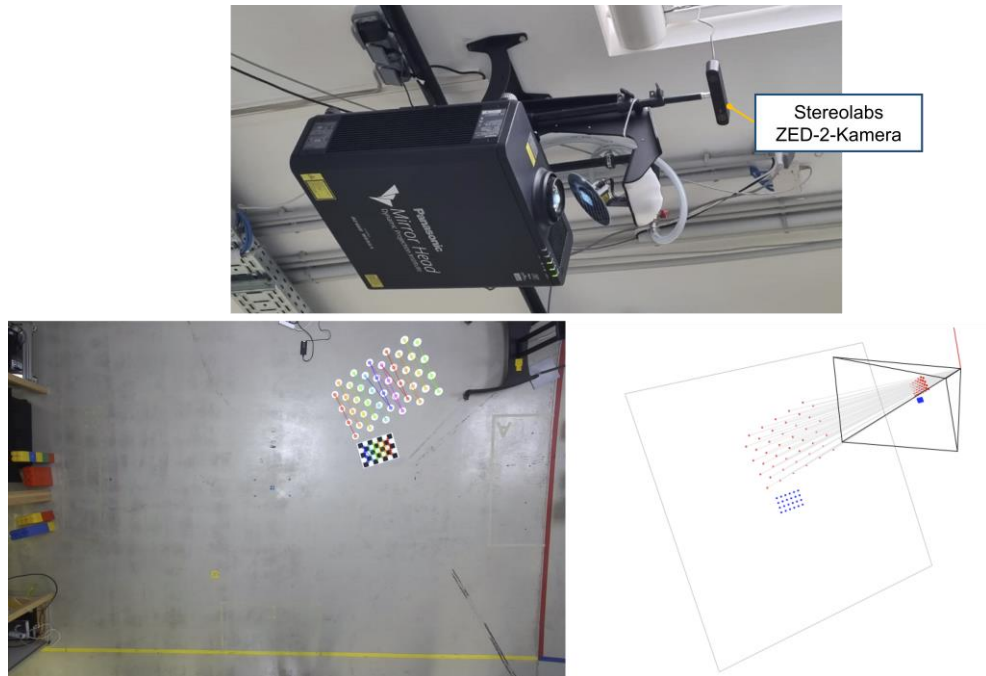


Abbildung 53 | Oben: Projektor mit Stereolabs ZED-2-Kamera, unten links: Schachbrettmuster und projizierte Punkte zur Kalibrierung, unten rechts: Schematische Darstellung der Projektionsfläche und Stereokamera, in Anlehnung an Hornacek et al. (2021)

Dabei wurde ein Schachbrettmuster mit definierten Abständen der Schwarz-Weiß-Felder verwendet, um den Abstand der jeweiligen Projektionsebene zur Kamera bzw. zum Projektor zu berechnen. Des Weiteren wurde ein spezifisches Punktemuster an den definierten Positionen der Arbeitsinformationen projiziert, welches gleichzeitig mit der Stereokamera aufgenommen wurde. Ausgehend vom projizierten Punktemuster wurde die jeweilige Homographie bzw. Verzerrung zu den einzelnen Positionen berechnet (Hornacek et al., 2021). Anschließend wurden die jeweiligen Arbeitsinformationen integriert und von der „Adaptiven Homographie“ so umgewandelt, dass das Bild für den/die Betrachter*in verzerrungsfrei auf der Projektionsebene dargestellt wurde (Abbildung 54) (Hornacek et al., 2021).



Abbildung 54 | Beispiel der automatisierten Verzerrung (Eigendarstellung)

Durch die Anwendung der automatisierten Verzerrung wird die Projektion (gelbe Linien) begradigt und dementsprechend „entzerrt“ dargestellt. Die „Adaptive Homographie“ ermöglicht dadurch einen verbesserten Prozess zur Erstellung von dynamischen Projektionsabfolgen, reduziert den Aufwand der manuellen Einstellung der Verzerrung und bietet neue Möglichkeiten zur Integration von Arbeitsinformationen. Demnach können definierte Projektionspositionen im Arbeitsraum vorab kalibriert, die zugehörige Homographie berechnet und anschließend die Arbeitsinformationen mit dem Projektionssystem verzerrungsfrei dargestellt werden. Die Adaptive Homographie ist eine Weiterentwicklung des dynamischen Projektionssystems und wurde in der Großgeräte- und Baustellenmontage als Prototyp getestet.

3.2.2 Visuelles Interaktionssystem

Das visuelle Interaktionssystem stellt das zweite Hauptsystem des Spatial Augmented Reality Assistenzsystems dar und wird zur Umsetzung der nutzeradaptiven Interaktion mit natürlichen Gesten eingesetzt. Dadurch dass ein Spatial Augmented Reality Assistenzsystem eine Interaktion in Echtzeit erfordert (vgl. Azuma, 1997), bedarf es der Auswahl geeigneter Software- und Hardwarekomponenten sowie des Einsatzes von echtzeitfähigen Objekterkennungsalgorithmen. Ebenso bedarf es unkomplizierter technischer Möglichkeiten, die Gesteninteraktion in der Montage zu implementieren sowie den wirtschaftlich sinnvollen Einsatz zu ermöglichen. Durch die Analyse von bestehenden Spatial Augmented Reality Assistenzsystemen und der State-of-the-Art-Recherche konnte ermittelt werden, dass sich visuelle Erfassungssysteme (Kamerasysteme) mit RGB-Videostreams besonders für die Umsetzung der natürlichen Mensch-Maschine-Interaktion, insbesondere der Gesteninteraktion, eignen und somit für den Einsatz in der Montage geeignete Lösungen darstellen (Dhiman et al., 2019; Preim & Dachsel, 2015). Des Weiteren konnte festgestellt werden, dass Objekterkennungsalgorithmen mit Deep Learning und „Single-stage“-neuronalen Netzwerken eine hohe Erfassungsgeschwindigkeit aufweisen und sich besonders für die Implementierung in der industriellen Montage sowie für die Umsetzung der Echtzeitfähigkeit eignen (Joseph Redmon & Ali Farhadi, 2018; Redmon et al., 2016; Redmon & Farhadi, 2017).

Demnach sind die zwei Hauptkomponenten des visuellen Interaktionssystems der Einsatz einer Full-HD-Kamera zur Erzeugung eines hochauflösenden RGB-Videostreams und die Verwendung einer Echtzeitobjekterkennung mit Deep Learning zur Umsetzung der Gesteninteraktion. Die detaillierten Anforderungen und die Auswahl der Komponenten und Spezifikationen sowie der Aufbau der Gesteninteraktion werden folgend dargestellt.

a) Full-HD-Kamera mit hochauflösendem RGB-Videostream

Für die visuelle Objekterkennung von Menschen, deren Positionen und Bewegungen bedarf es hochauflösender Kamerasysteme, geeigneter Schnittstellen sowie leistungsfähiger Verarbeitungsmöglichkeiten der Videostreams durch geeignete Hard- und Softwarekomponenten. In Forschungen zu statischen Spatial Augmented Reality Assistenzsystemen

in der Einzelplatzmontage werden als Interaktionskomponente meist RGB-/Stereokamerasysteme der Marke „Microsoft Kinect“ (Funk, 2016; Korn et al., 2015) oder hochauflösende Kameras der Marke „Intel“ (Abbildung 55) verwendet. Speziell für die Bild- und Objekterkennung sowie Trackinglösungen im industriellen Umfeld und vor allem für den Einsatz bei kollaborativen Robotern und Augmented Reality Lösungen wird der Typ „Intel RealSense D435i“ eingesetzt, um den Systemen das „Sehen“ zu ermöglichen (Intel Corporation, 2020).



Abbildung 55 | Rendering „Intel RealSense – D435i (Intel Corporation, 2020)

In Tabelle 13 werden die Spezifikationen der Kamera „Intel RealSense D435i“ dargestellt (Intel Corporation, 2020).

Tabelle 13 | Spezifikationen der Kamera „Intel RealSense Typ D435i“, in Anlehnung an Intel Corporation (2020)

Beschreibung	Spezifikationen
Abmessungen	90 × 25 × 25 mm
Preis	250 Euro ¹¹
RGB-Modul und Videostream-Spezifikationen	RGB-Sensor-Auflösung: 1920 × 1080 RGB-Frame Rate: 30 fps RGB-Sensor Field of View (H × V × D): 69.4° × 42.5° × 77° (±3°)
Tiefenkamera-Spezifikationen	Active-Infrarot-Stereo-Tiefentechnologie Depth Field of View: 86° × 57° (±3°) Depth Output Resolution: bis zu 1280 × 720 Depth Range: bis zu 10m Depth Frame Rate: bis zu 90 fps

Die Kamera ermöglicht mit dem RGB-Modul einen hochauflösenden Full-HD-Videostream mit einer Frame Rate von 30 fps, welche sich besonders für die weitere Analyse mit modernen Bilderkennungsalgorithmen zur Ermittlung von Objekten und Gesten eignet. Außerdem eignet sich die Kamera mit einem Field of View von 69.4° × 42.5° × 77° (H × V × D) für den Einsatz in der Baustellenmontage und kann durch die kompakte Abmessung im Spatial Augmented Reality System integriert werden¹².

Folgend wird der geeignetste Algorithmus für die Gegebenheiten der Baustellenmontage ausgewählt und der Aufbau und die Vorgehensweise zur Erkennung von Menschen und Gesten sowie zur Steuerung des dynamischen Projektionssystems beschrieben.

¹¹ www.amazon.de, abgerufen am 28.09.2020

¹² Anmerkung: Für die Objekterkennung zur Gesten- und Positionserkennung wurde ausschließlich das RGB-Modul der Kamera verwendet. Der Tiefensensor bzw. die Stereofunktion wurde in den ersten Versuchen für die „Adaptive Homographie“ getestet, jedoch aufgrund geeigneterer Spezifikationen (vgl. Tabelle 12), anschließend zur ZED-2-Kamera gewechselt.

b) Echtzeitobjekterkennung mit Deep-Learning-Algorithmus „YOLO“

Für die Objekterkennung wird der echtzeitfähige und leistungsfähige Deep-Learning-Algorithmus YOLOv3¹³ (vgl. Kapitel 2.3.4) ausgewählt, weil sich dieser aufgrund der schnellen Objekterkennungszeit und hohen Objekterkennungspräzision eignet und bereits im praktischen sowie industriellen Umfeld fundiert getestet wurde. Des Weiteren besteht eine große Community sowie eine gute Dokumentation der verwendeten Algorithmen, Frameworks und Softwarepackages in der Code-Verwaltungsplattform github.com. Um die Bilderkennung mit YOLO und dem Deep-Learning-Netzwerk umzusetzen, werden Trainingsdaten der zu erkennenden Objekte (Menschen, Gesten, Posen etc.) benötigt, welche für den Anwendungsfall aufbereitet wurden (vgl. Kapitel 4.4). Des Weiteren müssen die Frameworks, Algorithmen sowie die trainierten Daten implementiert und mit den ausführenden Systemen (Medien- und Projektorsteuerung) verbunden werden. Die schematische Vorgehensweise der Objekterkennung über die einzelnen Komponenten und Schnittstellen bis zur Ansteuerung der Projektion mithilfe des dynamischen Projektionssystems wird in Abbildung 56 dargestellt.

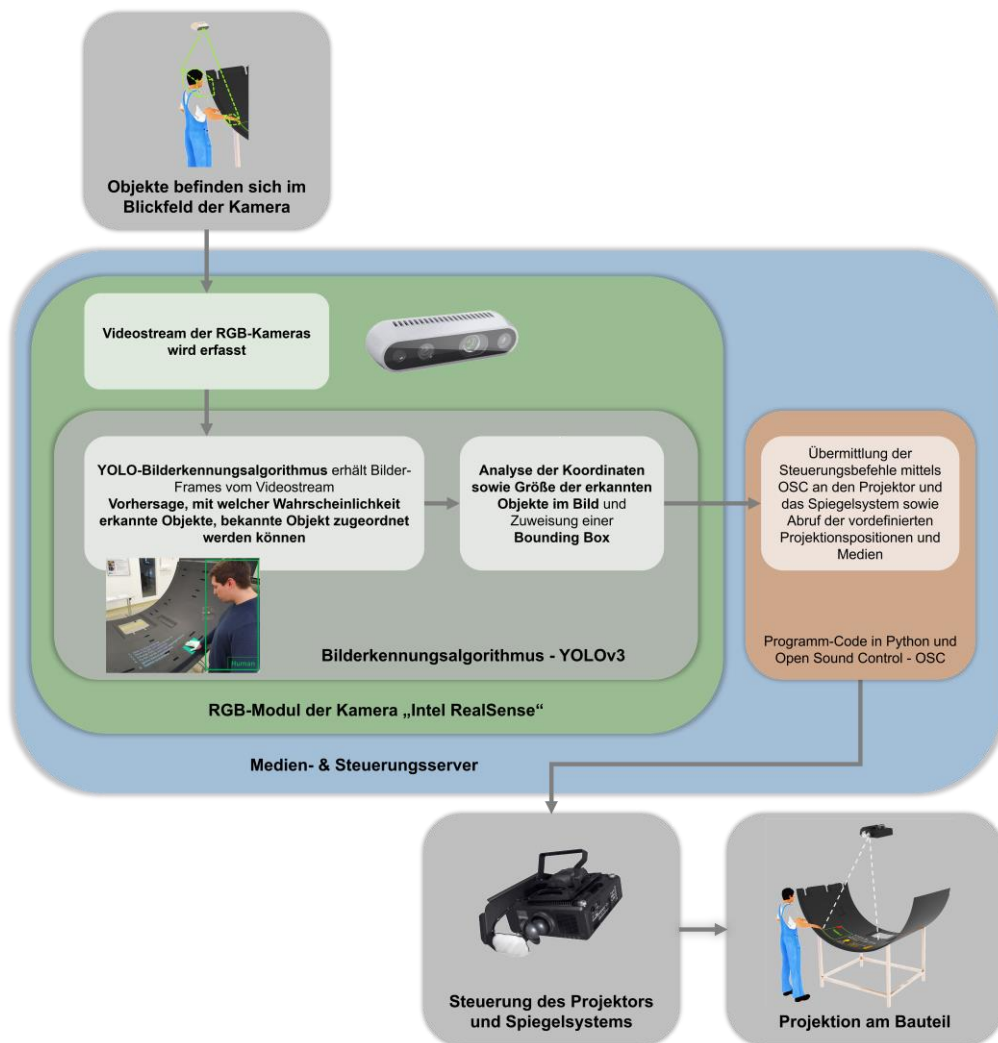


Abbildung 56 | Bilderkennung und Steuerung der dynamischen Projektion (Eigendarstellung)

¹³ Für die praktische Anwendung und zum Vergleich der Objekterkennung in den einzelnen Experimenten wurde in dieser Arbeit ausschließlich der Algorithmus YOLOv3 verwendet. Es bestehen jedoch Intentionen alternative Algorithmen, auch YOLOv4, in zukünftigen Forschungen zu testen und anzuwenden.

Zur Bereitstellung des hochauflösenden RGB-Streams wird die Kamera „Intel RealSense D435i“ verwendet, welche über der Montagestation platziert wird und die Objekte (Menschen, Gesten etc.) im Arbeitsraum erfasst. Zur Anbindung des Videostreams für die Verwendung mit dem YOLO-Algorithmus wird die pythonbasierte „librealsense“-Bibliothek von Intel herangezogen (<https://github.com/IntelRealSense/librealsense>). Für die optimale Objekterkennung im speziellen Anwendungsfall der industriellen Baustellenmontage wird der vorab trainierte Datensatz in den pythonbasierten YOLO-Programmcode „pytorch-yolov3“ (<https://github.com/ayoozhkathuria/pytorch-yolo-v3>) übertragen, um ihn für die Echtzeiterkennung im Montageprozess verwenden zu können. Die erfassten Koordinaten der erkannten Objekte bzw. der Bounding-Boxen werden in Echtzeit in den pythonbasierten Programmcode integriert, um damit anschließend die softwarebasierten Befehle auszulösen und den Projektor sowie das Spiegelsystem anzusteuern. Für die Übertragung der Daten an den Projektor und Spiegelkopf wird das pythonbasierte Modul Open Sound Control (OSC) verwendet (<https://github.com/Dynamic/ProjectionInstitute/MDCXexamples/blob/master/OSC.py>). Das OSC-Protokoll bietet die Möglichkeit der Kommunikation und des Datenaustausches mit IP-fähigen Geräten, wie dem Medien- und Steuerungsserver auf dem vorhandenen Ubuntu-Desktop-PC, dem Projektor und dem Spiegelkopf. Dazu werden die vorab definierten Arbeitsinformationen inklusive deren Projektionsposition und Orientierung sowie Verzerrungen abgerufen und mittels des Projektionssystems an der definierten Montageposition in der Baustellenmontage angezeigt. Mithilfe der Positionen des Menschen sowie der Posen und Gesten können damit die Arbeitsinformationen direkt in der Montagestation gesteuert werden, um z. B. von einer Information zur nächsten zu gelangen und so Schritt für Schritt durch den Prozess zu steuern (Rupprecht et al., 2020).

c) Synthetisch erzeugter Gestendatensatz

Aufgrund der Spezifika der Großgeräte- und Baustellenmontage (vgl. Kapitel 1.2.1), der verschiedenen Oberflächen und Farben der Montageobjekte bzw. des Hallenbodens und vor allem wegen des großen Abstandes zwischen Kamera und der zu erkennenden Objekte aus der Vogelperspektive (z. B. Menschen oder Handgesten von oben) (Abbildung 57) bedarf es, einen für das Setting spezifischen Gestendatensatz zu gestalten (Rupprecht et al., 2020; Rupprecht et al., 2021).

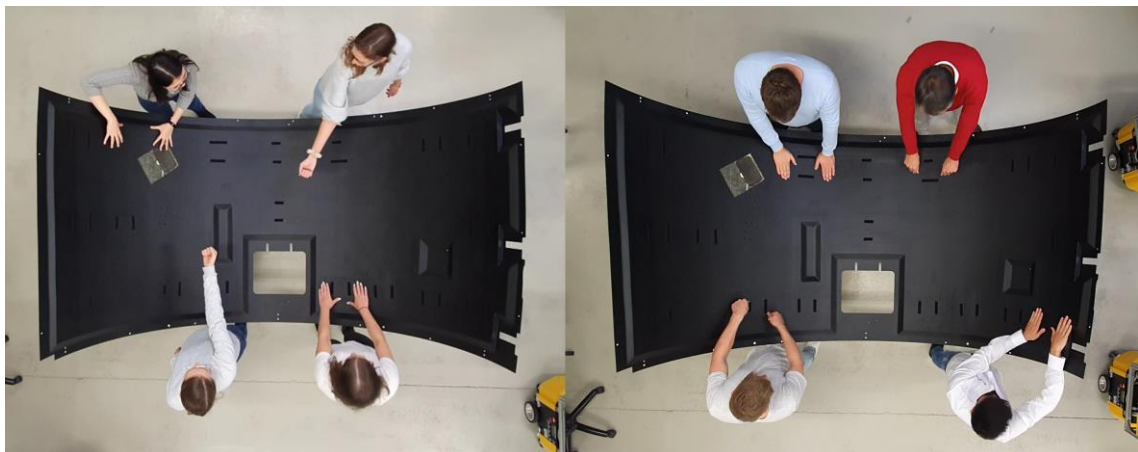


Abbildung 57 | Setting Baustellenmontage – Gestenerkennung von oben (Eigendarstellung)

Zur Erstellung des Datensatzes, speziell für die Großgeräte- und Baustellenmontage wurden Trainingsdaten im realen Setting der Montagestation aufgenommen. Dabei wurden in der ersten Iteration (Experiment 2, vgl. Kapitel 5.2.2) die Gestendaten manuell gelabelt und trainiert und im Experiment 3 (vgl. Kapitel 5.2.3) mit einem automatisierten Prozess erstellt („Synthetisch erzeugter Gestendatensatz“¹⁴). Für das Trainieren des Datensatzes standen Hochleistungsrechner des „Vienna Scientific Clusters“ (VSC; <https://vsc.ac.at>) zur Verfügung, welche auf leistungsfähige Grafikkarten der Marke „Nvidia GeForce GTX-1080“ zurückgreifen (Rupprecht et al., 2021). Der Datensatz wurde mit mehreren Instanzen des Hochleistungsrechners trainiert und der Objekterkennung zur Verfügung gestellt.

Tabelle 14 stellt eine Zusammenfassung der einzelnen Komponenten des dynamischen Projektionssystems und visuellen Interaktionssystems dar und gibt einen Überblick über die Preise/Kosten, die Nutzungsdauer sowie die Abschreibung (AfA). Des Weiteren werden die Gesamtanschaffungskosten des Spatial Augmented Reality Assistenzsystems dargestellt.

Tabelle 14 | Komponenten, Preise und Gesamtanschaffungskosten des SAR-Systems

Komponenten des SAR-Systems	Preise/Kosten	Nutzungsdauer	Afa
Dynamisches Projektionssystem	15 050,10 Euro		
Projektor	9 299,60 Euro	8 Jahre ¹⁵	1 162,00 Euro
Spiegelkopfsystem	2 295,00 Euro	8 Jahre	287,00 Euro
Deckenhalterung für den Projektor	187,00 Euro	Geringwertiges Wirtschaftsgut (GWG) ¹⁶	0,00 Euro
Medien- und Projektorsteuerung (Desktop-PC inklusive Lizenzen)	2 370,50 Euro	3 Jahre	790,00 Euro
ZED-2-Kamera	398,00 Euro	Geringwertiges Wirtschaftsgut (GWG)	0,00 Euro
Monitor und Kleinteile	500 Euro	Geringwertiges Wirtschaftsgut (GWG)	0,00 Euro
Visuelles Interaktionssystem	250 Euro		
Real Sense Kamera inkl. Kleinteile	250 Euro	Geringwertiges Wirtschaftsgut (GWG)	250 Euro
Gesamtanschaffungskosten (Dynamisches Projektionssystem + Visuelles Interaktionssystem)	15 300,10 Euro		
Afa + GWG Jahr 1	3 574,00 Euro		
Afa Jahr 2 bis 3	2 239,00 Euro		
Afa Jahr 4 bis 8	1 449,00 Euro		

Zusammenfassend ergibt sich aus den Summen der Preise der Einzelkomponenten ein Gesamtanschaffungswert des Spatial Augmented Reality Assistenzsystem von 15 300,10 Euro. Durch die Möglichkeit der Abschreibung einzelner Bauteile ergibt sich eine AfA von 3574,00 Euro im ersten Jahr, von 2239,00 Euro im zweiten und dritten Jahr und von

¹⁴ Details zum „Synthetisch erzeugter Datensatz“ finden sich in Kapitel 3.3.3.

¹⁵ Laut AfA-Tabelle des deutschen Bundesministeriums der Finanzen; vgl. Bundesministerium der Finanzen (2021)

¹⁶ Eine Anschaffung bis zu einem Wert von 800 Euro netto kann mit Beginn der Nutzungsdauer sofort abgesetzt werden; vgl. <https://www.usp.gv.at/steuern-finanzen/betriebseinnahmen-und-ausgaben/geringwertige-wirtschaftsgueter.html>

1449,00 Euro in den Jahren 4 bis 8. Der Aufbau stellt demnach, im Vergleich zu herkömmlichen Aufbauten (vgl. Kapitel 3.2.1), eine kostengünstige Variante zur Informationsbereitstellung in der Großgerätemontage dar.

3.3 Gestaltung der dynamischen Informationsbereitstellung und Mensch-Maschine-Interaktion

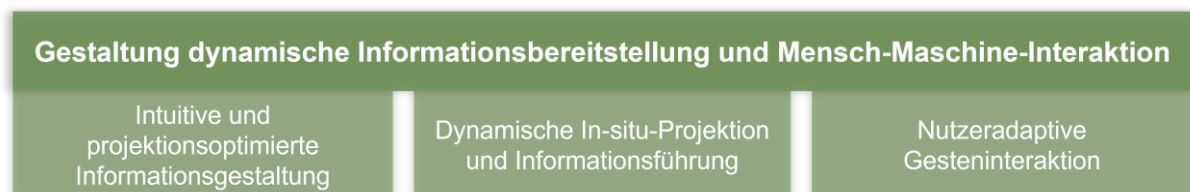


Abbildung 58 | Gestaltungsaspekte der dynamischen Informationsbereitstellung und Mensch-Maschine-Interaktion (Eigendarstellung)

Im vorangegangenen Kapitel wurde die Umsetzung der Informationsausgabe (dynamisches Projektionssystem) und Informationseingabe (visuelles Interaktionssystem) sowie das Zusammenspiel der einzelnen Hard- und Softwarekomponenten aus Sicht der technischen Gestaltung der Mensch-Maschine-Interaktion beschrieben.

Im Folgenden wird auf die **Gestaltung der dynamischen Informationsbereitstellung und Mensch-Maschine-Interaktion** (vgl. Abbildung 58) und vor allem auf die Informationsaufnahme und -abgabe aus Sicht des Menschen eingegangen (vgl. Kapitel 2.4, Abbildung 35). Demnach stellen aus Sicht des Users die Auswahl und Gestaltung der Gesten sowie das Training der Gesteninteraktion den Kern der Informationsabgabe dar. Die intuitive Aufbereitung und Bereitstellung der Arbeitsinformationen in situ am Ort des Arbeitsgeschehens sowie die Führung durch den Montageprozess stellen den Fokus der Informationsaufnahme dar.

Bei der **intuitiven und projektionsoptimierten Informationsgestaltung** wird auf die Art der Visualisierung und Strukturierung der Arbeitsinformationen eingegangen (vgl. Kapitel 2.3.2). Dabei wird vor allem auf die intuitive Verständlichkeit sowie die projektionsgerechte Darstellung geachtet, um eine geeignete kognitive Verarbeitung zu ermöglichen und somit den Menschen bei der effektiven und effizienten Ausführung der Montagetätigkeiten zu unterstützen.

Bei der **dynamischen In-situ-Projektion und Informationsführung** wird vor allem auf die In-situ-Gestaltung der Positionierung, Orientierung und Lage der Arbeitsinformationen im Raum eingegangen. Außerdem wird die Führung mittels dynamischer Projektion der Informationen bzw. des Projektionsstrahls gestaltet, um damit den Menschen gezielt durch den Montageprozess zu leiten. Dazu konnten geeignete Gestaltungsempfehlungen aus geltenden Normen und wissenschaftlichen Publikationen im Bereich der Montage recherchiert werden, welche nach den Anforderungen und Gegebenheiten der Groß- und Baustellenmontage adaptiert wurden (vgl. Kapitel 2.3.2 und 2.3.3). Die folgenden Gestaltungen der Arbeitsanweisungen wurden im iterativen Designprozess in diversen Experimenten am Demonstrator getestet, evaluiert und stetig verbessert.

3.3.1 Intuitive und projektionsoptimierte Informationsgestaltung

Wie in Kapitel 2.3.2 dargestellt, werden für standardisierte Montageanweisungen ein strukturierter Aufbau sowie eine verständliche Formulierung empfohlen. Des Weiteren sollen die Kerntätigkeiten der Montage, wie „Entnahme und Platzierung“ der Bauteile und „Montage der Bauteile“ integriert werden. Für die verständlichen Formulierungen eignen sich besonders kurze Texte, welche gemeinsam mit intuitiver Symbolik wie Piktogrammen, Konturen oder intuitiven Zeichen umgesetzt werden. Weitere Empfehlungen sind, dass die visuellen Benutzerschnittstellen klar erkennbar zu gestalten sind und jeweils an derselben Stelle der Arbeitsinformation platziert werden sollen. Für eine intuitive Verständlichkeit sollen internationale und kulturell konforme Zeichen und Farben verwendet und die Dichte der Information so gewählt werden, dass die Anweisung nicht überladen wirkt. Für die projektionsgerechte Darstellung sollen vereinfachte Schriftarten und optimale Schriftgrößen gewählt und kontrastreiche Farben im Vergleich zum vorliegenden Projektionshintergrund verwendet werden. Die im Folgenden präsentierten Arbeitsanweisungen wurden anhand dieser Gestaltungsempfehlungen aufbereitet und für die Verwendung am Demonstrator in der Baustellenmontage bereitgestellt. Aus den Evaluierungsexperimenten erhaltenes Feedback wurde wiederum in die Gestaltung eingearbeitet und die Informationsgestaltung damit verbessert.

Abbildung 59 stellt eine initiale Arbeitsanweisung für die Verwendung mittels Projektionssystem dar, welche die Anzeige von Arbeitstätigkeiten, Interaktionsmöglichkeiten und Positionen beinhaltet. Die Arbeitsinformationen sind für die Montage von Carbon-Matten (CFK-Layup) für die Fan-Cowl- bzw. Winglet-Montage¹⁷ (vgl. Kapitel 4.1) gestaltet.

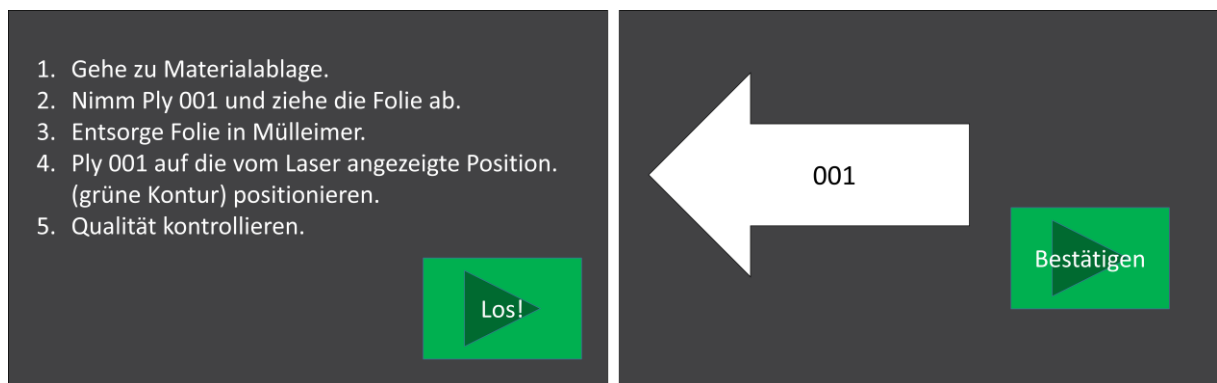


Abbildung 59 | Erste Arbeitsanweisung mit „Interaktive Buttons“. Links: Anzeige der Montagetätigkeiten, rechts: Anzeige der Montageposition mittels Symbolik (Eigendarstellung)

Die Arbeitsanweisung links zeigt fünf Montagetätigkeiten sowie einen Button für die Startinteraktion des Montageprozesses. Die Interaktion (Berühren des Buttons) wurde in diesem Use Case als Wizard-of-Oz integriert¹⁸. Die Anweisung teilt dem Menschen mit, dass

¹⁷ Anmerkung: Ein „Winglet“ ist ein Flugzeugbauteil am Ende der Tragfläche zur Reduktion des Treibstoffverbrauchs. Ein „Fan-Cowl“ bezeichnet ein Flugzeugturbinengehäuse, welches aus zwei Halbschalen besteht.

¹⁸ Wizard-of-Oz bedeutet nach Weiss et al. (2009), dass die Interaktionsfunktion nur prototypenhaft vorliegt und technisch nicht funktionsfähig ist. Der Proband wird jedoch im Glauben gelassen, dass er/sie mit einem funktionsfähigen Feature interagiert, der Wizard (Mensch) simuliert jedoch manuell die Interaktion im Hintergrund

die Carbon-Matte mit der Bezeichnung „Ply 001“ von der Materialablage entnommen, die Folie im Mülleimer entsorgt und anschließend auf der angezeigten Position montiert werden soll. Anschließend soll die Qualität der durchgeführten Montage überprüft werden. Die Arbeitsanweisung rechts zeigt, wo die exakte Montageposition zu finden ist, die exakte Kontur der Position wurde in diesem Fall mit einem zusätzlichen statischen Laser umgesetzt. Die „Interaktiven Buttons“ werden in Form von grünen Rechtecken und kurzen Textanweisungen umgesetzt und dienen dazu, um zur nächsten Arbeitsanweisung zu gelangen bzw. den Montageschritt zu bestätigen. Die Arbeitsanweisung wurde so gestaltet, dass ein strukturierter Aufbau sowie prägnante Formulierungen verwendet werden und die Kerntätigkeiten der Montage für Entnehmen (2) und Platzieren (4) sowie Montage der Bauteile (4 und 5) integriert sind. Es wurde darauf geachtet, dass Symbol- und Schriftfarben mit hohem Kontrast zur Projektionsfläche verwendet werden. Dabei wurde Reinweiß (RGB 255 255 255) und ein Standardgrün (RGB 0 176 80) aus der Farbpalette der Erstellungssoftware der Informationen verwendet. Spezielle Tests mit verschiedenen Farben wurden hier noch nicht durchgeführt. Das Weiß wird dabei für die Anzeige von neutralen Informationen und das Grün für die Anzeige der Aktionen verwendet. Ebenso wurde der Informations- und Aktionsteil voneinander getrennt dargestellt, um so eine gute Erkennbarkeit zu ermöglichen. Die Buttons zum Starten und Bestätigen wurden immer an derselben Stelle der Arbeitsanweisungen platziert, um die Gestaltungsempfehlung „Konsistenz“ sowie „schnelle Erkennung“ umzusetzen. In einem späteren Entwicklungsstand des Spatial Augmented Reality Assistenzsystems wurden die Interaktionen nicht mehr über interaktive Buttons umgesetzt, sondern eine funktionsfähige Gestensteuerung verwendet. Dabei wurde die Komplexität der Montagetätigkeiten erhöht, um ein realitätsnahes Setting umzusetzen. Dementsprechend wurde besonders auf die intuitive Verständlichkeit der Arbeits- und Gesteninformationen geachtet (Abbildung 60).

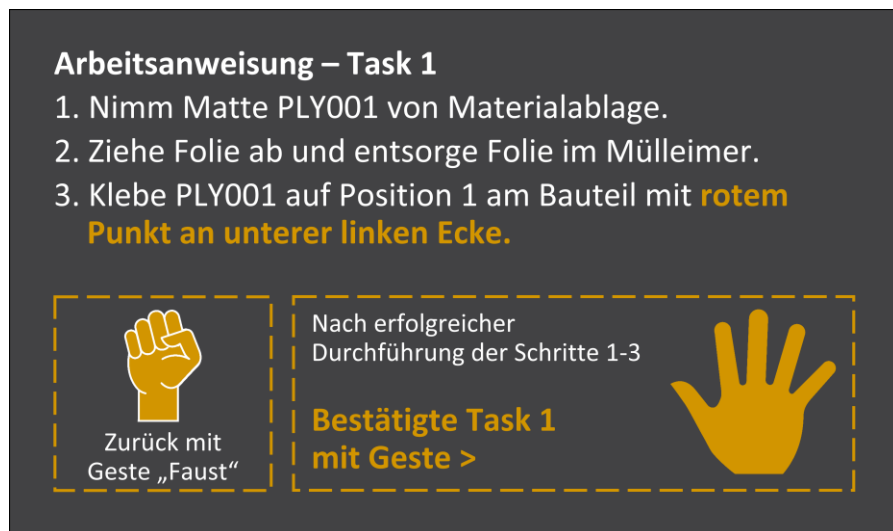


Abbildung 60 | Komplexere Arbeitsanweisung mit Information über die Gesteninteraktion
(Eigendarstellung)

Die „Arbeitsanweisung – Task 1“ stellt drei Montagetätigkeiten sowie die Informationen für die möglichen Gesteninteraktionen dar. Die Information gibt an, dass eine Carbon-Matte mit der Bezeichnung „PLY001“ aus der Materialablage zu entnehmen ist, die Folie im Mülleimer

entsorgt und diese auf die Position 1 mit angegebener Orientierung montiert werden soll. Im Vergleich zu Abbildung 59 wurden die Tätigkeiten 1 bis 3 hier in den Punkten 1 und 2 zusammengefasst und klarer strukturiert, woraus eine kompaktere Darstellung resultiert. Hierbei wurden die Empfehlungen wie oben erwähnt umgesetzt und zusätzlich besonders relevante Arbeitsinformationen in Form von farblichen Texten (orange) hervorgehoben. Für die Information über die Gesteninteraktion wurden Gestenpiktogramme, welche den echten Gesten ähneln, verwendet sowie kurze Kommandotexte integriert, um eine intuitive Gestaltung umzusetzen. Um den Informationsteil und Aktionsteil noch optimaler zu erkennen, wurden die Teile durch eine gestrichelte Linie getrennt. Ebenso wurden für die projektionsgerechte Aufbereitung kräftige Farben mit hohem Kontrast herangezogen, welche aus diversen Tests am Demonstrator ermittelt wurden (vgl. Kapitel 4). Dabei wurde für die Anzeige von neutralen Informationen ein Reinweiß (RGB 255 255 255) verwendet und besonders relevante Informationen im kräftigem Gold-Orange (RGB 210 149 0) umgesetzt. Für die Anzeige der Gestenaktionen wurde hier ebenso Orange anstatt das Standardgrün (vgl. Abbildung 59) verwendet, um damit noch nicht abgeschlossene Aktionen zu charakterisieren. Eine weitere Anforderung bei der Umsetzung der Informationsgestaltung bei der Interaktion mit Gesten war, dass dem Menschen zu jedem Zeitpunkt bekannt sein sollte, wie der Status des Systems bzw. der Gesteninteraktion ist. Demnach wurde nach jeder Arbeitsanweisung eine Feedbackinformation angezeigt, um der Person mitzuteilen, dass der Montagetausk erfolgreich durch die Geste bestätigt und zugleich die Interaktion mit der Geste erfolgreich ausgeführt wurde (Abbildung 61).



Abbildung 61 | Feedbackinformation und Bestätigung der erfolgreichen Task-/Gestenausführung
(Eigendarstellung)

Die Feedbackinformation wurde direkt nach erfolgreicher Ausführung der Geste projiziert und so gestaltet, dass mit prägnantem Text sowie dem Einsatz eines Piktogramms angezeigt wird, dass der vorangegangene Montagetausk mit der Geste bestätigt wurde. Dabei symbolisiert bzw. unterstreicht der integrierte Haken im Piktogramm sowie die Verwendung von grüner Farbe die erfolgreiche und abgeschlossene Durchführung der Aktion. Als Farbgestaltung wurde hierbei statt einem Standardgrün ein kräftigeres Grün (RGB 36 145 0) verwendet. Bei den Evaluierungsexperimenten (vgl. Kapitel 5) konnte aus dem qualitativen Feedback der Teilnehmer*innen ermittelt werden, dass weniger Text in den Arbeitsanweisungen, speziell bei der Gesteninformationen, verwendet werden soll. Dafür wurden folgende Vorschläge zur

Verbesserung der intuitiven Gestaltung der Anzeige der Gesteinsinformation präsentiert (Abbildung 62).



Abbildung 62 | Verbesserung der intuitiven Gestaltung der Gesteninformation (Eigendarstellung)

Dazu wurden die verbesserten Gesten mit gängigen und intuitiv verständlichen Zusatzsymbolen ausgestattet, welche stellvertretend für die textlichen Informationen „Bestätigen“, „Starten“ und „Weiter“ stehen und diese und weitere textliche Informationen dementsprechend ersetzen sollen. Aus dem Feedback konnte festgestellt werden, dass nach den ersten durchgeführten Tasks der Informationsinhalt bereits verstanden wurde und darauffolgende gleiche Informationen mit geringerem Detailgrad angezeigt werden sollten. Dementsprechend wurde folgender Vorschlag für die Verringerung der detaillierten Informationen nach zeitlichem Montagefortschritt integriert (Abbildung 63).

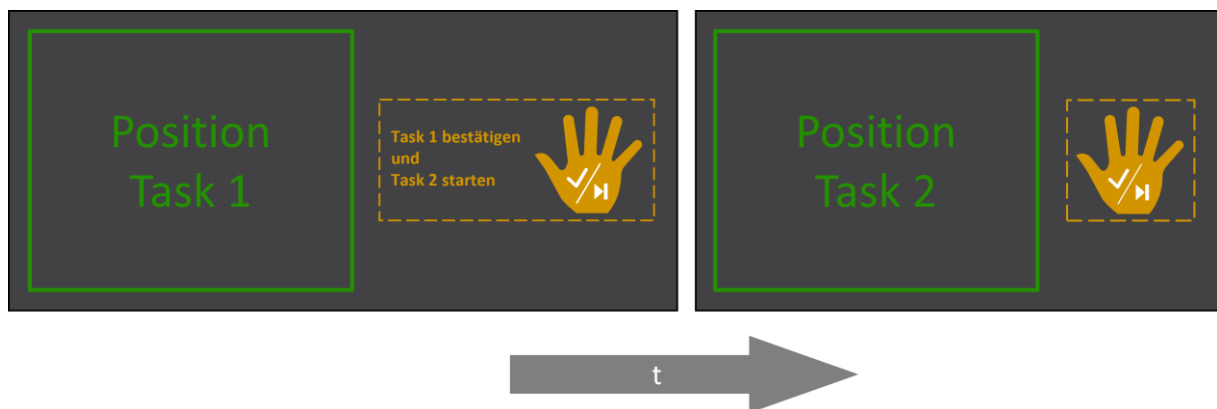


Abbildung 63 | Verringerung der Informationsdetails mit zeitlichem Montagefortschritt (Eigendarstellung)

Dabei wurde die Information zur Gesteninteraktion anfangs im Montageprozess noch mit textlichen Informationen und einem Piktogramm mit Zusatzsymbolen angezeigt und im weiteren Zeitverlauf der Montagetätigkeiten vereinfacht nur mehr als Piktogramm. Daraus leitet sich auch ab, dass ebenso wie in der Langzeitstudie nach Funk et al. (2017) beschrieben wird, speziell erfahrene Werker*innen oder diejenigen, die den Prozess gut kennen, dementsprechend weniger Informationen für die Umsetzung der Montagetätigkeiten benötigen. So wäre es für erfahrene Werker*innen bzw. auch für ähnliche und wiederholende Informationen denkbar, dass die Arbeitsinformationen so intuitiv wie möglich angezeigt

werden. Jedoch sollte stets darauf geachtet werden, dass besonders relevante und neue Informationen oder Änderungen von bekannten Anweisungen markiert und hervorgehoben werden, um Fehler zu vermeiden und Tasks nicht zu übersehen (vgl. Abbildung 60).

Zusammengefasst wurden die Arbeitsanweisungen für das Spatial Augmented Reality Assistenzsystem in der industriellen Großgeräte- und Baustellenmontage nach den geltenden Gestaltungsempfehlungen sowie dem Feedback aus den Experimenten intuitiv aufbereitet und somit für die effiziente und effektive Unterstützung zur Umsetzung der Montagetätigkeiten bereitgestellt.

3.3.2 Dynamische In-situ-Projektion und Informationsführung

Wie in Kapitel 2.3.3 beschrieben, stellt die Informationsbereitstellung mit dynamischer In-situ-Projektion und dynamischer Informationsführung neben der intuitiven und projektionsoptimierten Informationsgestaltung (vgl. Kapitel 3.3.1) relevante Aspekte bei Verwendung von Spatial Augmented Reality in der Großgeräte- und Baustellenmontage dar. Dabei konnte aus der State-of-the-Art-Analyse sowie dem Feedback der Demonstration und Evaluierung (vgl. Kapitel 4 und 5) festgestellt werden, dass sich die Informationsbereitstellung mit **dynamischer In-situ-Projektion** (Anzeige der Information an der richtigen Montageposition, Lage und Orientierung) und **dynamischer Informationsführung**, also die Führung „step by step“ von Arbeitsanweisung zu Arbeitsanweisung, zur Entnahmestelle oder zur exakten Montageposition, besonders für den effektiven Einsatz in der Großgeräte- und Baustellenmontage eignet. Dementsprechend wird folgend auf die Gestaltungsvarianten der In-situ-Projektion sowie auf die Umsetzung der dynamischen Informationsführung durch den Montageprozess eingegangen.

Das Ziel dabei ist, dass die Anzeige der Arbeitsinformationen sowie der Montageposition und Orientierung allein mithilfe des dynamischen Projektionssystems umgesetzt wird und dementsprechend auf die Verwendung anderer Systeme (z. B. statische Lasersysteme oder PC-Terminals) verzichtet wird. Für die technische Umsetzung wird das dynamische Projektionssystem (vgl. Kapitel 3.2.1) eingesetzt, um damit Arbeitsinformationen an den jeweiligen Positionen anzuzeigen und den Menschen dynamisch durch den Montageprozess zu führen. Für die Gestaltung der In-situ-Projektion und Informationsführung sind folgende Fragestellungen zu beantworten und dementsprechende Gestaltungsmerkmale zu definieren.

- Wer oder was steuert den Prozess bzw. die In-situ-Projektion und Informationsführung?
 - a) Mensch durch Interaktion
 - b) System durch prozessintegrierte Steuerung
 - c) Kombination durch Interaktion und prozessintegrierte Steuerung
- Was ist die Zielposition? Wohin wird geführt?
 - a) Zur (nächsten) Arbeitsanweisung/Information
 - b) Zur (nächsten) Montageposition
 - c) Zur (nächsten) Bereitstellungsfläche
 - d) Zum (nächsten) Montageobjekt

- Wie exakt soll die Zielposition erreicht werden?
 - a) Ins Nahefeld der Zielposition (in view)
 - b) Exakt zur Zielposition (Teil – in situ)
 - c) Exakt zur Zielposition inklusive Orientierung und Lage (in situ)
- Wie soll die Führung durch Information stattfinden?
 - a) Durch rein textliche/grafische Information (ohne dynamische Führung)
 - b) Durch In-situ-Projektion mit exakter Anzeige der Positionen ohne dynamische Führung zur Position
 - c) Durch In-situ-Projektion mit exakter Anzeige der Positionen mit dynamischer Führung zur Position

Demnach entstehen bei der Gestaltung der Informationsbereitstellung verschiedene Möglichkeiten der Umsetzung der In-situ-Projektion sowie der Informationsführung. Bei allen Varianten wird eine konkrete Abfolge an Arbeitstätigkeiten im System definiert, welche den Menschen durch den Montageprozess (von Position zu Position oder von Arbeitsanweisung zu Arbeitsanweisung) führt. Der Mensch bestimmt jedoch die Geschwindigkeit der Ausführung der einzelnen Montagetätigkeiten und steuert durch die Gesteninteraktion das System. Im interaktiven Designprozess wurden verschiedene Varianten der Gestaltungsmerkmale anhand der Experimente und Use Cases demonstriert und verglichen.

1. **Experiment 1 – Fan-Cowl-Montage:** In-view-Projektion von textlicher/grafischer Arbeitsinformation grob im Nahefeld der exakten Montageposition, In-situ-Projektion mit statischem Lasersystem (vgl. Kapitel 5.2.1 und Abbildung 64).

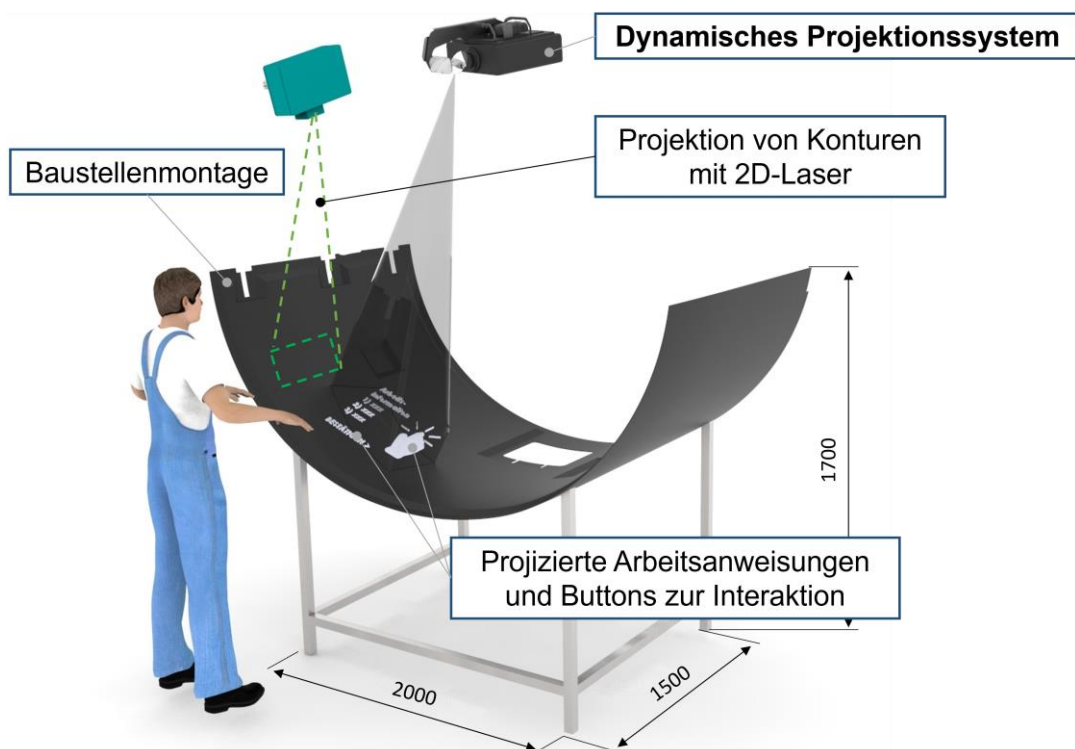


Abbildung 64 | In-view-Projektion mit Projektionssystem und In-situ-Projektion mit statischem Lasersystem, in Anlehnung an Rupprecht et al. (2020)

Im Experiment 1 mit dem Use Case „Fan-Cowl-Montage“ wird die textliche/grafische Arbeitsinformation mit dem dynamischen Projektionssystem als In-view-Projektion grob im Nahfeld der exakten Montageposition projiziert. Die exakte Montageposition wird in situ mit einem statischen Lasersystem umgesetzt (vgl. Kapitel 2.3.1).

2. Mini-Use-Case – Montage von Nutensteinen: In-situ-Projektion mit exakter Anzeige der Positionen ohne dynamische Führung zur Position (Abbildung 65).

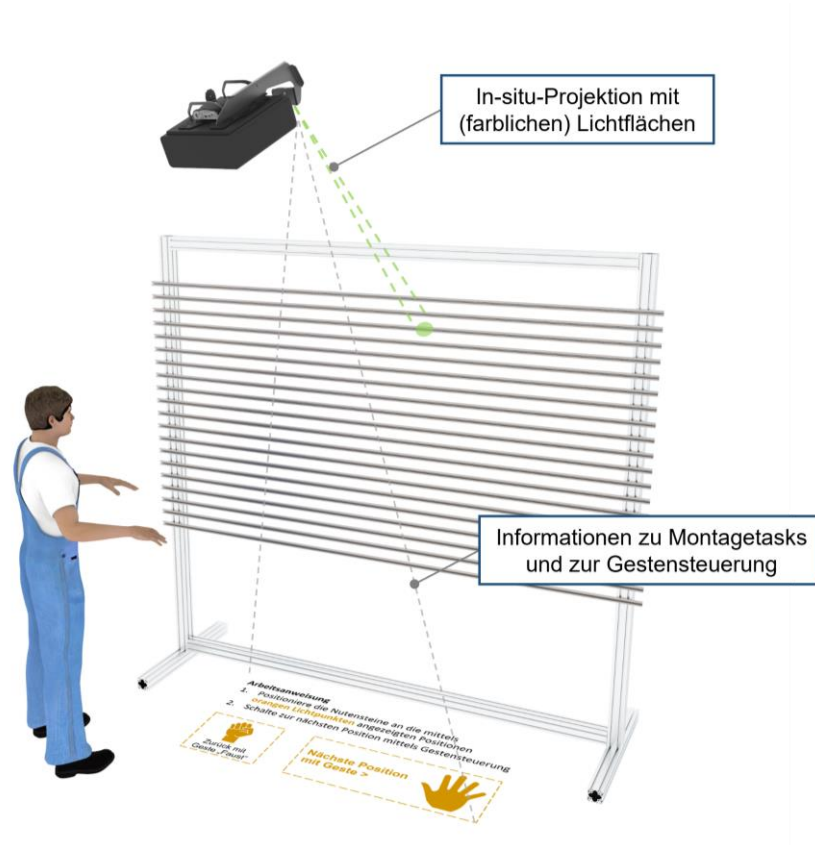


Abbildung 65 | In-situ-Projektion mit farbigen Lichtpunkten (Eigendarstellung)

Eine weitere Gestaltungsvariante der In-situ-Projektion stellt die Anzeige mittels **farblicher Lichtpunkte bzw. Hervorhebung durch Lichtflächen** dar, welche in einem Mini-Use-Case der Montage von Nutensteinen umgesetzt wurde. Dabei war die Anforderung, die Nutensteine exakt an bestimmten Positionen in den vorhandenen C-Profilen zu platzieren. Dabei wurden die Montagepositionen vorab in der Software definiert und die exakten Positionen der Nutensteine in Form von grünen Lichtpunkten an den Profilen angezeigt. Nach Beendigung aller Montagetaetigkeiten wurde die Montagestation mit einem grünen Licht hervorgehoben, was die erfolgreiche Durchfuehrung signalisiert. Zusätzliche Informationen zum Montageprozess sowie Informationen zur Interaktion wurden mittels Projektion auf den Hallenboden umgesetzt.

3. Experiment 2 – Fan-Cowl-Montage: Dynamische In-situ-Projektion mit Informationsführung zur Montageposition (vgl. Kapitel 5.2.2; Abbildung 66).

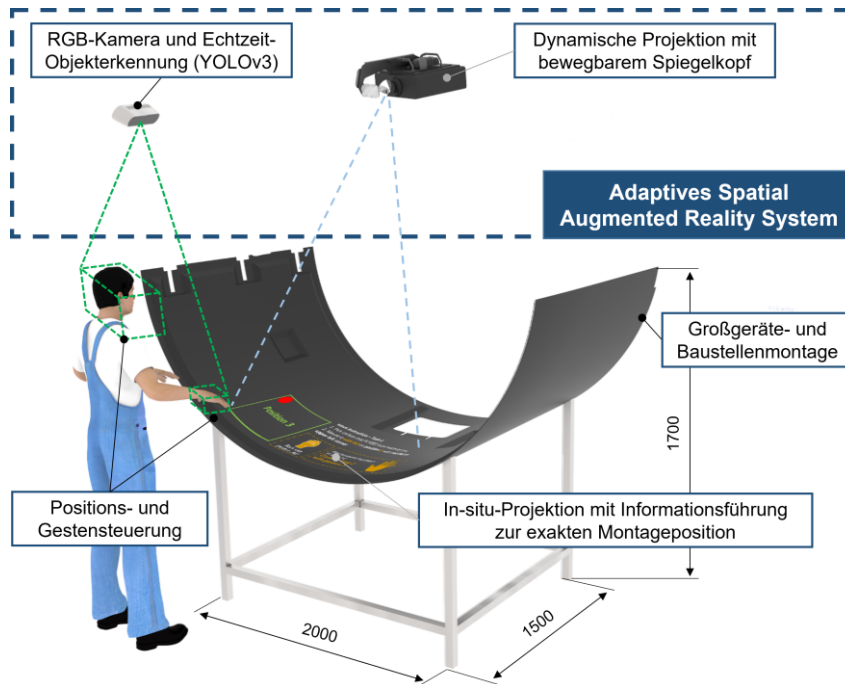


Abbildung 66 | In-situ-Projektion mit dynamischer Informationsführung zur Montageposition (Eigendarstellung)

Für das Experiment 2 und den Use Case „Fan-Cowl-Montage“ wurde als Baseline-Variante die textliche/grafische Arbeitsinformation (vgl. Abbildung 60) als In-view-Projektion grob im Nahfeld der exakten Montageposition angezeigt. Die exakten Montagepositionen sind ausschließlich über die textliche Anweisung am Bauteil zu ermitteln. Bei der Variante mit In-situ-Projektion mit dynamischer Informationsführung zur Montageposition wird einerseits die exakte Montageposition mittels Projektionssystem angezeigt und gleichzeitig führt der Projektionsstrahl dynamisch von Montageposition zu Montageposition. Für die Gestaltung der In-situ-Projektion wurde die textliche Arbeitsanweisung (vgl. Abbildung 60) herangezogen und durch zusätzliche Information ergänzt (Abbildung 67).

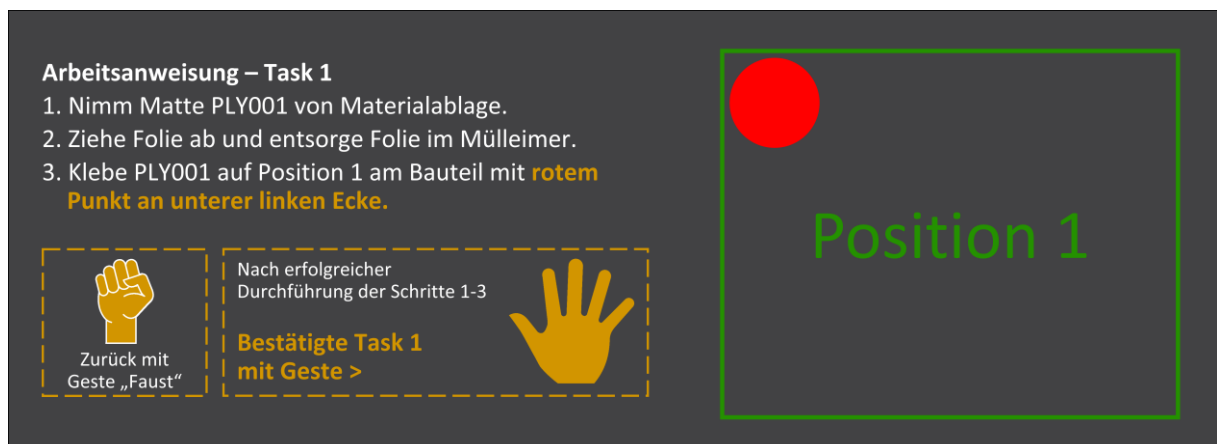


Abbildung 67 | Anzeige der Montageposition und -orientierung (Eigendarstellung)

Dabei werden die Montagetasks und die Informationen für die Gesteninteraktion wiederum als Kombination aus textlichen und grafischen Inhalten gestaltet und die Anzeige der Montageposition und Orientierung als grafische Zusatzinformation integriert. Die Anzeige der exakten Montageposition wird dabei mittels grüner Kontur, welche der Grundform des Montageobjektes entspricht, umgesetzt und die Orientierung des Montageobjektes wird mit einem roten Punkt angegeben. Zur schnellen Auffindung und Vermeidung von Fehlern wird zusätzlich noch die Positionsnummer als kurzer Text in die Kontur integriert. Durch die Zusatzinformation wird die Montageposition nicht nur textlich als Information präsentiert, sondern zusätzlich grafisch angezeigt, was die schnelle Auffindung und Erkennbarkeit sowie die kognitive Verarbeitung des Menschen fördert. Diese Arbeitsinformation wird anschließend dynamisch und in situ an der richtigen Position an der Montagestation projiziert und der Mensch schaltet durch die Interaktion mit dem Informationssystem von einer In-situ-Projektion zur nächsten.

4. Experiment 3 – Spielzeugbaustein-Montage: In-situ-Projektion mit dynamischer Informationsführung zur Entnahme- und Montageposition in Kombination mit textlichen/grafischen Anweisungen (vgl. Kapitel 5.2.3; Abbildung 68).

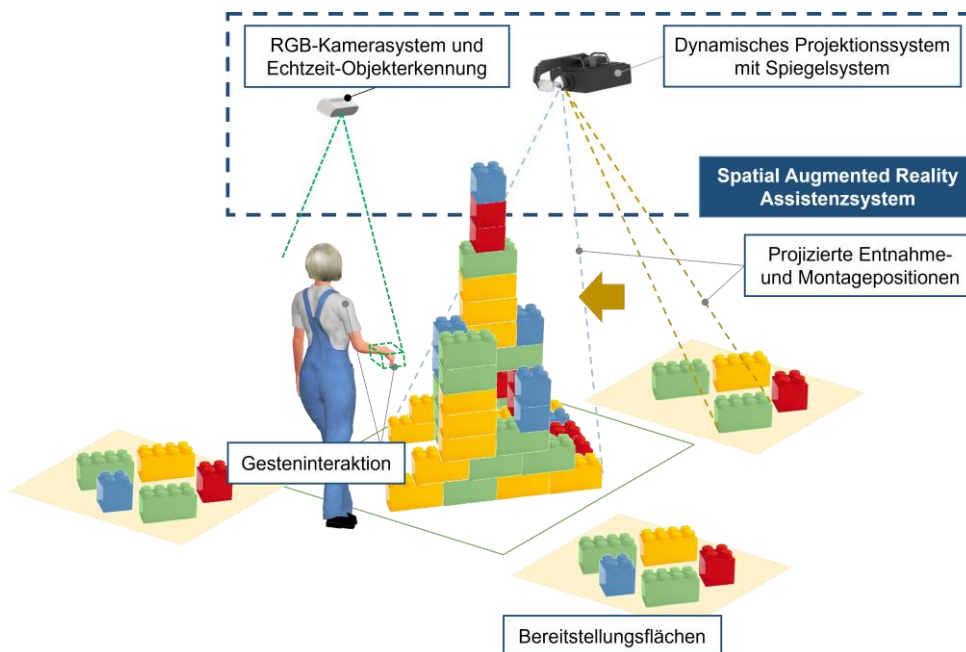


Abbildung 68 | Experiment 3: In-situ-Projektion mit dynamischer Informationsführung (Eigendarstellung)

Die In-situ-Projektion mit dynamischer Informationsführung zur Entnahme- und Montageposition wird im Use Case der Spielzeugbaustein-Montage umgesetzt. Dabei wird mittels dynamischer Animation und Projektionsstrahl zur jeweiligen Bereitstellungsfläche geführt und die exakte Entnahmeposition der Bauteile angezeigt. Mithilfe farbiger oder weißer Lichtflächen wird die exakte Montageposition am Hallenboden oder direkt auf dem Montageobjekt projiziert. Weitere Informationen zu Montagetasks und der Interaktion werden einzeln (Single-Point-Projektion) oder in Kombination mit den Lichtflächen (Multi-Point-Projektion) projiziert (Abbildung 69).

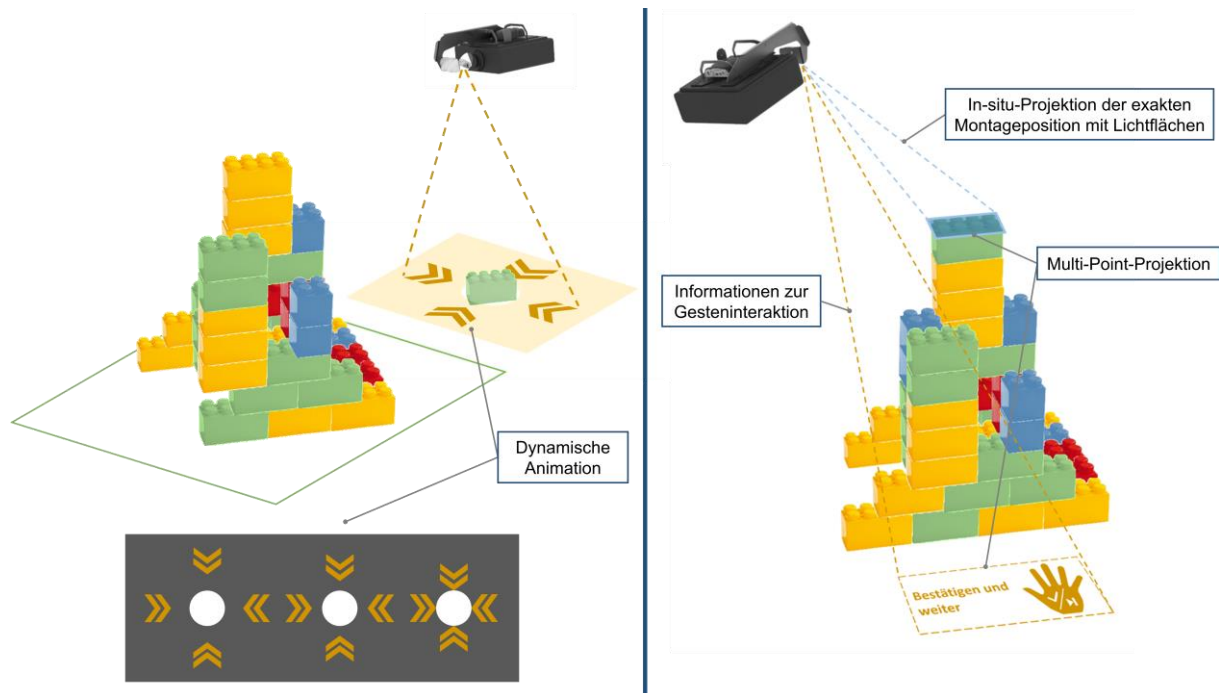


Abbildung 69 | Links: Dynamische Animation zur Anzeige der Entnahmeposition, rechts: Multi-Point-Projektion zur Anzeige der Montageposition und Informationen zur Gesteninteraktion (Eigendarstellung)

Die Arbeitsanweisung mit dynamischer Animation besteht aus einem weißen Lichtpunkt in der Mitte und animierten orangenen Pfeilen, welche sich kontinuierlich zum Punkt bewegen. Die Animation wurde mithilfe von Adobe After Effects¹⁹ erstellt und als MP4-Dateiformat dem Projektionssystem bereitgestellt. Durch die bewegten Pfeile wird die Information schnell im Raum wahrgenommen, was sich besonders für den Einsatz in der Großgeräte- und Baustellenmontage eignet. Bei der Multi-Point-Projektion wird eine Kombination aus Arbeitsinformationen mit hohem (Piktogramm und kurze Text) und geringem (Lichtfläche) Detailgrad verwendet. Der Fokus des Projektors wird dabei auf die Position der Arbeitsanweisung mit hohem Detailgrad eingestellt und durch z. B. Lichtflächen können weitere Informationen angezeigt werden. Zusammengefasst eignet sich die dynamische In-situ-Projektion und Informationsführung, umgesetzt mit der Anzeige von Lichtflächen oder dynamischen Animationen, besonders für die Informationsbereitstellung und als Unterstützung für simple Montagetätigkeiten („Pick & Place“) und vor allem zum schnellen Auffinden von Positionen oder Bauteilen im großen Arbeitsraum. Dabei wird der Suchvorgang verbessert und dem Menschen intuitiv vermittelt, welches Teil zu entnehmen ist und wo es, je nach Anforderung grob oder exakt, platziert werden soll. Die Umsetzung der dynamischen Informationsbereitstellung stellt damit ein großes Potenzial für die Großgeräte- und Baustellenmontage dar, um den Menschen bei der effektiven und effizienten sowie auch bei der ergonomischen Umsetzung der Montagetätigkeiten zu unterstützen.

¹⁹ www.adobe.com

3.3.3 Nutzeradaptive Gesteninteraktion

Um in der industriellen Baustellenmontage sowie am Demonstrator in der TU Wien-Pilotfabrik Menschen und deren Gesten erkennen zu können, müssen Trainingsdaten in Form von vorklassifizierten bzw. gelabelten Objektdaten dem Deep-Learning-Algorithmus zum „Lernen“ zur Verfügung gestellt werden (vgl. Kapitel 2.3.4 und 2.4.2). Aus den Trainingsdaten ermittelt der Algorithmus, welche Objekte welchen Klassen zugeordnet werden sollen. Speziell in der Baustellenmontage bedarf es aufgrund des höheren Abstandes zwischen Kamera und Objekten (ca. 4 Meter) und aufgrund der Tatsache, dass die Personen, Posen und Gesten von oben aus erkannt werden sollen, und wegen der verschiedenen Oberflächen und Farben der Montageobjekte bzw. des Hallenbodens spezieller Trainingsdaten. Demnach müssen eigene Trainingsdaten aufgenommen, gelabelt und dem Algorithmus zur Verfügung gestellt werden.

Positions- und Gestenerkennung mit manuell gelabeltem Datensatz

Für die Verwendung zur Steuerung der dynamischen Projektion sowie zur Interaktion mit den Arbeitsinformationen (vgl. Kapitel 3.2.2) werden folgende Objektklassen für die Umsetzung der Positions-, Posen- und Gesteninteraktion in der industriellen Baustellenmontage definiert:

- Objektklasse 1: „Mensch von oben“ (männlich und weiblich)²⁰
- Objektklasse 2: Geste „offene Hand“
- Objektklasse 3: Geste „Faust“

Die Anforderung bei der Generierung des Datensets „Mensch von oben“ war, dass ein Mix aus männlichen und weiblichen Personen mit unterschiedlichen äußeren Merkmalen (Haarlänge, Haarfarbe, Kleidung etc.) integriert wird. Für die Gestaltung der Gesten wurden intuitive verständliche Gesten der Hand (vgl. Kapitel 2.4.2) ausgewählt, welche auch von der großen Entfernung und von oben gut unterscheidbar sind (Abbildung 70).



Abbildung 70 | Links: Geste „offene Hand“, rechts: Geste „Faust“ (Eigendarstellung)

Die Geste „offene Hand“ sowie die Geste „Faust“ eignen sich demnach als gut unterscheidbar und sind weit verbreitete Gesten zum „Weiterschalten“ bzw. zum An- und Abwählen von z. B. Informationen (vgl. Kapitel 2.4.2).

²⁰ Die Objektklasse 1: „Mensch von oben“ diente in den ersten Iterationen um ein Tracking des Projektionssystem, analog zur Bewegung des Menschen im Raum, umzusetzen. Für das finale Assistenzsystem und die Gesteninteraktion wurden nur Objektklasse 2 und 3 verwendet.

Um die Trainingsdaten zu generieren, wurde eine Kamera zur Aufzeichnung des RGB-Videostreams über der Montagestation montiert und ein Beispielmontageprozess mit 10 Trainingsteilnehmer*innen durchgeführt. Die Personen bewegten sich dazu ca. 40 Minuten (20 Min. Männer, 20 Min. Frauen)²¹ um die Baugruppe, nahmen verschiedene Positionen ein und führten die zwei definierten Gesten an verschiedenen Stellen aus. Aus dem 40-minütigen Videostream wurden anschließend die einzelnen Bilder (Frames) erzeugt und für die weitere manuelle Klassifizierung bzw. zum Labelling zur Verfügung gestellt. Das Labelling wurde mithilfe des pythonbasierten Labelling-Tool „BBbox“ durchgeführt, bei dem manuell die einzelnen Bounding-Boxen (Labels) über die jeweiligen definierten Objektklassen der Bilder gezogen und die Koordinaten der Boxen anschließend als Textformat abgespeichert wurden (Abbildung 71).²²

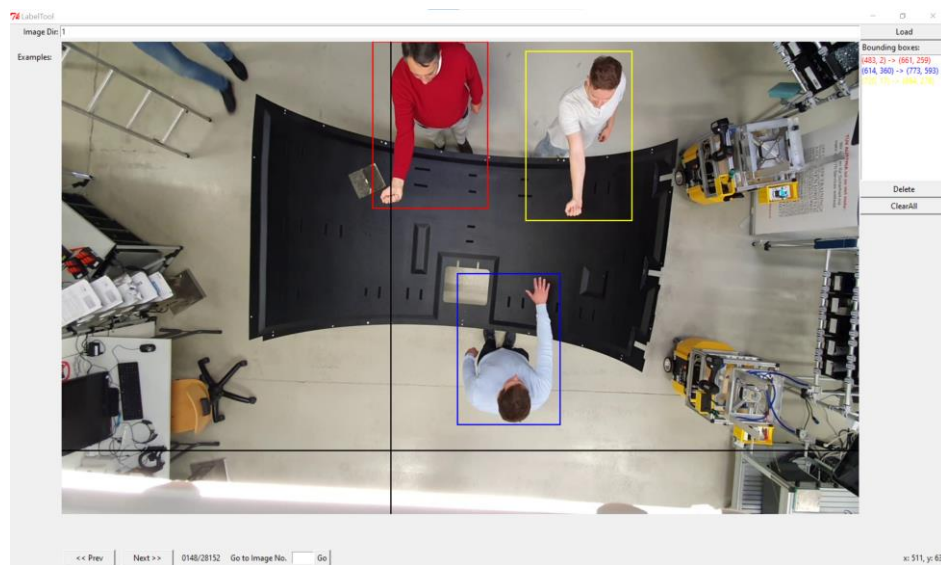


Abbildung 71 | Vorklassifizierung der Objektklasse „Mensch von oben“ mit den Labels (Eigendarstellung)

Insgesamt wurden aus dem Videostream bzw. den einzelnen Frames über 60 000 Menschen (30 000 Frauen, 30 000 Männer) und über 50 000 Gesten (Gesten „offene Hand“ und „Faust“) in verschiedenen Positionen und Ausrichtungen manuell gelabelt. Die gelabelten Daten inklusive Koordinaten wurden anschließend mit mehreren Instanzen der Hochleistungsrechner trainiert (vgl. Kapitel 3.2.2). Das Ergebnis eines dieser Trainings wird in Abbildung 72 präsentiert.

²¹ Es wurde bei der Datensatzerstellung darauf geachtet, einen ausgeglichenen Mix an Männern und Frauen zu integrieren. Um jedoch einen diskriminierungsfreien Datensatz zu gestalten, bedarf es weiterer diverser Trainingsdaten (z. B. Männer und Frauen mit verschiedenen Hautfarben, Körper- und Handgrößen etc.) sowie Gestendaten mit Verwendung bzw. Tragen von Handschuhen, Schmuck oder Werkzeug.

²² Der Zeitaufwand für das manuelle Labeln eines der Objekte, ähnlich Abbildung 71, beträgt durchschnittlich 2 Sekunden.

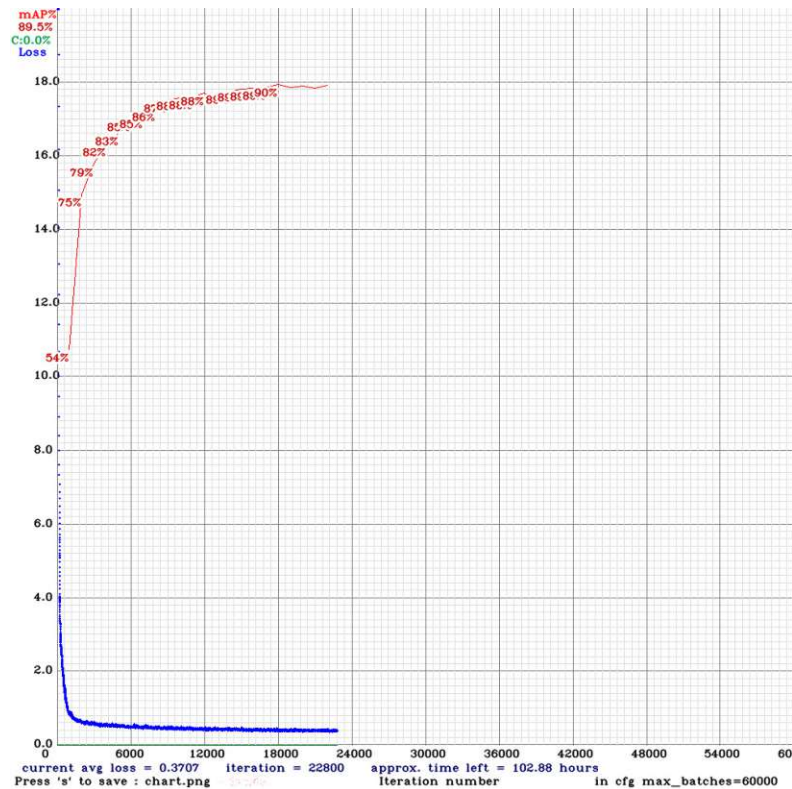


Abbildung 72 | Trainingsergebnisse YOLO (Eigendarstellung)

Dabei wurden 60 000 Bilder in mehreren Iterationen trainiert bis eine „Mean Average Precision“ (mAP), Maßeinheit für die Genauigkeit der Erkennung, von ca. 90 %, und ein „Average Loss“, Maßstab für die Fehlerrate, von ca. 0,37, erreicht wurde. Der trainierte Gestendatensatz wurde anschließend dem Objekterkennungsalgorithmus für die Verwendung in der Montage zur Verfügung gestellt. Die Vorgehensweise der manuellen Datensatzerstellung stellt einen möglichen Ansatz dar, um einen eigenen Gestendatensatz für die Großgeräte- und Baustellenmontage zu erzeugen, benötigt jedoch auch einen hohen Zeitaufwand für die Aufnahme der Montageszene und das manuelle Labeln der Gestenbilder.

Gesteninteraktion mit synthetisch erzeugtem Datensatz

Um den Aufwand des manuellen Labelns zu optimieren, wurde nach alternativen Methoden zur Erstellung des Gestendatensatzes gesucht und versucht die Anzahl der Gestendaten, bei gleichbleibender bzw. verbesserter Qualität der Erkennung, zu reduzieren. Durch mehrere Iterationen und Tests konnte die Gesamtanzahl an Gestenbildern auf 15 000 reduziert und der Prozess mit einem synthetisch erzeugten Datensatz vereinfacht werden. Dabei wurde bei der Erstellung des synthetischen Datensatzes wie folgt vorgegangen:

1. Aufnahme von definierten Gesten auf neutralem Hintergrund (vgl. Abbildung 73)
2. Entfernung des Hintergrundes mit „GrabCut“ (Rother et al., 2004)
3. Aufnahme der realen Szene (Montagestation) als Standbild
4. Größenskalierung der Geste in Bezug zur Szene und Abstand zur Kamera
5. Automatisierte Integration der Gesten an unterschiedlichen Positionen im Standbild (inklusive Winkel- bzw. Rotations-Augmentierung) (vgl. Abbildung 74)
6. Automatisierte Speicherung der Frames inklusive Koordinaten- und Gestenlabels

Für die Erzeugung des synthetischen Datensatzes wurden je fünf Gesten „offene Hand“ und „Faust“ von fünf verschiedenen Personen (vier männlich, eine weiblich) mittels der RealSense-Kamera auf neutralem Hintergrund aufgenommen und anschließend der Hintergrund mit dem Tool „GrabCut“ (Rother et al., 2004) entfernt (Abbildung 73).



Abbildung 73 | Links: Aufnahme der Gesten vor neutralem Hintergrund, rechts: entfernter Hintergrund mittels GrabCut (Rother et al., 2004) (Eigendarstellung)

Um eine realistische Szene der Montage als Background aufzunehmen, wurde die Kamera in vier Metern Höhe über dem Montagebereich positioniert und einige statische Bilder des Hallenbodens, einschließlich einiger Bausteine und Personen, aufgenommen. Um die getrennt aufgenommenen Gesten anschließend in der richtigen Größenskalierung in den Background zu integrieren, wurde der Abstand zwischen Kamera und Gestenausführungsebene ermittelt und die Gestenbilder dementsprechend skaliert. Mithilfe eines automatisierten Prozesses (Python-Skript) wurden die ausgeschnittenen Gestenbilder zufällig an 15 000 verschiedenen Positionen mit unterschiedlichen Orientierungen in die vorab aufgenommenen Standbilder der Szene bzw. des Hintergrunds integriert. Abbildung 74 gibt einige Beispiele von verschiedenen Positionen in unterschiedlichen Szenen.

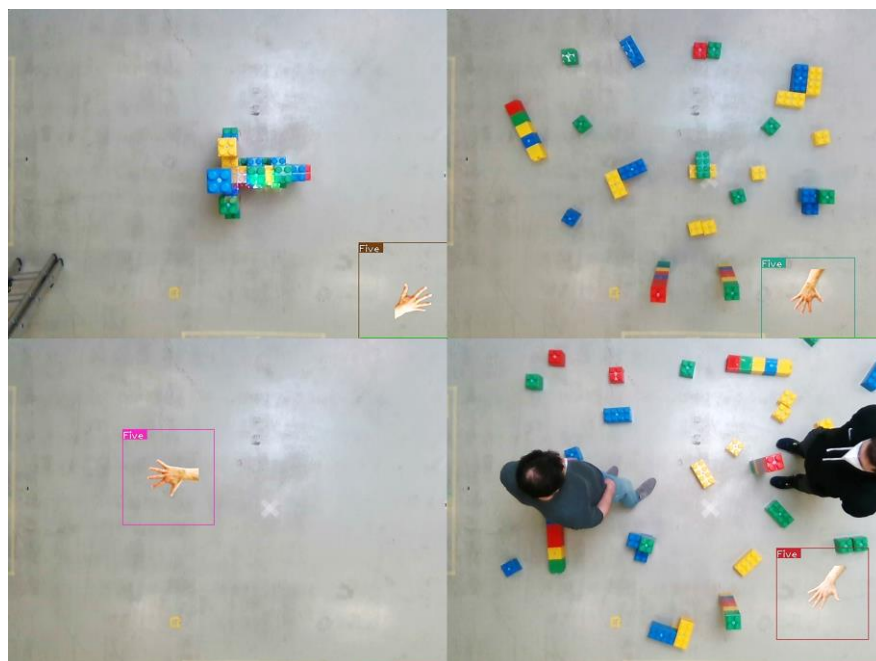


Abbildung 74 | Automatisierte Integration der Gesten in die Szene (Eigendarstellung)

Die 15 000 Gestenbilder, inklusive Textdateien mit zugehörigen Labels und Koordinaten, wurden anschließend automatisiert abgespeichert und dem Hochleistungsrechner (vgl. Kapitel 3.2.2) zum Trainieren zur Verfügung gestellt. Der trainierte Gestendatensatz wurde anschließend dem YOLO-Algorithmus für die Objekterkennung in der Montage bereitgestellt.

Vergleich Arbeitsaufwand: manuell und synthetisch erzeugter Gestendatensatz

Eine Abschätzung des Arbeitsaufwands zur Aufbereitung der manuell und synthetisch erzeugten Gestendatensätze mit 15 000 Gestenbildern wird in Tabelle 15 präsentiert.

Tabelle 15 | Vergleich Arbeitsaufwand der Gestendatensatzerstellung

Datensatzgestaltung	manuell erstellt	synthetisch erstellt
Aufnahme/Aufbereitung der Gestenbilder und Montageszenen	3 h für Videoaufnahme + Frame-Erstellung	0,5 h für die Aufnahme von 10 Gestenbildern + Hintergrundentfernung
Labeln der Gestenbilder	8 h für manuelles Labeln und Speichern von 15 000 Gestenbilder (~2 Sekunden pro Gestenbild)	2,5 h für die automatisierte Integration und Labeln von 15 000 Gestenbilder (~0,6 Sekunden pro Gestenbild)
Gesamtdauer	11 h	3 h
Verbesserung		8 h (73 %)

Durch die Optimierung der Anzahl der benötigten Gesten und die Automatisierung des Labeling-Prozesses können (neue) Gestensätze, im Vergleich zum manuellen Labeln, deutlich schneller integriert und flexibler in verschiedenen Szenen bzw. Hintergründen positioniert werden. Des Weiteren konnte durch den Einsatz des synthetischen Datensatzes mit verschiedenen Variationen von Händen und Handgrößen experimentiert werden und damit die Gesteninteraktion in Bezug auf Stabilität und Erkennungsgeschwindigkeit im Experiment 3 (vgl. Kapitel 5.2.3) optimiert werden.

4. Demonstration von Spatial Augmented Reality in der industriellen Großgeräte- und Baustellenmontage

Für die Demonstration und Evaluierung des Spatial Augmented Reality Assistenzsystems steht eine industrienähe Pilotfabrik der TU Wien (<https://www.pilotfabrik.at>) zur Verfügung, in der mehrere Demonstratoren, welche die industrielle Großgeräte- und Baustellenmontage repräsentieren, aufgebaut werden. An den Demonstratoren werden die gestaltenden Konzepte und Artefakte aus Kapitel 3 demonstriert und im iterativen Designprozess verbessert. In den folgenden Abschnitten wird der Aufbau des Spatial Augmented Reality Assistenzsystems an verschiedenen Use Cases vorgestellt und praxisrelevante Ergebnisse der einzelnen Gestaltungsaspekte präsentiert.

Um folgend die Gestaltungsprozesse, Tests und Evaluierungen durchzuführen, wurden die Komponenten des Spatial Augmented Reality Assistenzsystems (dynamisches Projektionssystem und visuelles Interaktionssystem) (vgl. Kapitel 3.2) in der Pilotfabrik aufgebaut und vernetzt (Abbildung 75).

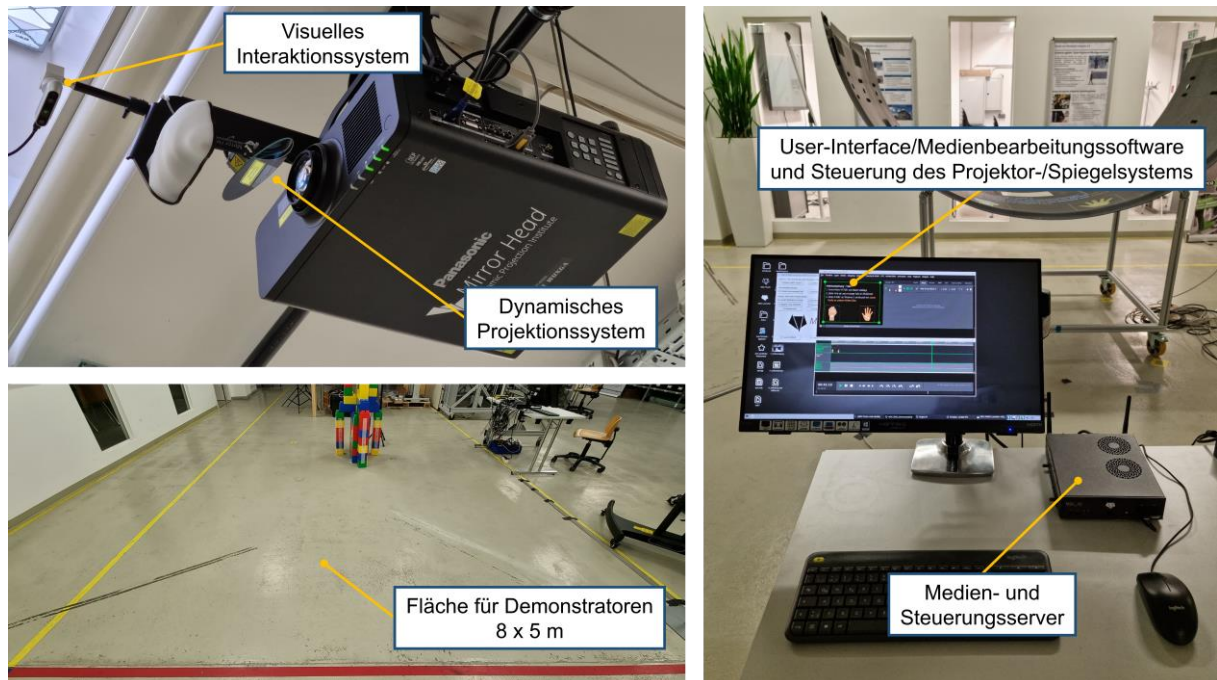


Abbildung 75 | Spatial Augmented Reality Assistenzsystem in der TU Wien-Pilotfabrik
(Eigendarstellung)

Das Projektionssystem mit dem Spiegelsystem wird dazu an der 4 Meter hohen Decke der Pilotfabrik montiert und ermöglicht damit die dynamische Projektion über einen großen Arbeitsbereich. Ebenso wurde das visuelle Interaktionssystem (Kamerasystem) an der Decke nahe dem Projektor platziert und ermöglicht die Erfassung von Objekten im Arbeitsbereich. Der PC zur Steuerung der Medien und des Projektors inklusive Bildschirm mit User-Interface befindet sich auf einem Arbeitstisch seitlich der Montagefläche und ist mit diversen kabelgebundenen und kabellosen Verbindungselementen mit den Systemen vernetzt.

4.1 Demonstratoren und Use Cases

Um die Demonstration des Spatial Augmented Reality Assistenzsystems zu ermöglichen und die einzelnen Design-Aspekte zu gestalten, werden praxisrelevante Use Cases verwendet und geeignete Demonstratoren dazu aufgebaut (Abbildung 76).



Abbildung 76 | Demonstratoren und Use Cases der Forschungsarbeit (Eigendarstellung)

Der Use Case „Fan-Cowl-Montage“ charakterisiert die Großgeräte- und Baustellenmontage und wird verwendet, um Tests des Spatial Augmented Reality Systems durchzuführen und die Projektion von Arbeitsanweisungen auf amorphem und dunklem Untergrund zu evaluieren. Die Spielzeugbaustein-Montage dient als erweiterter und komplexerer Use Case, bei dem die Montagebauteile an mehreren Bereitstellungsflächen zur Verfügung gestellt werden und ein komplexes Montageobjekt in mehreren Lagen gebaut wird. Die Demonstratoren eignen sich durch ihre Größe und Gestalt sowie den zugrunde liegenden Montageprozess für die Demonstration des Spatial Augmented Reality Assistenzsystems in der Großgerätemontage.

Demonstrator 1: Fan-Cowl-Montage

Bevor die Fan-Cowl-Montage für die Demonstration eingesetzt wurde, stand eine Winglet-Montagestation für die Herstellung eines Winglets/Sharklets,²³ einer Tragfläche des Flugzeugtyps Airbus A-350, für erste Tests zur Verfügung (Abbildung 77)²⁴.

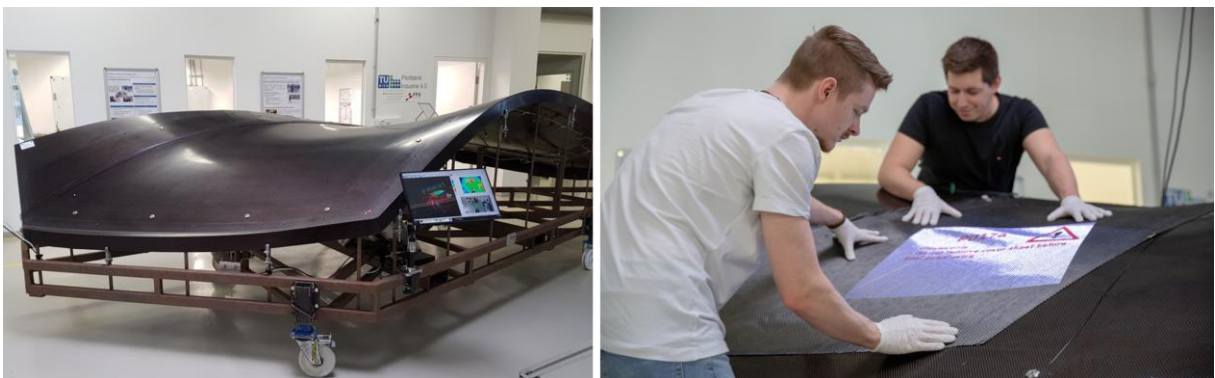


Abbildung 77 | Demonstrator „Winglet-Montagestation“ (Eigendarstellung)

²³ Anmerkung: Winglets (Airbus) bzw. Sharklets (Boeing) sind Flugzeugbauteile am Ende der Tragfläche zur Reduktion des Treibstoffverbrauchs.

²⁴ Der Demonstrator wurde vom Unternehmen FACC AG für Forschungstätigkeiten der TU Wien bereitgestellt.

Der Demonstrator hat eine Abmessung von 5500 × 2000 × 1000 Millimeter und ein Gewicht von ca. 800 Kilogramm. Die Oberfläche ist glatt spiegelnd, die Form ist polymorph gestaltet und die Farbe ist schwarz-grünlich und besitzt eine Carbonfaser-Texturierung. Diese Gegebenheiten der Projektionsfläche stellen für die Projektion von Arbeitsinformationen eine besondere Herausforderung dar. Der Demonstrator repräsentierte einen Teilprozess der Winglet-Herstellung, bei dem Materialkomponenten (Carbonfaser-Mattenzuschnitte) in mehreren Lagen von einem oder mehreren Mitarbeiter*innen aufgeklebt bzw. „montiert“ werden. Die Winglet-Montagestation wurde aus organisatorischen Gründen anschließend gegen die **Fan-Cowl-Montagestation** (Abbildung 78) ausgetauscht und stand anschließend mit Maßen von 2000 x 1500 x 1700 Millimetern als Hauptdemonstrator zur Verfügung.



Abbildung 78 | Fan-Cowl-Montagestation in der TU Wien-Pilotfabrik (Eigendarstellung)

Ein „Fan-Cowl“ bezeichnet ein Flugzeugturbinengehäuse, welches aus zwei Halbschalen besteht und analog zum Winglet hergestellt wird. Die Oberfläche des Demonstrators ist ebenso polymorph gestaltet, jedoch mit geringerer Komplexität als bei der Winglet-Montage. Die Farbe der Oberfläche ist mattes Schwarz und enthält Strukturelemente wie Einbuchtungen, Aussparungen sowie spezielle Wabenstrukturelemente, die ebenso eine Herausforderung für die Projektion darstellen.

Demonstrator 2: Spielzeugbaustein-Montage

Um die Erkenntnisse aus dem ersten Demonstrator zu verallgemeinern und das Assistenzsystem in einem weiteren Use Case zu demonstrieren und zu evaluieren, wurde die Spielzeugbaustein-Montage als zweiter Demonstrator verwendet (Abbildung 79).



Abbildung 79 | Spielzeugbaustein-Montage (Eigendarstellung)

Dabei wurden ein komplexerer Montageprozess mit 18 Montagetasks und 64 Bauteilen definiert und durch gezielte Informationsbereitstellung im großen Arbeitsraum die Montage eines Turmes angeleitet. Die Bausteine standen dabei vorkommissioniert als Baugruppe oder einzeln an drei verschiedenen Bereitstellungsflächen zur Verfügung. Bei dem Use Case wurden einerseits am Hallenboden über eine Fläche von 5 x 3 Metern Arbeitsinformationen sowie dynamische Animationen angezeigt und andererseits durch In-situ-Projektion mit Lichtflächen die Montageposition der Bauteile in der jeweiligen Lage angezeigt.

4.2 Projektion von Arbeitsinformationen im industriellen Umfeld

Um eine optimale Gestaltung der Projektion direkt in der Montagestation der Baustellenmontage umzusetzen, bedarf es iterativer Gestaltungsprozesse sowie experimenteller Tests am Demonstrator. Dabei werden verschiedenen Gestaltungsvarianten von Arbeitsinformationen sowie einzelne Aspekte zu Visualisierungsarten, Struktur und Farbgestaltung demonstriert und die projektionsgerechte Aufbereitung evaluiert. Folgend werden die Möglichkeiten des dynamischen Projektionssystem mit dem Full-HD+ Projektor und dem steuerbarem Spiegelsystem (Abbildung 80) am Demonstrator dargestellt und Empfehlungen für die Umsetzung projektionsoptimierter Arbeitsanweisungen präsentiert.



Abbildung 80 | Full-HD+ Projektor mit digital steuerbarem Spiegelsystem (Eigendarstellung)

Dabei wird besonders auf die Herausforderungen der dynamischen Projektion bei der Anzeige auf gekrümmten polymorphen Projektionsoberflächen sowie dunklen Materialien eingegangen, im Weiteren werden praxisrelevante Tests durchgeführt. Die Ergebnisse dienen der projektionsoptimierten Gestaltung, Arbeitsinformationen und die Parameter des Projektors und Spiegelsystems optimal einzustellen.

4.2.1 Herausforderungen bei projizierten Arbeitsinformationen

Da in der Großgeräte- und Baustellenmontage kaum planare Projektionsflächen (ausgenommen der Hallenboden) zur Verfügung stehen, sondern größtenteils polymorphe Objekte vorhanden sind, ist es erforderlich die Gegebenheiten der physischen Projektionsflächen zu analysieren und die Projektionen den polymorphen Gegebenheiten

anzupassen, um auf unterschiedlichen Formen, Oberflächen und Materialien hochwertige Projektionen zu erzeugen. Die Erkenntnisse aus diversen Tests und Demonstrationen sollen dabei unterstützen, eine optimale Gestaltung der Arbeitsinformationen durchzuführen sowie geeignete Einstellungen der Projektionsparameter zu definieren. Nach ersten Versuchen konnten folgende Herausforderungen und Einflussgrößen festgestellt werden, welche die Qualität der projizierten Informationen beeinflussen:

- **Form der physischen Projektionsfläche**
 - Polymorphe, gekrümmte Form
 - Einbuchtungen und Aussparungen
- **Oberflächenbeschaffenheit, Struktur und Rauheit**
 - Matte, texturierte, glatte Oberfläche
 - Dreidimensional strukturierte Oberfläche (z. B. Wabenkern-Strukturelement)
- **Helligkeit und Farbe der Projektionsfläche sowie der Umgebung**
 - Kontrast, Farbe der Projektionsfläche
 - Helligkeit der Umgebung (Tageslicht etc.)
- **Schärfe und Fokus bei der Multi-Point-Projektion**
 - Gleichzeitige Projektion auf Projektionsflächen mit unterschiedlichem Abstand zum Projektor
- **Beschattung der Projektionsfläche**
 - Beschattung durch Montageobjekte
 - Beschattung durch den Menschen

Im Folgenden werden die Herausforderungen und Einflussgrößen erörtert sowie die Erkenntnisse und potenziellen Gegenmaßnahmen am Projektionssystem zur Verbesserung der Projektion dargestellt (Tabelle 16).

Tabelle 16 | Herausforderungen der Projektion und Gegenmaßnahmen

Herausforderung	Erkenntnis durch Praxistests	Gegenmaßnahme/Einstellungsparameter
Polymorphe, gekrümmte Oberfläche	Gekrümmte Oberflächen verzerren die Projektion, sodass sie kaum lesbar ist.	Durch den Einsatz der manuellen oder automatischen Verzerrung kann die Krümmung der Projektionsfläche gut ausgeglichen werden und die Projektion lesbar sowie unverzerrt dargestellt werden.
Glatte Oberfläche	Glatten Oberflächen spiegeln die Projektion stark, sodass eine Blendung des Menschen auftritt und die Projektion schlecht lesbar ist.	Bei glatten Oberflächen kann durch Reduktion der Helligkeit des Projektors oder durch Anpassung des Projektionswinkels die Blendung reduziert werden.
Strukturierte Oberfläche (z. B. Waben-Strukturelement)	Strukturierte Oberflächen brechen die Projektion, Details der Medieninhalte werden unzureichend dargestellt.	Projektionen auf strukturierten Oberflächen sollten generell vermieden werden. Als weitere Maßnahme können Projektionsinhalte so aufbereitet werden, dass kleine Details vermieden und nur grob strukturierte Medien wie Symbole, Warnzeichen oder Piktogramme verwendet werden.
Farbe und Helligkeit des Materials sowie der Arbeitsumgebung	Die Farbe und Helligkeit des Materials sowie die Helligkeit der Arbeitsumgebung (Tageslicht, Beleuchtung etc.) wirkt sich auf die Lesbarkeit der Projektion	Bei der Aufbereitung der Projektion sollen die Farbe und Helligkeit der Projektionsfläche sowie der Arbeitsumgebung berücksichtigt (vgl. Kapitel 4.2.2) und die Inhalte danach angepasst werden. Dabei sollen grelle Farben mit hohem

	aus. So werden bei dunklem Material der Projektionsfläche dunkle Inhalte nicht optimal angezeigt. Ebenso können bei zu grellem Umgebungslicht die Informationen nur unzureichend erkannt werden.	Kontrast zur Oberfläche verwendet werden. Außerdem sollen die Einstellungen bzgl. Kontrast und Helligkeit des Projektors dementsprechend angepasst werden.
Schärfe und Fokus der Arbeitsinformation bei der Multi-Point-Projektion	Bei gleichzeitiger Projektion auf Flächen mit unterschiedlichem Abstand zum Projektor besteht die Herausforderung, dass sich der Fokus bzw. die Schärfe des Bildes nur auf eine Position (Single-Point-Projektion) einstellen lässt.	Beim Erstellen der Projektionsshow bzw. Festlegen der Projektionspositionen muss der Fokus so eingestellt werden, dass die Arbeitsinformationen scharf lesbar sind. Multi-Point-Projektion: Bei der Umsetzung einer Multi-Point-Projektion kann der Fokus nur auf einer Position scharf gestellt werden. Dementsprechend müssen Arbeitsinformationen mit hohem Detailgrad (z. B. Texte etc.) mit Arbeitsinformationen mit geringem Detailgrad (z. B. Lichtflächen, Lichtpunkte etc.) kombiniert werden und der Fokus auf diejenige mit hohem Detailgrad eingestellt werden (vgl. Kapitel 3.3.2). Alternative: Single-Point-Projektion + dynamische Projektion Als alternative Möglichkeit kann die Single-Point-Projektion mit Einsatz des dynamischen Spiegelsystems verwendet werden. Arbeitsinformationen sind dabei mit geringerem Umfang zu gestalten (z. B. Aufteilung einer Arbeitsanweisung in mehrere) und dynamisch mit Einsatz des Spiegelsystems und automatisierter Einstellung des Fokus umzusetzen.
Beschattung der Projektionsfläche durch Montageobjekte oder Menschen	Da der Projektor zentral über der Arbeitsstation montiert wurde, kann es temporär zur Beschattung der projizierten Arbeitsinformationen kommen.	Durch intuitive und projektionsoptimierte Gestaltung der Arbeitsinformationen mithilfe von Symbolen, Piktogrammen und dynamischen Animationen etc. (vgl. Kapitel 3.3.1) sind Informationen auch bei Teil-Beschattung noch ausreichend lesbar. Weiters kann, aufgrund der Möglichkeiten des Spiegelsystems und der Einstellung der Verzerrung, der Projektor auch seitlich der Montagestation installiert werden, um damit Beschattungen zu vermeiden.

Dabei stellen vor allem gekrümmte, glatte und strukturierte Oberflächen sowie verbaute Strukturelemente, welche zur Verstärkung des Montagebauteils eingesetzt werden, eine Herausforderung für die Projektion dar. Des Weiteren spielen die Farbe des Materials und die Helligkeit der Arbeitsumgebung sowie eine mögliche Beschattung durch andere Montageobjekte oder durch den Menschen bei der Gestaltung der Informationen und Einstellung der Parameter des Projektors eine essenzielle Rolle. Die Versuche am Demonstrator helfen bei der projektionsgerechten Aufbereitung der Informationen, je nachdem, welche Farben verwendet werden sollen, welche Größen und Detailgrade optimal sind und auf welchen Elementen und Oberflächen eine Projektion vermieden werden soll. Außerdem wird durch die praktischen Tests ermittelt, was bei der Erstellung von

Projektionsshow und Definition der Projektionspositionen in Bezug zur Beschattung zu beachten ist und wie sich der Abstand von Projektionsflächen zum Projektor (z. B. bei der Multi-Point-Projektion) auf die Scharfstellung der Arbeitsinformation auswirkt. Abbildung 81 zeigt einen Auszug von verschiedenen Projektionstests auf unterschiedlichen Oberflächen und Positionen der Demonstratoren.



Abbildung 81 | Praxisbeispiele von Projektionen auf verschiedenen Oberflächen und Materialien (Eigendarstellung)

Dabei konnte festgestellt werden, dass Projektionen bei optimaler Aufbereitung (Detailgrad, Farben, Größe etc.) sowohl auf glatter und matter Oberfläche, auf gekrümmten Formen sowie auf strukturierteren und texturierten Elementen ausreichend erkannt werden konnten. Dazu müssen jedoch stets die Projektionsflächen analysiert und die Arbeitsinformationen dementsprechend danach angepasst werden. Es konnte ermittelt werden, dass durch den Einsatz des Full-HD+ Projektors mit einer möglichen Helligkeit von 6200 ANSI Lumen und einem Kontrast von 10 000:1 die Projektionen auch bei Tageslicht und heller Umgebungsbeleuchtung gut lesbar sind, wodurch sich ein Einsatz für das industrielle Umfeld besonders eignet. Des Weiteren eignet sich die Aufbereitung und Kombination von Arbeitsinformation mit hohem (z. B. Texte etc.) und geringen Detailgrad (z. B. Lichtflächen) vor allem auch für die Multi-Point-Projektion.

Demnach eignet sich das dynamische Projektionssystem mit den ausgewählten Komponenten und Spezifikationen optimal für den Einsatz in der industriellen Großgeräte- und Baustellenmontage und bietet den Mehrwert der direkten Darstellung von Informationen an den Oberflächen der Montage sowie der intuitiven Positionierung und Bearbeitung für unterschiedliche polymorphe Projektionsflächen.

Des Weiteren stellt die projektionsgerechte Aufbereitung von Arbeitsinformationen einen relevanten Part für die hochwertige Anzeige von Projektionen direkt am Bauteil dar. Durch die Demonstration von State-of-the-Art-Arbeitsanweisungen zur Herstellung eines Winglets

konnte festgestellt werden, dass herkömmliche Arbeitsinformationen meist textlastig, umfangreich und strukturell unübersichtlich sind und sich demnach für den Einsatz mittels Projektionssystem nicht eignen. Es ist deswegen erforderlich, Arbeitsinformationen projektionsgerecht zu gestalten, damit der sinnvolle Einsatz der direkten Projektion am Arbeitsplatz gewährleistet werden kann. Die Herausforderungen, Erkenntnisse sowie potenzielle Vorschläge für die Gestaltung werden in Tabelle 17 dargestellt.

Tabelle 17 | Erkenntnisse und Vorschläge für projektionsgerechte Medienaufbereitung

Herausforderung	Erkenntnis durch Praxistests	Vorschläge
Art und Struktur des Inhalts	Viel Text wird meist nicht ordnungsgemäß gelesen. Besonders relevante Informationen und Warnungen können dadurch schnell untergehen.	Die Projektion ermöglicht auch Bilder und Videos sowie Animationen zu verwenden. Durch geeignete Aufbereitung der Informationen mit Piktogrammen, Symbolen, knalligen Farben sowie Kurzvideos können Informationen intuitiver dargestellt werden. Warnungen können dadurch farblich hervorgehoben werden und gehen nicht unter.
Größe der Inhalte	Kleine Schriftarten, Bilder und Grafiken sowie Fließtexte eignen sich nur geringfügig für die Projektion.	Kurz prägnante Texte mit großer Schriftgröße sowie selbsterklärende Grafiken bieten eine gute Lesbarkeit für die Projektion direkt am Bauteil.
Farbe und Kontrast des Inhalts	Blasse Farben und geringer Kontrast im Vergleich zur Projektionsfläche fördern eine schlechte Lesbarkeit der Inhalte.	Bei der Projektion sollen kräftige Farben mit hohem Kontrast im Vergleich zur Projektionsfläche verwendet werden.

Durch die Erkenntnisse der Analyse der bestehenden Arbeitsinformationen sowie durch erste Tests von verschiedenen Projektionsinhalten am Demonstrator konnten bereits relevante Inputs für die projektionsgerechte Aufbereitung gewonnen werden. Ausgehend davon werden im folgenden Kapitel detaillierte Gestaltungsaspekte aus Kapitel 4 demonstriert und Erkenntnisse gesammelt, welche Visualisierungsarten, und Formen sowie welche Farbgestaltung sich am besten für die projektionsgerechte sowie intuitiv verständlich Gestaltung eignen.

4.2.2 Projektionsoptimierte Arbeitsinformationen

Der essenzielle Nutzen der Test und Demonstrationen ist, eine optimale visuelle Darstellung der Arbeitsinformationen direkt auf den jeweiligen Oberflächen der Montagestation zu ermitteln. Dabei wurden praxisrelevante Experimente mit verschiedenen Varianten der Arbeitsinformationen durchgeführt, um herauszufinden, welche Gestaltungsaspekte am geeignetsten für die Verwendung mittels Projektionssystem sind. Folgend wird ein Auszug der demonstrierten Arbeitsanweisungen und Gestaltungsvarianten präsentiert.

Die Erkenntnisse aus den Tests am Demonstrator werden dazu verwendet, um Arbeitsanweisungen projektionsgerecht aufzubereiten und die intuitive Verständlichkeit zu testen.

Abbildung 82 zeigt einen Auszug der demonstrierten Arbeitsinformationen am Demonstrator in der industriellen Baustellenmontage, welche in Kapitel 4 gestaltet wurden.

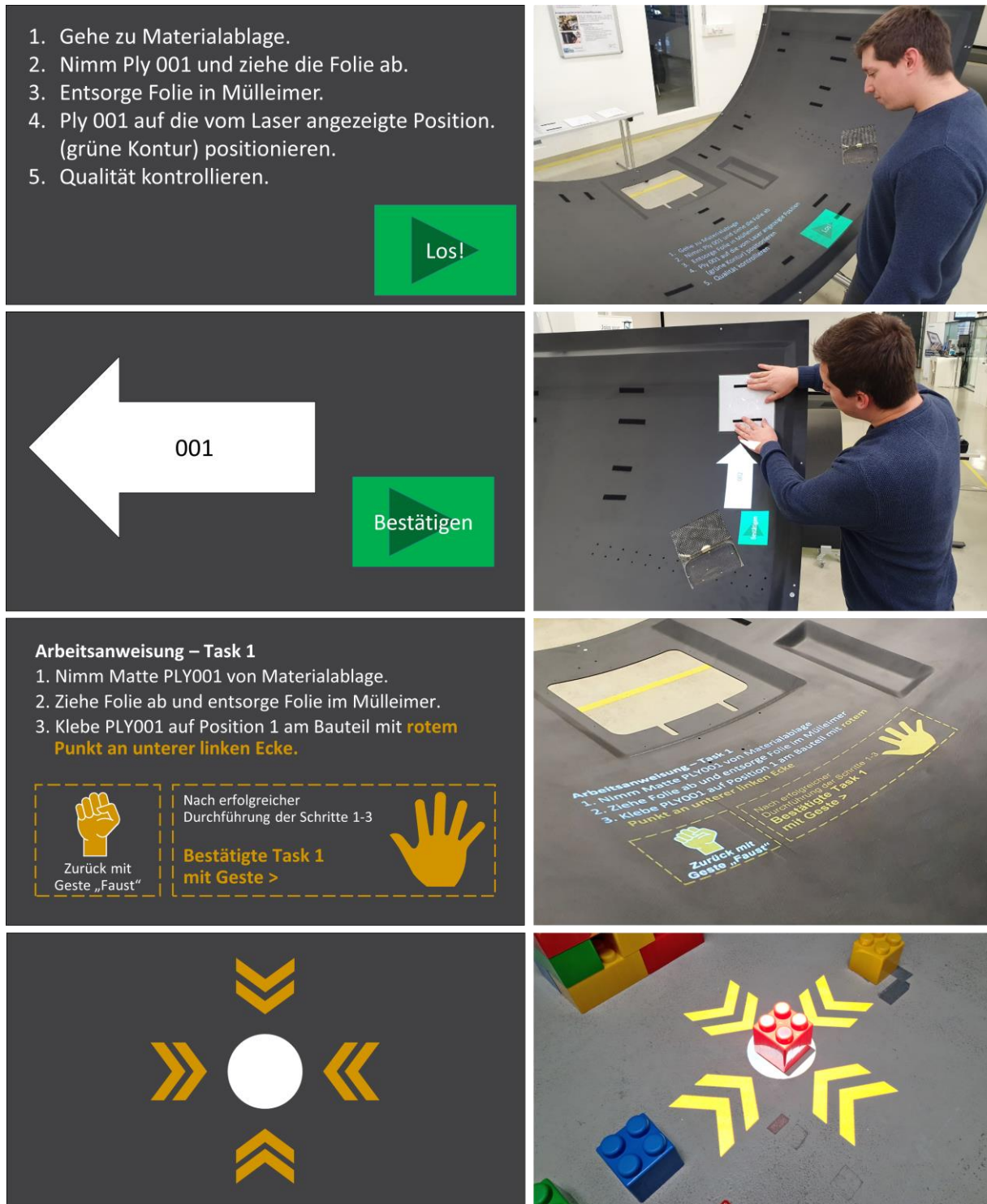


Abbildung 82 | Projektionsgerechte Aufbereitung der Arbeitsinformationen (Eigendarstellung)

Durch die praxisrelevanten Tests am Demonstrator konnten die aufbereiteten und optimiert gestalteten Arbeitsinformationen auf Lesbarkeit, Kontrast etc. evaluiert und ggf. optimiert werden. Folgend werden zwei Varianten der Darstellung der Gesteninformation verglichen (Abbildung 83).



Abbildung 83 | Vergleich von Gesteninformationen: textlastige Information (links), Verwendung von Piktogrammen (rechts) (Eigendarstellung)

Dabei konnte festgestellt werden, dass die intuitive Aufbereitung einen potenziellen Mehrwert in Form einer strukturierteren und kompakteren Darstellung bietet. Diese Art der intuitiven Darstellung mit Piktogrammen und Symbolen kann nach Unterweisung über deren Bedeutung durchgängig bei allen Anweisungen umgesetzt oder erst im Zeitverlauf der Montagetätigkeit integriert werden.

Demonstration von Hintergründen und Kontrasten

Für die Gestaltung der Arbeitsinformationen wurden zu Beginn der Versuche Arbeitsinformationen mit schwarzer oder färbiger Schrift auf weißem Hintergrund verwendet und für die Projektion in der Baustellenmontage bereitgestellt. Durch mehrere Tests und Iterationen der projektionsgerechten Gestaltung wurde erkannt, dass sich weiße oder helle Farben auf transparentem Hintergrund für die Darstellung in der Baustellenmontage besonders eignen. Abbildung 84 stellt dazu einen Vergleich dar.

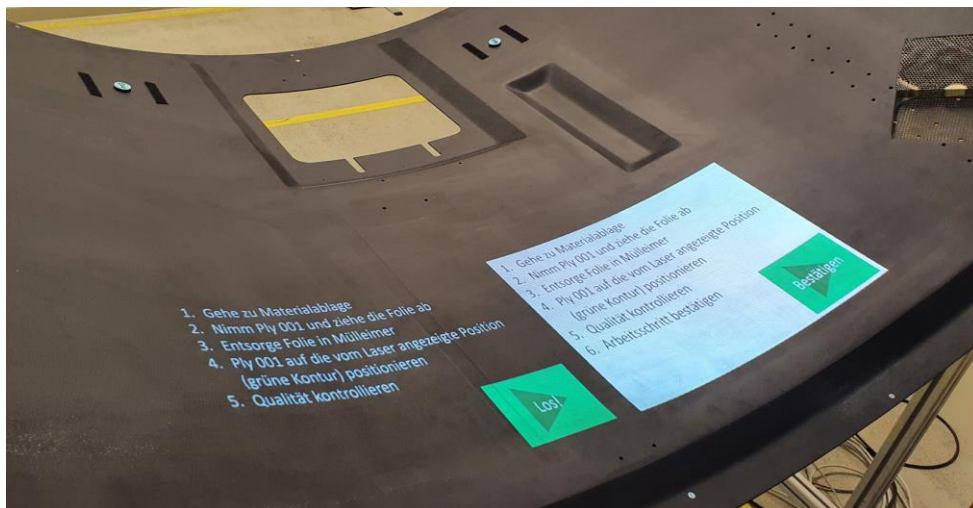


Abbildung 84 | Demonstration von Arbeitsinformationen mit weißem und transparentem Hintergrund (Eigendarstellung)

Es konnte ebenso festgestellt werden, dass bei der Gestaltung ein schwarzer Hintergrund zu vermeiden ist, da der Projektor diesen ebenso projiziert und die Darstellung der Informationen verschlechtert bzw. bei bereits dunkler Projektionsfläche eine sichtliche schwarze Fläche entsteht. Demnach stellt die Verwendung von heller oder weißer Farbe auf transparentem

Hintergrund eine optimale Gestaltung bzgl. Kontrast und Lesbarkeit der Projektionen dar und soll stets für die Aufbereitung der Arbeitsinformationen verwendet werden.

Demonstration von Farben und Kontrasten

Wie oben beschrieben, stellt die Verwendung von geeigneten hellen Farben auf dunkler Projektionsfläche einen essenziellen Gestaltungsaspekt dar. Dementsprechend wurden neun verschiedene Farben (rot, blau, orange/gelb, grün) in jeweils zwei Abstufungen (hell und dunkel) sowie die Farbe „weiß“ demonstriert, um damit die Lesbarkeit und den subjektiven Kontrast zu evaluieren (Abbildung 85).

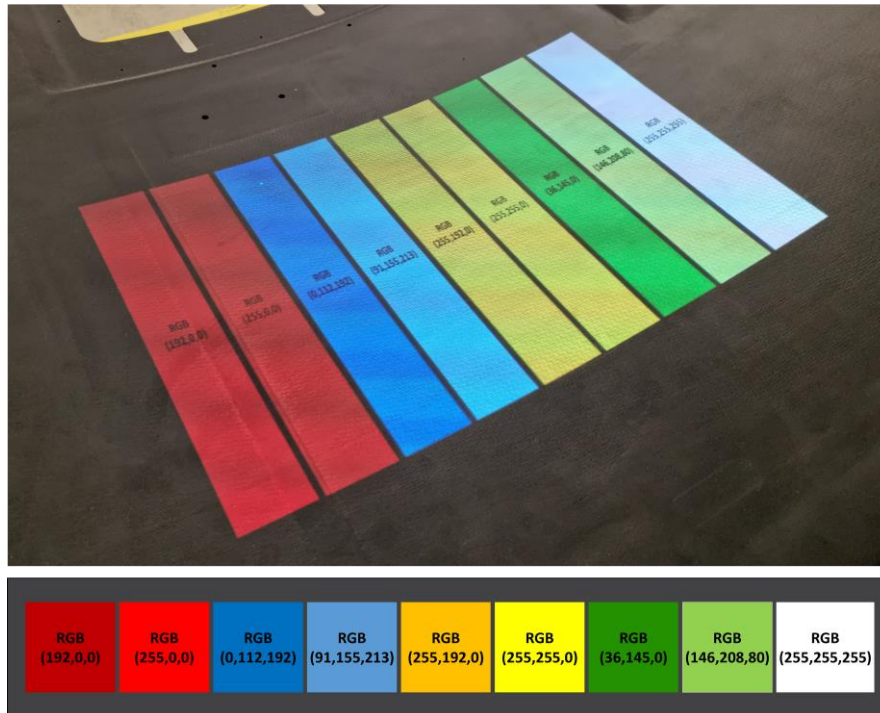


Abbildung 85 | Farbdemonstration in der Baustellenmontage (Eigendarstellung)

Dabei wurden folgende Standardfarben aus der Farbpalette von MS PowerPoint mit zugehörigen RGB-Farbcodes verwendet und am Demonstrator getestet:

- Dunkelrot RGB (192 0 0)
- Rot RGB (255 0 0)
- Dunkelblau RGB (0 112 192)
- Blau RGB (91 155 213)
- Gold-Orange RGB (255 192 0)
- Gelb RGB (255 255 0)
- Dunkelgrün RGB (36 145 0)
- Gelbgrün RGB (146 208 80)
- Reinweiß RGB (255 255 255)

Es konnte festgestellt werden, dass alle Farben im Vergleich zum dunklen Hintergrund einen hohen Kontrast liefern und gut lesbar sind. Jedoch wurde ermittelt, dass die Farben der Projektion etwas vom digitalen Abbild abweichen. Die projizierten Farben am Demonstrator werden heller bzw. kräftiger angezeigt als am Bildschirm und ähnliche Farben (z. B. Dunkelrot und Rot oder Gold-Orange und Gelb) werden am Demonstrator als gleiche Farben dargestellt. Dieser Aspekt der Unterscheidbarkeit hat Auswirkungen auf die Gestaltung der Arbeitsinformationen und trägt dazu bei, Fehlinterpretationen bei falscher Anzeige zu vermeiden und wurde anhand der Farben Gold-Orange und Gelb näher untersucht (Abbildung 86).

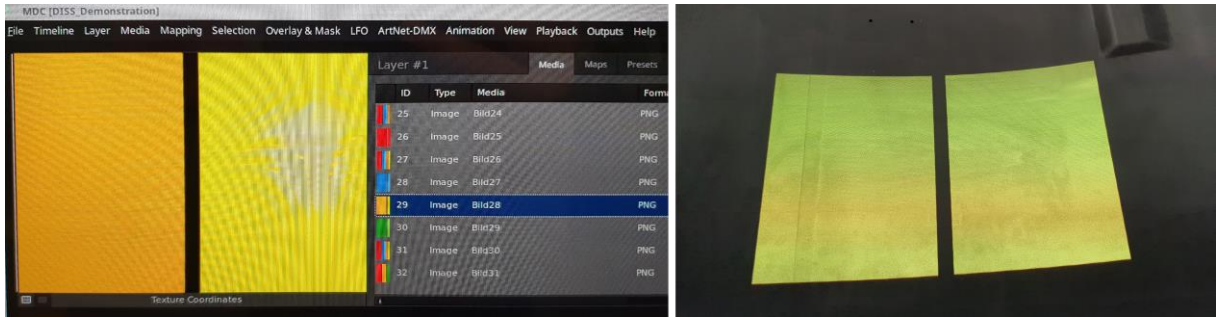


Abbildung 86 | Vergleich der Farben Gold-Orange und Gelb in der Software des Projektors und am Demonstrator (Eigendarstellung)

Dabei ist ersichtlich, dass die verwendeten Farben in der Medienbearbeitungssoftware des Projektors unterscheidbar angezeigt werden, jedoch der Unterschied bei der Projektion nicht mehr erkannt wird. Daraus ist zu schließen, dass Farben, welche sich nur geringfügig unterscheiden, für die Gestaltung von Arbeitsinformationen vermieden werden sollen, da sie vom Projektor nicht unterscheidbar angezeigt werden können²⁵.

Aus den Erkenntnissen wurden die Farbauswahl überarbeitet und kräftigere sowie unterscheidbarere Farben herangezogen. Dabei wurde nur ein Farbton pro Farbe ausgewählt und geringfügige Abstufungen vermieden. Die finalen Farben wurden am Demonstrator getestet und für die weitere Gestaltung der Arbeitsinformationen verwendet (Abbildung 87).



Abbildung 87 | Finale Farbgestaltung (Eigendarstellung)

Demnach eignen sich die Farben Gold-Orange (RGB 210 149 0), Grün (RGB 36 145 0) und Reinweiß (RGB 255 255 255) besonders als kontrastreiche Darstellung von Informationen und werden mit der Wahl der Farben Dunkelrot (RGB 192 0 0) und Dunkelblau (RGB 0 112 192) für projektionsgerecht gestaltete Arbeitsinformationen am Demonstrator verwendet.

Daraus resultiert, dass sich kräftige Farben gut für die Projektion von Arbeitsanweisungen eignen und dass darauf geachtet werden soll, dass große Unterschiede im Kontrast bzw. in der Abstufung der Farben Berücksichtigung finden, um mögliche Fehlanzeigen und demzufolge Fehlinterpretationen bei der Ausführung der Montagetasks auszuschließen.

²⁵ Hierbei ist anzumerken, dass die Darstellung ausschließlich mit dem vorhandenen Projektor (vgl. Kapitel 3.2.1) getestet wurde.

4.3 Dynamische In-situ-Projektion und Informationsführung

Wie oben erwähnt, stellen die In-situ-Projektion und Informationsführung essenzielle Gestaltungsaspekte bei der Verwendung eines Spatial Augmented Reality Assistenzsystems dar. Dazu wird folgend die Informationsbereitstellung der Arbeitsinformationen am Demonstrator der Fan-Cowl-Montage, beim Use Case der Nutensteinmontage sowie der Spielzeugbaustein-Montage präsentiert.

In-view-Projektion als Baseline

Bei der In-view-Projektion (vgl. Kapitel 2.3.3 und 3.3.2) werden die Arbeitsinformationen im Blickfeld des Menschen angezeigt, sodass sie von allen Arbeitspositionen rund um den Demonstrator einzusehen sind. Die Anweisungen werden statisch in der Mitte des Demonstrators platziert, wobei mittels der Gesteninteraktion auf die nächste Information „weitergeblättert“ werden kann (Abbildung 88).



Abbildung 88 | In-view-Projektion in der Fan-Cowl-Montage (Eigendarstellung)

Aus der Demonstration sowie aus der Durchführung der Experimente wurde erkannt, dass sich diese Art der Informationspositionierung gut für die Anzeige von Montageanweisungen eignet, da die Informationen immer an der gleichen Stelle projiziert werden und für den Menschen schnell ersichtlich sind. Dabei wird der Person vermittelt, dass sie die „Führende“ im Prozess ist. Der große Nachteil ist, dass hier die Position der Entnahmestelle von Objekten sowie auch die exakte Montageposition nur textlich angezeigt wird. Daraus resultieren lange Suchzeiten für die Positionsfindung bzw. müssen andere Systeme (z. B. ein statisches Lasersystem) für die exakte Anzeige eingesetzt werden. Die In-view-Arbeitsanweisungen stellen dabei den Versuch dar, die Anzeige der Arbeitsinformationen wie am PC-Terminal statisch umzusetzen, diesen aber durch die Bereitstellung direkt in der Montage abzulösen.

In den folgenden Kapiteln werden die dynamischen Möglichkeiten des Projektionssystem verwendet und auf die erweiterte Positionierungsart der In-situ-Projektion eingegangen sowie die dynamische Informationsführung durch den Montageprozess demonstriert.

4.3.1 Dynamische In-situ-Projektion der Montageposition

Im Laufe des Gestaltungsprozesses wurden verschiedene Varianten der In-situ-Projektion und Informationsführung in die jeweiligen Demonstratoren integriert und die Möglichkeiten in diversen Experimenten evaluiert. Die dynamische In-situ-Projektion und Informationsführung zur Montageposition (Abbildung 89) wurden am Demonstrator der Fan-Cowl-Montage demonstriert und in verschiedenen Experimenten evaluiert.

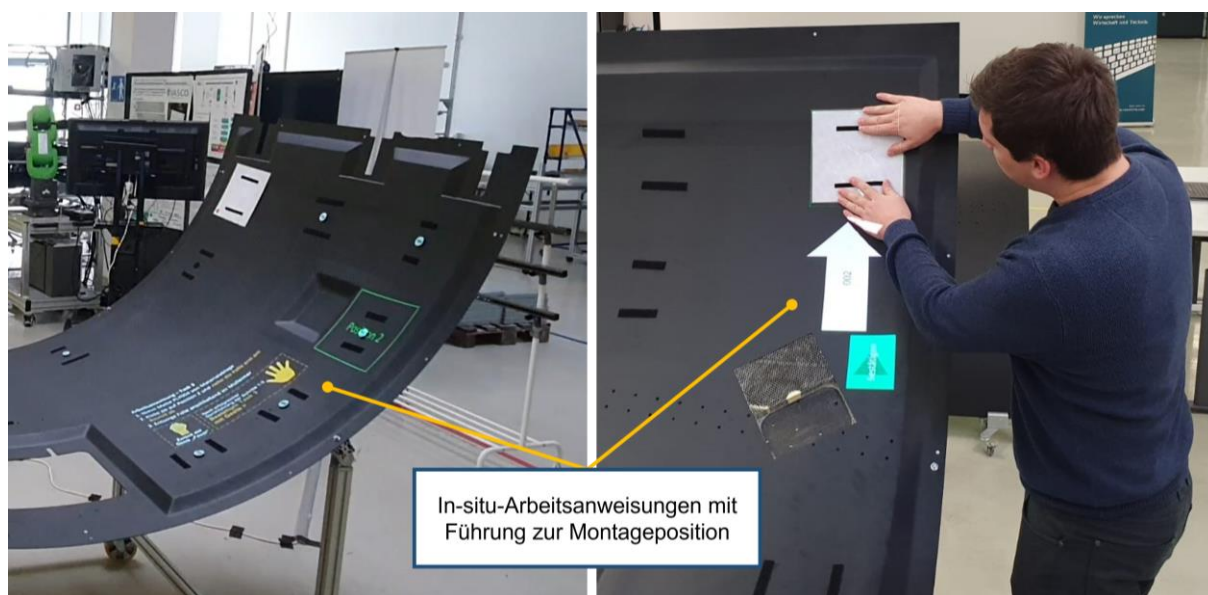


Abbildung 89 | In-situ-Projektion von Arbeitsanweisungen in der Fan-Cowl-Montage (Eigendarstellung)

In Abbildung 89 links wurde die In-situ-Projektion von Arbeitsanweisungen inklusive der Anzeige der exakten Montageposition über das dynamische Projektionssystem umgesetzt (vgl. Kapitel 5.2.2). Bei Abbildung 89 rechts wurde vom Projektionssystem in Form von Pfeilen angezeigt, wo die exakte Montageposition am Bauteil zu finden ist. Die Kontur und exakte Montageposition wurden hier mit einem statischen Laser umgesetzt (vgl. Kapitel 5.2.1). Beim Fan-Cowl-Montageprozess wurde die dynamische Informationsführung so durchgeführt, dass durch Anzeige von textlicher/grafischer Arbeitsinformation der Mensch zum Nahefeld der Montageposition geführt wurde (dynamische Informationsführung zur Montageposition). Gleichzeitig wurde die Person durch In-situ-Projektion der exakten Montageposition beim Montieren der Bauteile unterstützt. Ein Ausschnitt des Montageprozesses an der Fan-Cowl-Montage wird in Abbildung 90 gezeigt.

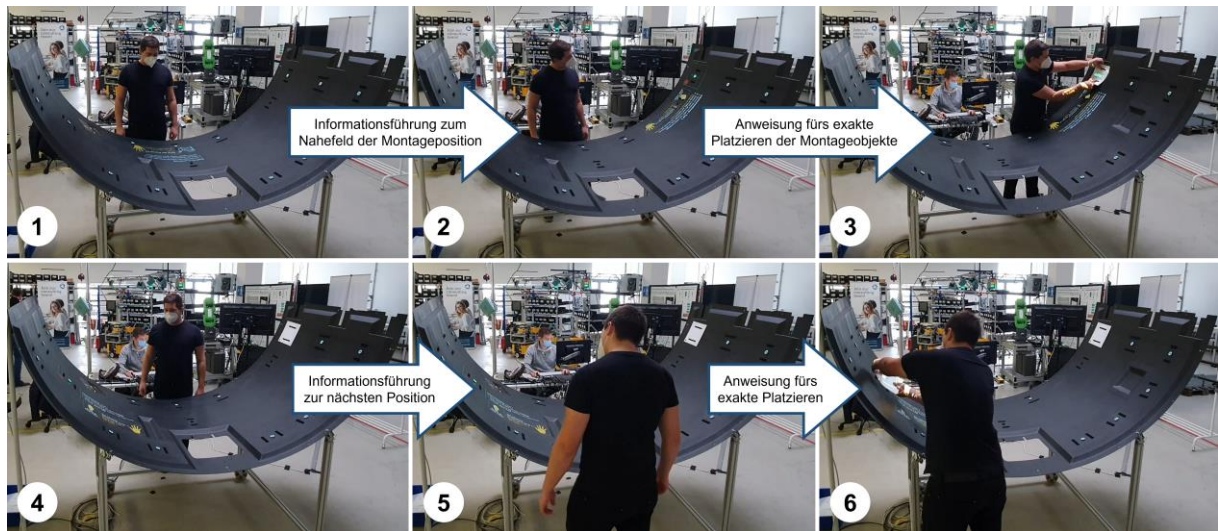


Abbildung 90 | Dynamische Informationsführung und In-situ-Projektion (Eigendarstellung)

Dabei wird in (1) die Start-Arbeitsinformation direkt in der Montage angezeigt. Nach Ausführung der Gesteninteraktion wird der Mensch mittels Anzeige der Arbeitsinformation ins Nahefeld der nächsten Montageposition geführt (2). Der Mensch liest die Anweisungen, holt die zu montierenden Objekte aus der Materialablage (nicht in der Abbildung dargestellt) und setzt die Montagetätigkeiten, unterstützt durch die exakte Projektion der Montageposition, um (3). Nach Bestätigung des Tasks mit einer Geste führt die Information zur nächsten Position (4, 5) und unterstützt bei der weiteren Montage (6).

Eine weitere Umsetzung der In-situ-Projektion stellt die Anzeige mittels farblicher Lichtpunkte dar. Dabei wurde ein Demonstrator für die Montage von Nutzensteinen in C-Profilen aufgebaut (Abbildung 91) und ein Montageprozess mit neun Montagepositionen demonstriert.



Abbildung 91 | Anzeige der Montageposition durch Lichtpunkte und Lichtvorhang bei erfolgreicher Durchführung (Eigendarstellung)

Die Montagepositionen wurden vom Projektionssystem mit Lichtpunkten in situ an den C-Profilen angezeigt und die erfolgreiche Durchführung des Montageprozesses durch einen grünen „Lichtvorhang“ angezeigt. Durch die Demonstration konnte festgestellt werden, dass sich die Anzeige mittels farbiger Lichtpunkte besonders für die schnelle Auffindung von Positionen oder Objekten eignet. Bei der Gestaltung ist die Herausforderung, die Genauigkeit der Positionierung zu evaluieren sowie festzustellen, welche Zusatzinformationen zum Montageprozess benötigt werden. Diese Zusatzinformationen können am Hallenboden oder auf dem Montageobjekt projiziert werden.

4.3.2 Dynamische In-situ-Projektion der Entnahme- und Montageposition

Eine Erweiterung der In-situ-Projektion stellt die Anzeige und Informationsführung zur Entnahmeposition und Montageposition dar, welche im Use Case der Spielzeugbaustein-Montage umgesetzt wurden (Abbildung 92).



Abbildung 92 | In-situ-Projektion und Informationsführung zur Entnahme- und Montageposition bei der Spielbaustein-Montage (Eigendarstellung)

Dabei werden mithilfe dynamischer Animation durch bewegende Elemente (vgl. Kapitel 3.3.2) und kontinuierlicher Bewegung des Projektionsstrahls die Bauteile an den Bereitstellungsflächen angezeigt. Mit den Experimenten konnte festgestellt werden, dass so die Lokalisierungszeit der Bauteile schneller ist, als wenn die Bauteile rein durch textliche Informationen gefunden werden müssen. Des Weiteren wurde durch die In-situ-Projektion von Lichtflächen die exakten Montagepositionen in der richtigen Lage und Orientierung im Raum angezeigt. Die Informationen zur Interaktion wurden am Anfang des Prozesses am

Hallenboden als Single-Point-Projektion und später in Kombination mit den Lichtflächen als Multi-Point-Projektion angezeigt. Aus der Demonstration konnte festgestellt werden, dass sich die In-situ-Projektion und die dynamische Informationsführung mit bewegenden Animationen besonders für das schnelle Auffinden von Positionen und Bauteilen im zweidimensionalen (Hallenboden) und zweieinhalbdimensionalen (Montageobjekte) eignet.

4.4 Nutzeradaptive und natürliche Gesteninteraktion

Um die Mensch-Maschine-Interaktion umzusetzen und den Menschen sowie dessen Posen und Gesten im Montageraum zu erfassen, wurde das visuelle Objekterkennungssystem (Kamerasystem und Objekterkennungsalgorithmus) in den Demonstratoraufbau integriert. Dazu wurde die Kamera über dem Demonstrator platziert, mit dem Medien- und Steuerungsserver vernetzt und die Programmcodes zur Ausführung der Objekterkennung am Medienserver installiert. Damit können anschließend die Objekte rund um den Demonstrator in Echtzeit erfasst werden und die Befehle (z. B. der Gesten) direkt zur Steuerung des Projektionssystems und des Spiegelkopfes weitergeleitet werden. Um die Funktionsweise sowie die praktische Relevanz aufzuzeigen, werden im Folgenden die Erkennung des Menschen und dessen Position sowie die Ausführungen der Gesten demonstriert.

Die Objekterkennungsklassen („Mensch“, Gesten „offene Hand“ und „Faust“) (vgl. Kapitel 3.3.3) wurden mit einem Hochleistungsrechner trainiert und anschließend für die Echtzeit-Objekterkennung zur Erkennung der Objekte zu Verfügung gestellt (Abbildung 93).



Abbildung 93 | Mensch-, Positions- und Gestenerkennung, in Anlehnung an Rupprecht et al. (2021)

Dabei wird der Mensch (grüne Bounding-Box mit Label „Human“) sowie die Gesten (Bounding-Boxen mit den Labels „Fist“ und „Five“) in der Montagestation vom Algorithmus erkannt und der Schwerpunkt (grüner Punkt innerhalb der Bounding-Box) als Position mit x- und y-Koordinaten erfasst. Die Objekterkennung agiert dazu in Echtzeit mit einer Frame Rate von 30 fps und kann dadurch für die intuitive Steuerung der Systeme (z. B. Weiterschalten der Arbeitsinformationen) verwendet werden.

4.4.1 Nutzeradaptive Interaktion durch Positionserkennung des Menschen

Um die Steuerung des Projektionssystem mit der Position bzw. Bewegung des Menschen im Arbeitsraum zu demonstrieren, wurde ein praktischer Test durchgeführt, bei dem die Montagestation in vier Quadranten unterteilt wurde (Abbildung 94).



Abbildung 94 | Veranschaulichung: Projektion folgt dem Menschen (Eigendarstellung)

Dabei wurde programmtechnisch festgelegt, dass, wenn der Schwerpunkt der Objektklasse „Mensch“ von der Objekterkennung in einem der Quadranten ermittelt wurde, die Projektion der Arbeitsanweisung im Nahefeld des Menschen (Offset von 500 Millimetern) angezeigt wird. Wechselt der Mensch nun in einen anderen Quadranten, dann folgt die Information dem Menschen. Diese Art der Projektionsverfolgung kann eingesetzt werden, um Arbeitsinformationen jeweils dynamisch im Nahefeld des Menschen zu projizieren. Dabei konnte jedoch festgestellt werden, dass nach dem Wechsel des Quadranten die Projektion ordnungsgemäß folgt, jedoch bei kleinen Änderungen der Position des Schwerpunktes innerhalb des Quadranten die Projektion ebenso versucht, dem Menschen zu folgen, und „wackelt“. Demnach muss bei der Gestaltung darauf geachtet werden, dass die Projektionsverfolgung mittels Zusatzaktionen (z. B. einem Gestenbefehl) ein- und ausgeschaltet werden kann. Dieser Ansatz bezieht sich darauf, dass der Mensch, der/die „Führende“ im Montageprozess ist und durch den Prozess leitet und die Informationen (je nach Bedarf) dementsprechend dem Menschen folgen. Mit der Demonstration wurde gezeigt, dass die Projektion durch die Erkennung und Bewegung des Menschen gesteuert werden kann, im Folgenden wird jedoch nur mehr auf die Möglichkeiten der Interaktion mit Gesten fokussiert.

4.4.2 Nutzeradaptive Gesteninteraktion

In den weiteren Demonstrationen wird vor allem der Aspekt eingesetzt, dass das System den Menschen durch gezielte Projektion der Arbeitsinformationen durch den Prozess führt, der Mensch jedoch durch die Interaktion mit den Gesten die Weiterschaltung zur nächsten Information oder Montageposition steuert. Dazu wurde der definierte Arbeitsprozess im Projektionssystem eingestellt und die Positionen der Arbeitsinformationen sowie die Verzerrungen und Orientierungen als Pre-Sets abgespeichert. Des Weiteren wurde mittels Programmcode die Gesteninteraktion in den Prozess integriert, sodass die

Arbeitsinformationen (Pre-Sets) damit abgerufen werden können. Das Setting wurde am Demonstrator integriert und diverse Experimente zur Informationspositionierung und -führung damit umgesetzt (Abbildung 95).



Abbildung 95 | Interaktion mittels Handgesten inklusive Feedbackinformation (Eigendarstellung)

Die Information über die möglichen Gesteninteraktionen wurden dem Menschen in den Arbeitsanweisungen angezeigt (vgl. Kapitel 4). Die Geste „offene Hand“ wurde dazu eingesetzt, um die Arbeitsinformationen weiterzuschalten, die Geste „Faust“, um die Anweisungen zurückzuschalten. Für die Ausführung der Geste wird die Hand über dem Demonstrator ausgestreckt und die Geste von oben mithilfe des visuellen Interaktionssystems erkannt. Nach erfolgreicher Durchführung der Gesteninteraktion wird eine Feedbackinformation projiziert, damit der Mensch zu jedem Zeitpunkt des Prozesses Bescheid weiß, wie der Status ist bzw. was das System aktuell ausführt.

Auch bei der Spielzeugbaustein-Montage wird die Interaktion mittels Gesteninteraktion umgesetzt und damit das dynamische Projektionssystem gesteuert. Die Informationen zur Interaktion werden dazu vom Projektionssystem am Hallenboden in Form von kurzen Textanweisungen und Piktogrammen angezeigt. Durch Ausführung der Geste werden die Arbeitsinformationen laut einem definierten Montageprozess abgerufen und mittels Spiegelsystem zur nächsten Position geführt (Abbildung 96).

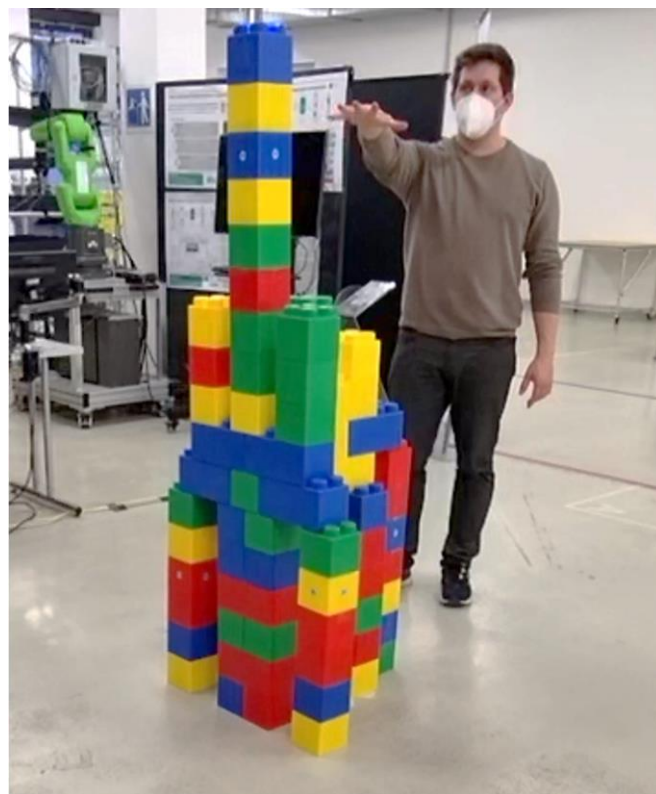


Abbildung 96 | Gesteninteraktion Spielzeugbaustein-Demonstrator (Eigendarstellung)

Im Zuge der Demonstrationen konnte festgestellt werden, dass die implementierte Gesteninteraktion für die Interaktionen mit den Arbeitsinformationen stabil funktioniert und eine intuitive Bedienung ermöglicht. Der große Vorteil dabei ist, dass die Steuerung mittels Gesten keine tragbaren Zusatzgeräte benötigt und von jeder Stelle der Montage (sofern im FOV der Kamera) aus durchgeführt werden kann. Die detaillierten Ergebnisse der Evaluierung der Gesteninteraktion werden folgend in den Evaluierungsexperimenten noch näher dargestellt. Abschließend kann zusammengefasst werden, dass sich die Interaktion mit natürlichen Gesten unter dem Einsatz der Echtzeit-Objekterkennung für den Einsatz im großen Arbeitsbereich der industriellen Großgeräte- und Baustellenmontage eignet und für die Steuerung des Projektionssystems sowie zur Interaktion mit den projizierten Arbeitsanweisungen optimal eingesetzt werden kann.

5. Evaluierung des Spatial Augmented Reality Assistenzsystems

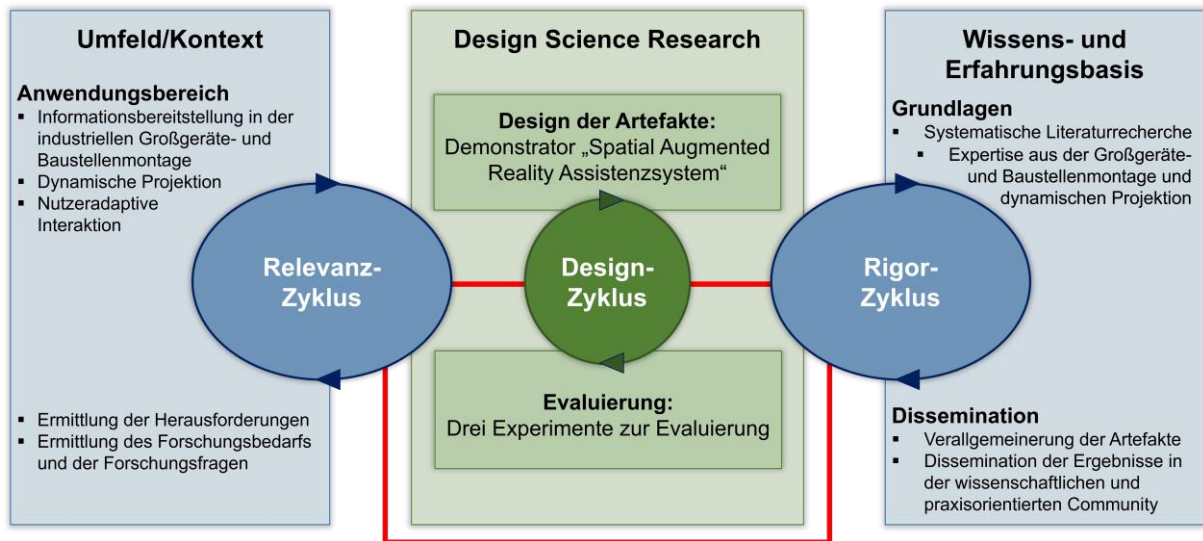


Abbildung 97 | Evaluierung mittels drei Experimenten, in Anlehnung an Alan R. Hevner et al. (2004)

Das **Spatial Augmented Reality Assistenzsystem mit dynamischer In-situ-Projektion und nutzeradaptiver Gesteninteraktion** wird im Folgenden mit zwei Demonstratoren und drei aufeinander aufbauenden Experimenten sowie mit geeigneten Evaluierungsmethoden bewertet, um die Nützlichkeit, Effizienz sowie Qualität des Design-Artefakts (Alan R. Hevner et al., 2004) (vgl. Abbildung 97) sowie die Gebrauchstauglichkeit (DIN EN ISO 9241-11:2018-11) festzustellen. Dabei werden wissenschaftlich anerkannte Methoden im Bereich der Nutzerakzeptanz (Davis et al., 1989), Usability (Brooke, 1996) sowie der kognitiven Arbeitsbelastung (Hart & Staveland, 1988) herangezogen, um den Mehrwert für Mitarbeiter*innen zu ermitteln. Außerdem werden Methoden zur Ermittlung der detaillierten Montage- und Taskzeiten und Fehlerrate (Funk, 2016; Funk, Kosch et al., 2015) zur Evaluierung des Mehrwerts für Unternehmen und zur Prozessverbesserung im Vergleich zu herkömmlichen Systemen (PC-Terminal, Tablet) verwendet.

5.1 Evaluierungsmethoden

Für die Evaluierung werden fundierte Methoden im Bereich Usability, Technologieakzeptanz und kognitiver Arbeitsbelastung sowie der Erhebung der Prozesseffizienz und Fehlerrate eingesetzt. Dabei werden sowohl quantitative Methoden (z. B. Ermittlung der Taskzeit durch Videoanalyse) als auch qualitativen Methoden wie Kurzinterviews und offene Fragen in den Fragebögen verwendet.

Um allgemeine Informationen über die Teilnehmer*innen zu erhalten und um die Evaluierungsergebnisse besser einschätzen zu können, wurde im Fragebogen das Alter, Geschlecht, Beschäftigungsart (Schüler*in, Student*in, berufstätig) sowie die Affinität zur Verwendung von digitalen Tools abgefragt. Bei der Durchführung der Evaluierung wurde stets darauf geachtet, eine anonymisierte Datenerhebung anzuwenden und für die Videoaufzeichnung und Datenverarbeitung die Zustimmung der einzelnen Teilnehmer*innen

einzuholen. Dabei wurde zum Schutz der Anonymität beim Fragebogen sowie bei der Dateiablage eine anonymisierte Kennung verwendet.

5.1.1 Usability, Technologieakzeptanz und kognitive Arbeitsbelastung

Zur Bewertung der Benutzerfreundlichkeit bzw. Usability wird der „System Usability Scale“ (SUS) (Brooke, 1996), zur Evaluierung der Nutzerakzeptanz neuer Technologien das „Technology Acceptance Model“ (TAM) (Davis et al., 1989) und für die Erfassung der kognitiven und physischen Belastung, Leistungsfähigkeit, Aufwand und Frustrationslevel die vereinfachte NASA-RTLX-Methode (Hart, 2006; Hart & Staveland, 1988) eingesetzt.

System Usability Scale (SUS)

Jakob Nielsen beschreibt Usability mit den Aspekten der Lernfähigkeit, Effizienz, Gedächtnisleistung, Fehlertoleranz und Zufriedenheit eines Benutzers während der Benutzung von Produkten und Dienstleistungen (Nielsen, 1993). Die Norm DIN EN ISO 9241-11:2018-11 beschreibt Usability (deutsch „Gebrauchstauglichkeit“) als das „Ausmaß, in dem ein System, Produkt oder eine Dienstleistung von bestimmten Benutzern in einem bestimmten Nutzungskontext verwendet werden kann, um bestimmte Ziele effektiv, effizient sowie zufriedenstellend zu erreichen“.

Der SUS ist eine fundierte Methode, um die Gebrauchstauglichkeit von Produkten oder Dienstleistungen speziell für die Gegebenheiten der Industrie abzufragen (Brooke, 1996). Der Fragebogen besteht aus zehn Items, die mit einer fünfstufigen Likert-Skala beantwortet werden. Das Ergebnis ist ein „Total SUS Score“ (Abbildung 98), welcher nach den Aspekten „Akzeptanz – Ränge“, „Notenskala von A–F“ (A bedeutet „gute Usability“) und einer „Adjektivbewertung“ interpretiert wird (Bangor et al., 2009).

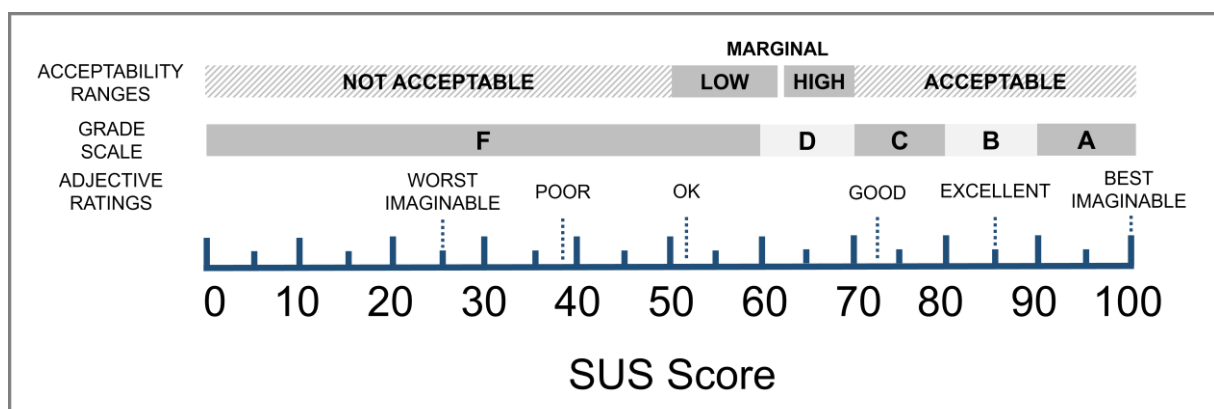


Abbildung 98 | Akzeptanzbereiche, Notenskala und Adjektivbewertungen von SUS-Ergebnissen, in Anlehnung an Bangor et al. (2009)

Technology Acceptance Model (TAM)

Das „Technology Acceptance Model“ dient als praxiserprobtes Modell zur Bewertung der Nutzerakzeptanz von eingesetzter (neuer) Technologie. Davis et al. (1989) entwickelten das TAM, um die Benutzerakzeptanz besser vorherzusagen, zu erklären sowie zu erhöhen und um zu verstehen, warum Menschen Computersysteme eher akzeptieren oder ablehnen (Davis et al., 1989). Die Technologieakzeptanz wird anhand eines Fragebogens mit 13 Fragen und einer Skala von 1 bis 6 (--- bis +++) bewertet (Tabelle 18).

Tabelle 18 | TAM-Fragebogen, in Anlehnung an Davis et al. (1989)

Fragen	Ich stimme auf keinen Fall zu	Ich stimme nicht zu	neutral	Ich stimme zu	Ich stimme voll zu
Die Nutzung des Assistenzsystems ist eine neue Erfahrung für mich.					
Die Nutzung des Assistenzsystems ist nicht mit dem vergleichbar, was ich bisher gemacht habe.					
Die Nutzung des Assistenzsystems unterscheidet sich von anderen Erfahrungen, die ich gemacht habe.					
Die Nutzung des Assistenzsystems ist eine neue Arbeitserfahrung für mich.					
Die Anwendung des Assistenzsystems in der Praxis würde meinen eigenen Werten widersprechen.					
Die Nutzung des Assistenzsystems passt nicht zu meiner Sicht der Welt.					
Die Verwendung des Assistenzsystems steht im Widerspruch zu dem, wofür Computer meiner Meinung nach eingesetzt werden sollten.					
Die Verwendung des Assistenzsystems ist für eine Person mit meinen Wertvorstellungen bezüglich der Rolle von Computern nicht geeignet.					
Die Nutzung des Assistenzsystems widerspricht meinen Wertvorstellungen, wie gearbeitet werden sollte.					
Die Verwendung eines solchen Assistenzsystems würde zu meiner bevorzugten Arbeitsroutine passen.					
Das Assistenzsystem würde es mir ermöglichen, in der von mir bevorzugten Weise zu arbeiten.					
Die Nutzung des Assistenzsystems würde gut zu meiner Arbeitsweise passen.					
Die Nutzung des Assistenzsystems würde meiner bevorzugten Arbeitsmethode entsprechen.					

NASA – Raw Task Load Index (NASA-(R)TLX)

Der „NASA Task Load Index“ (NASA-TLX) ist ein anerkanntes Bewertungsinstrument zur Evaluierung der Arbeitsbelastung von spezifischen Arbeitsaufgaben oder eines Arbeitssystems. Im Fragebogen werden Aspekte zu mentaler, physischer und temporärer Belastung sowie der Leistungsfähigkeit, Bemühung/Aufwand sowie des Frustrationslevels einer Arbeitsaufgabe ermittelt (Hart & Staveland, 1988). In Tabelle 19 werden die sechs Fragen sowie die fünfstufigen Antwortmöglichkeiten dargestellt.

Tabelle 19 | NASA-RTLX zur Bewertung von Assistenzsystemen, in Anlehnung an Hart und Staveland (1988)

Fragen	Sehr hoch	hoch	Mittel	wenig	Sehr wenig
Wie geistig anspruchsvoll war die Aufgabe?					
Wie körperlich anstrengend war die Aufgabe?					
Wie eilig oder gehetzt war das Arbeitstempo?					
Wie erfolgreich waren Sie bei der Erfüllung der Ihnen gestellten Aufgabe?					
Wie sorgfältig/hart mussten Sie arbeiten, um Ihr Leistungsniveau zu erreichen?					
Wie unsicher, entmutigend, irritiert, gestresst und verärgert waren Sie?					

Der NASA-(R)TLX-Score gibt eine gute Auskunft über die kognitive Arbeitsbelastung des Arbeitssystems und wird, wie auch von Funk (2016) in der Kleingerätemontage eingesetzt, in der Großgerätemontage zur Bewertung des Spatial Augmented Reality Assistenzsystems verwendet.

5.1.2 Montage- und Taskzeit (General Assembly Task Model)

Zur Evaluierung der Leistung und Effizienz des Montageprozesses sowie der Qualität der eingesetzten Informationsbereitstellung wird die „**Task Completion Time**“ (TCT) nach dem „**General Assembly Task Model**“ (GATM) nach Funk, Kosch et al. (2015) verwendet. Dabei werden die jeweiligen ausgeführten Montageschritte (Tasks) in vier standardisierte Phasen gegliedert und die zugehörigen Zeiten ermittelt. Durch die Ermittlung und Interpretation der einzelnen Taskzeiten je Phase, können auf die Qualität und Verständlichkeit von Arbeitsanweisungen geschlossen sowie die Verbesserungen durch Einsatz des dynamischen Projektionssystems evaluiert werden. Dadurch eignet sich die TCT besser für die Analyse von Montageprozessen und Montagetätigkeiten als die reine Ermittlung der Gesamtmontagezeit (Funk, Kosch et al., 2015). Das GATM ist angelehnt an die Philosophie von „Methods Time Measurement“ (MTM) mit dessen fünf Grundbewegungen „Hinlangen – Greifen – Bringen – Fügen – Loslassen“ (Bokranz & Landau, 2012; MTM ASSOCIATION e. V., 2020) und macht eine Evaluierung von verschiedenen Assistenzsystemen vergleichbar (Funk, 2016; Funk, Kosch et al., 2015). Beim GATM wird jeder Montageschritt den folgenden vier Phasen zugeteilt (Funk, Kosch et al., 2015):

- Lokalisieren Bauteil/Entnahmeposition („locate part“)
- Aufnehmen („pick“)
- Lokalisieren Montageposition („locate position“)
- Montieren („assemble“)

Um Fehlinterpretationen zu vermeiden und die Montageschritte den Phasen exakt zuweisen zu können, bedarf es einer genauen Definition jeder einzelnen Phase (Blattgerste et al., 2017). Die Definition nach Blattgerste et al. (2017) wird in Tabelle 20 herangezogen und um die Gegebenheiten der Baustellenmontage erweitert.

Tabelle 20 | Phasendefinition GATM für die Baustellenmontage, in Anlehnung an Blattgerste et al. (2017)

GATM (Funk, Kosch et al., 2015)	Definition der Phasen nach (Blattgerste et al., 2017) inklusive den Gegebenheiten in der Baustellenmontage
Lokalisieren Bauteil/Entnahmeposition („locate part“)	Beginn des Lesens der Arbeitsanweisung bis zum Eintritt der Finger des Menschen in den korrekten Behälter des Montagebauteils. Baustellenmontage: Bis zum Austrecken des Armes zum Erlangen des korrekten Montagebauteils. Da die Materialablage meist räumlich zum Montageobjekt entfernt ist, ist auch der Weg zur Materialablage in diese Phase inkludiert.
Aufnehmen („pick“)	Aufnehmen des Montagebauteils, sodass es zur Montageposition transportiert/bewegt werden kann.
Lokalisieren Montageposition („locate position“)	Lokalisieren der Montageposition am Montageobjekt, Weg zum Montageobjekt und anschließendes „Bringen“ des Bauteils zur korrekten Montageposition (Bauteil schwebt im geringen Abstand über [erster] Montageposition) Baustellenmontage: Die Wegstrecke von der Materialablage bis zum Montageobjekt ist inkludiert.
Montieren („assemble“)	Montieren der aufgenommenen Bauteile auf die jeweiligen (weiteren) Montagepositionen laut Arbeitsanweisung. Dazu gehören z. B. folgende Tätigkeiten: Verwendung von benötigtem Werkzeug, Ausrichtung, Qualitätskontrolle etc.

Die Zusammensetzung der Gesamtmontagezeit aus den einzelnen ausgeführten Taskzeiten der jeweiligen Phasen nach dem GATM wird in folgender Formel (F1.1) dargestellt (Funk, Kosch et al., 2015):

$$t_{montagezeit} = n \left(\underbrace{t_{lokalisieren_{Bauteil}} + t_{aufnehmen} + t_{lokalisieren_{Montageposition}} + t_{montieren}}_{t_{montage_task}} \right) \quad (F 1.1)$$

Die Gesamtmontagezeit $t_{montagezeit}$ ist die Zeit, in der ein komplettes Produkt montiert bzw. alle Montagetasks des definierten Montageprozesses durchgeführt wurden. Dabei ist „n“ die Anzahl der jeweiligen Montagetasks. $t_{lokalisieren_Bauteil}$ ist die Zeit zum Lesen der Arbeitsanweisung, für das Aufsuchen des Materialablageortes sowie Finden des richtigen Bauteils laut Arbeitsanweisung. Die Zeit zur Aufnahme des Bauteils wird mit $t_{aufnehmen}$ angegeben. Wenn beim Montageschritt kein Bauteil aufgenommen werden muss, dann sind $t_{lokalisieren_Bauteil}$ und $t_{aufnehmen}$ mit der Zeit = 0 zu versehen. Die Zeit, um die Montageposition am Montageobjekt zu ermitteln, sich zur Montageposition zu bewegen sowie das Montagebauteil an die (erste) Montageposition zu bringen (kurz vor physischen Kontakt) wird unter $t_{lokalisieren_Montageposition}$ ermittelt. $t_{montieren}$ ist die Zeit, um die Montageaufgabe laut Arbeitsanweisung durchzuführen bzw. bei gleichzeitiger Aufnahme mehrerer Bauteile die weiteren Teile zu montieren (Funk, Kosch et al., 2015).

Für die Zuweisung der Montageschritte zu den vier Phasen des GATM und die Ermittlung der zugehörigen Zeiten wird der Montageprozess per Video aufgezeichnet und für die Auswertung

das Open-Source-Video-Annotation-Programm „ANVIL2“²⁶ (Abbildung 99) verwendet (Pimminger et al., 2020).

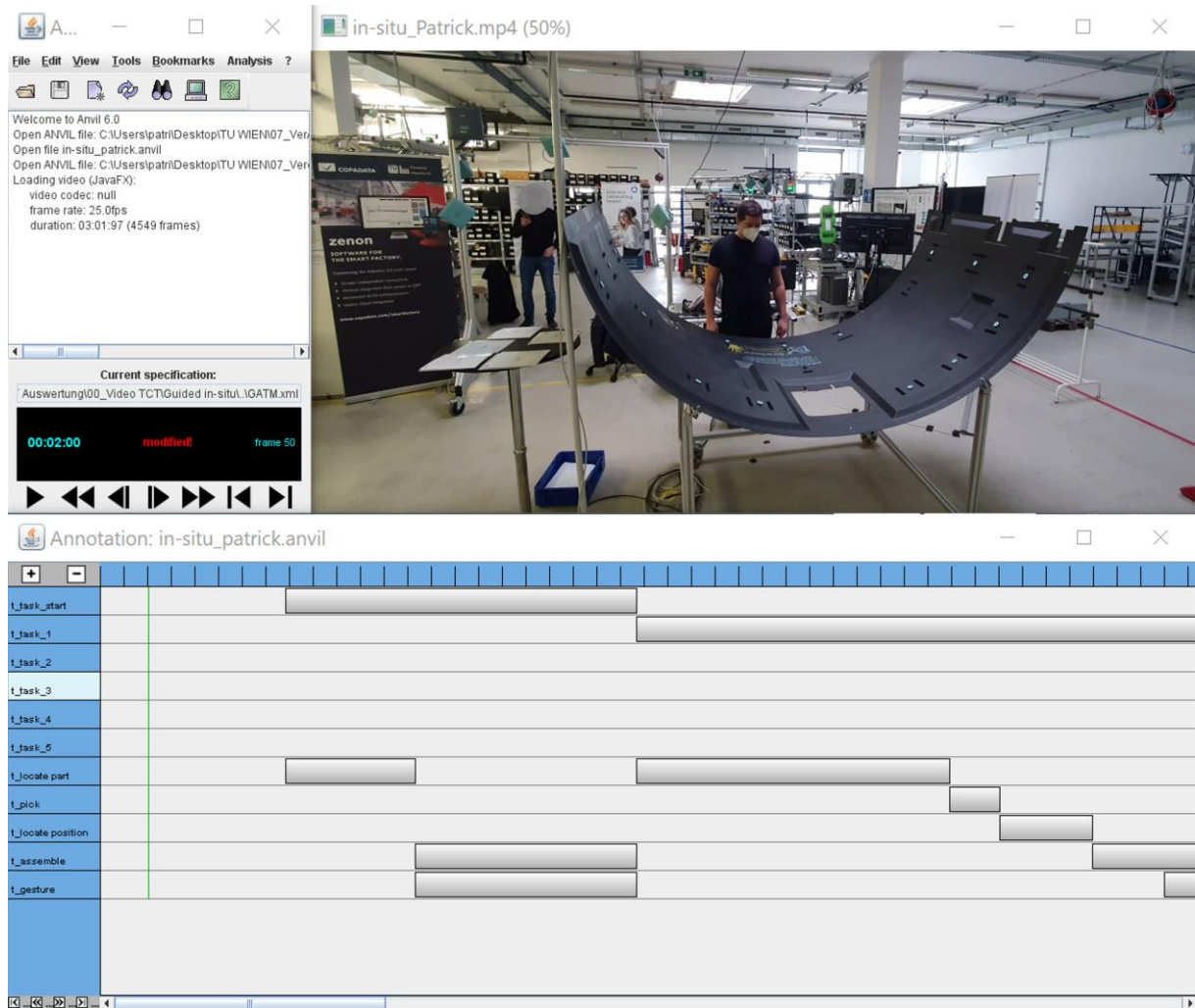


Abbildung 99 | Video-Annotation-Tool „ANVIL2“ (Eigendarstellung)

Mit dem Video-Annotation-Tool werden die einzelnen Montagetaasks sowie die vier Phasen nach GATM angelegt, um bei der Analyse die einzelnen Videosequenzen dementsprechend zuweisen zu können. Dabei wird gleichzeitig die jeweilige Dauer der Phasen annotiert, welche anschließend als Dateiformat CSV bzw. Textdatei exportiert wird. Die Daten werden anschließend in MS Excel importiert und die jeweiligen Mittelwerte über die Teilnehmer*innen sowie die Summenwerte der einzelnen Phasen je Experiment ausgewertet. Die Daten werden anschließend grafisch in einem Diagramm aufbereitet.

Analyse und Interpretation der TCTs

Zur Analyse der jeweiligen TCTs werden die Phasen in **aufgabenabhängige** ($t_{\text{montieren}}$, $t_{\text{aufnehmen}}$) und **aufgabenunabhängige** ($t_{\text{lokalisieren Bauteil}}$, $t_{\text{lokalisieren montageposition}}$) Zeiten unterteilt (Funk, Kosch et al., 2015) (Tabelle 21). Dabei werden die durchschnittlichen TCTs der jeweiligen Phasen herangezogen und für die quantifizierbare Interpretation verwendet. Die

²⁶ <https://www.anvil-software.org>.

aufgabenunabhängigen Zeiten eignen sich dabei besonders für die Evaluierung der Qualität von Arbeitsanweisungen bzw. eingesetzten Informationssystemen sowie für die Bewertung der verschiedenen Informationsführungsarten.

Tabelle 21 | Interpretation der TCTs, in Anlehnung an Funk, Kosch et al. (2015)

Aufgabenabhängige Zeiten	
Achtung: Es können nur ähnlich komplexe Tätigkeiten verglichen werden!	
TCTs laut Phasen des GATM	Interpretation der TCTs
$t_{\text{montieren}}$	Grad der Verständlichkeit der Arbeitsanweisung - für die korrekte Montage des Objekts, - für das Verständnis, welche Werkzeuge benötigt werden und wie diese zu verwenden sind, - für die Orientierung und Art der Montage des Objektes, - für die Kontrolle der Qualität.
$t_{\text{aufnehmen}}$	Kann nicht für die Interpretation herangezogen werden, weil nicht angewiesen werden kann, wie ein Bauteil aufgenommen werden soll.
Aufgabenunabhängige Zeiten	
$t_{\text{lokalisieren}_{\text{bauteil}}}$ $t_{\text{lokalisieren}_{\text{montageposition}}}$	Interpretation der kognitiven Leistung, wie die gelesene Arbeitsanweisung auf die Arbeitstätigkeiten am Arbeitsplatz transferiert und umgesetzt wird. Grad der Verständlichkeit der Arbeitsanweisung/Anzeige, - wo das Montageobjekt zu finden ist und welches Objekt zu montieren ist, - wo die Entnahme- und Montageposition zu finden ist. Vergleich von eingesetzten Technologien, Systemen und Methoden - Informationsbereitstellungssysteme - Informationsführungsarten (in view/in situ)

Die Dauer der **aufgabenabhängigen Zeiten** hängt direkt mit der gestellten Aufgabe und Komplexität der Arbeitsanweisung zusammenhängen. Es kann daraus interpretiert werden, ob in der Arbeitsanweisung klar und verständlich vermittelt wurde, was beim „Montieren“ konkret zu tun ist, wie das Objekt montiert werden soll (Orientierung o. Ä.), welche Werkzeuge erforderlich sind und wie die Qualität der Montage sichergestellt wird. Demnach kann diese Zeit nur im Vergleich zu ähnlich komplexen Tätigkeiten herangezogen werden.

Die Dauer der **aufgabenunabhängigen Zeiten** hängt nicht direkt mit dem der jeweiligen Komplexität der Aufgabe zusammen. Diese Zeit kann dementsprechend interpretiert werden, wie schnell die gelesene Arbeitsanweisung in den räumlichen Kontext (Arbeitsraum, Montageprozess usw.) transferiert werden kann. Im Speziellen kann damit analysiert werden, wie schnell die einzelnen Montagebauteile im Arbeitsraum bzw. am Arbeitsplatz aufgefunden werden ($t_{\text{lokalisieren_Bauteil}}$) und wie schnell die jeweilig zugehörigen Montagepositionen lokalisiert werden können ($t_{\text{lokalisieren_Montageposition}}$). In weiterer Folge kann damit ermittelt werden, ob mit dem jeweiligen Informationsbereitstellungssystem sowie den vorhandenen Informationen der Arbeitsanweisung klar dargestellt wurde, wo das Montageobjekt zu finden ist, welches Objekt zu entnehmen ist und wo dieses exakt montiert werden sollen.

Durch Interpretation dieser Zeiten können die eingesetzten Systeme zur Informationsbereitstellung (PC-Terminal, Tablet, Spatial Augmented Reality Assistenzsystem) sowie die verschiedene Informationsführungsarten (in view/in situ) quantifizierbar verglichen werden.

5.1.3 Fehlerrate

Zur Evaluierung der Leistung und Qualität des Arbeitsprozesses, des eingesetzten Informationssystems und Arbeitsanweisung ist die Ermittlung der Fehlerrate im Montageprozess essenziell (Blattgerste et al., 2017; Funk, 2016; Funk, Kosch et al., 2015). Dabei werden vorab potenzielle Fehlergründe ermittelt und in z. B. einer Fehlerliste (Abbildung 100) eingetragen. Durch Beobachtung der Teilnehmer*innen im Montageprozess werden die Anzahl der Fehler pro Fehlergründe dementsprechend eingezeichnet.

Fehlerliste Use Case „Fan-Cowl“	
Experiment: 2	Teilnehmer: xyz
Variante: 1	Datum: dd.mm.2020
Beobachter: Hans	Uhrzeit: hh:mm
Fehlergründe	Anzahl Fehler
Falsches Bauteil aufgenommen	I
Falsches Bauteil montiert	II
Matte zum falschen Zeitpunkt entfernt	III
Bauteilorientierung nicht korrekt	I
Bauteil vergessen	
„Prozess beenden“ nicht laut vorgelesen	
...	
...	
...	

Abbildung 100 | Fehlerliste des Montageprozesses (Eigendarstellung)

Blattgerste et al. (2017) beschreiben, dass es essenziell ist, in welcher Phase des GATM der Fehler aufgetreten ist, um detaillierter Aussagen über die Qualität und Leistung des Prozesses sowie der eingesetzten Informationsbereitstellung zu erfahren. Dabei werden im Speziellen die Zeiten der GATM-Phasen und deren Ausreißer analysiert und Fehlermöglichkeiten interpretiert. Durch die Fehleranalyse kann Verbesserungspotenzial bei der Gestaltung der Arbeitsinformationen und beim Einsatz des Projektionssystems aufgezeigt sowie auf die Stabilität und Erkennungsgeschwindigkeit der Gesteninteraktion geschlossen werden.

5.1.4 Qualitative Evaluierungsmethoden

Um weitergehend neben den quantifizierbaren Methoden auch die inhärenten Meinungen, Eindrücke und Gefühle der Teilnehmer*innen zu erfahren, wurden offene Fragen in den jeweiligen Fragebögen²⁷ integriert und Kurzinterviews nach jedem Durchgang durchgeführt. Bei den offenen Fragen wurde speziell nach den positiven und negativen Eindrücken bei der Durchführung des Experiments sowie nach den Herausforderungen zu den jeweiligen Systemen (dynamische Informationsbereitstellung, Gesteninteraktion) gefragt. Des Weiteren wurde ermittelt, wie lang die Teilnehmer*innen bereit wären, das jeweilige Assistenzsystem zu verwenden. Bei der Durchführung der Kurzinterviews wurden gezielte Fragen zu den jeweiligen Experimenten und eingesetzten Assistenzsystemen eingesetzt, um die Meinungen zu erfahren. Dabei wurden nach jedem Interview die Kernaussagen dokumentiert und anschließend analysiert.

²⁷ Vorlagen der Fragebögen befinden sich im Anhang.

5.2 Experimente und Use Cases

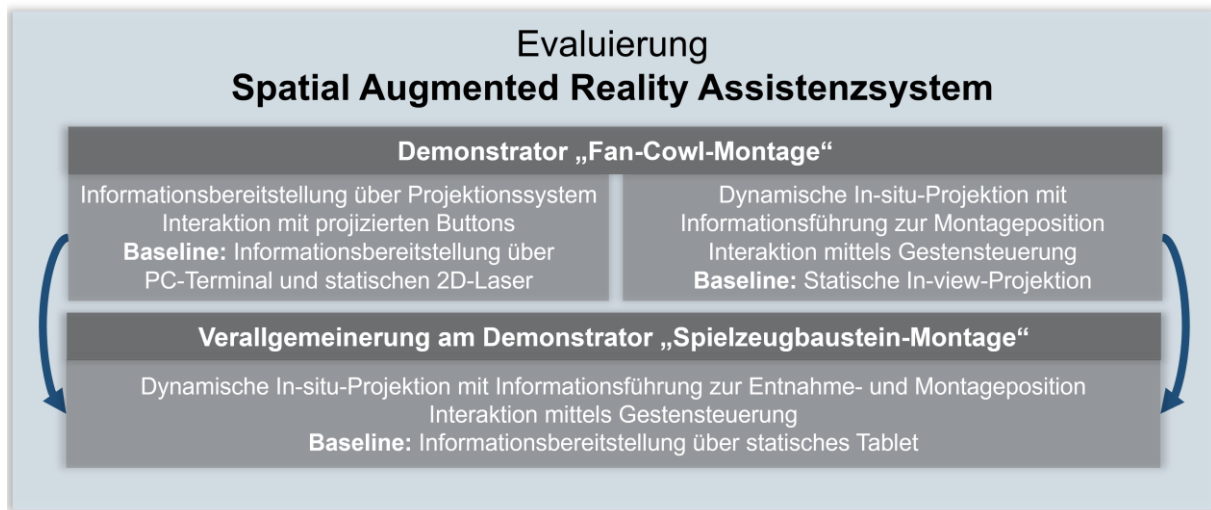


Abbildung 101 | Evaluierungsexperimente und Use Cases (Eigendarstellung)

Die Evaluierung des Spatial Augmented Reality Assistenzsystems wird an den Demonstratoren „Fan-Cowl-Montage“ und „Spielzeugbaustein-Montage“ in der TU Wien-Pilotfabrik durchgeführt (vgl. Abbildung 101). Dabei werden drei aufeinander aufbauende Experimente mit unterschiedlichen Usergruppen durchgeführt und die Erkenntnisse in einem iterativen Designprozess wieder in die Gestaltung des Assistenzsystems integriert.

Im ersten Experiment wird im Use Case „Fan-Cowl-Montage“ das dynamische Projektionssystem zur Anzeige von Arbeitsanweisungen direkt am Bauteil mit herkömmlicher Informationsbereitstellung mittels PC-Terminal verglichen und für die Interaktion projizierte Buttons („Wizard of Oz“) verwendet (Rupprecht et al., 2020).

Im zweiten Experiment wird in der Fan-Cowl-Montage die Informationsbereitstellung ausschließlich mittels Projektionssystem umgesetzt und verschiedene Gestaltungsaspekte der Informationsführung evaluiert. Dabei wird die statische In-view-Projektion (Projektion der Arbeitsschritte in der Mitte der Montage) mit der dynamischen In-situ-Projektion, bei der die Montageposition zusätzlich vom Projektionssystem angezeigt wird und der Mensch mittels Projektion zu Montageposition geführt wird, verglichen (Rupprecht et al., 2021).

Das dritte Experiment dient zur Verallgemeinerung der Ergebnisse aus den ersten beiden Experimenten im weiteren Use Case der Spielzeugbaustein-Montage. Dabei wird das Spatial Augmented Reality Assistenzsystem in einem komplexeren Montageprozess mit 18 Montagetasks eingesetzt und unterstützt bei der Lokalisierung der Montageobjekte an den jeweiligen Entnahmestellen sowie bei der Auffindung der zugehörigen Montagepositionen. Als Baseline dient hierbei die Anzeige von Arbeitsinformationen über ein statisches Tablet, welches auf einem Ständer in der Montage bereitsteht. Folgend werden die eingesetzten Evaluierungsmethoden sowie die einzelnen Experimente beschrieben und anschließend die Ergebnisse präsentiert.

5.2.1 Experiment 1 – Fan-Cowl-Montage: Informationsbereitstellung über ein Projektionssystem, Interaktion mittels projizierten Buttons

Im Experiment 1 wird erstmalig das dynamische Projektionssystem im Use Case „Fan-Cowl-Montage“ zur Informationsbereitstellung eingesetzt und dabei die Arbeitsinformationen direkt in die Montagestation projiziert. Die Interaktion mit dem System wird dabei als Wizard-of-Oz-Methode mit projizierten Buttons umgesetzt, um damit die Interaktion ebenso direkt in der Montage durchzuführen. Bei der Wizard-of-Oz-Methode wird bei Berührung des projizierten Buttons durch die Teilnehmer*innen die nächste projizierte Arbeitsinformation manuell von den Versuchsleiter*innen weitergeschaltet. Als Baseline-Variante des Experiments wird die in der industriellen Baustellenmontage vorherrschende Informationsbereitstellung mit PC-Terminal und 2D-Lasersystem verwendet, um damit einen ersten Vergleich zum State-of-the-Art-System zu erhalten. Abbildung 102 zeigt schematisch den Aufbau der beiden Varianten des Experiments 1 (Rupprecht et al., 2020).

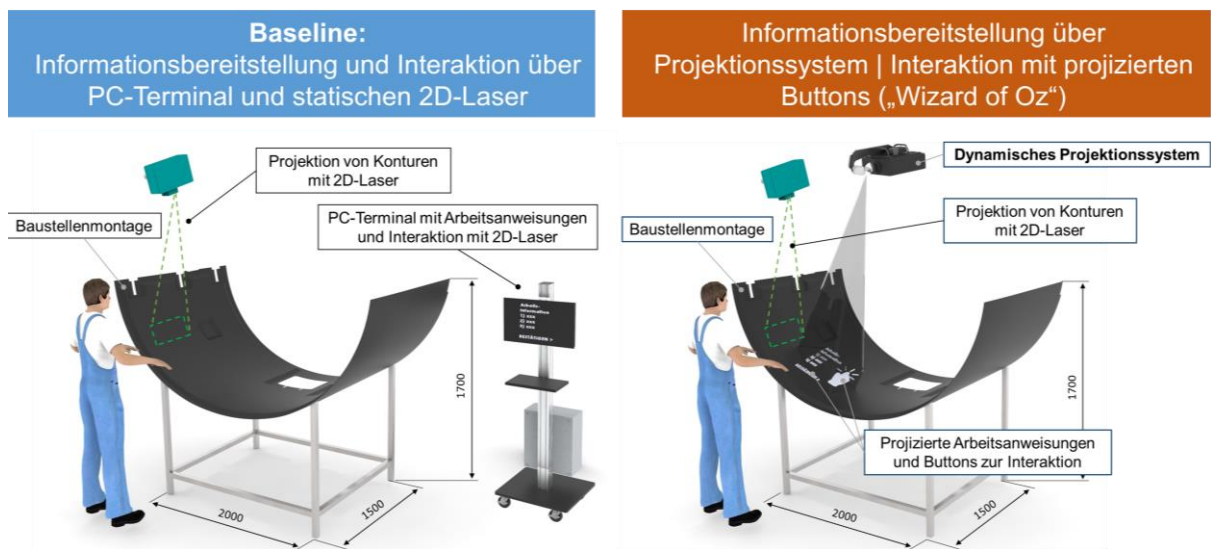


Abbildung 102 | Experiment 1: Dynamische Projektion vs. PC-Terminal, in Anlehnung an Rupprecht et al. (2020)

Die Ziele des Experiments sind (Rupprecht et al., 2020):

- Durchführung des ersten Tests zur Anzeige von Arbeitsinformationen mit dem Projektionssystem direkt in der Baustellenmontage
- Durchführung des ersten Tests zur Interaktion des Informationssystems direkt aus der Montage
- Erhalt erster Erkenntnisse über Usability und Prozesseffizienz des dynamischen Projektionssystems
- Erlangung von Erkenntnissen über die Informationsgestaltung und -führung

Durch die Evaluierung soll folgende Frage beantwortet werden (Rupprecht et al., 2020): „Welchen Mehrwert bietet die dynamische Informationsbereitstellung und Interaktion direkt am Montageplatz gegenüber der Bereitstellung und Interaktion mit PC-Terminal und statischer Laserprojektion?“

Prozessablauf Experiment 1

Um die beiden Varianten vergleichbar zu gestalten, wurde ein vereinheitlichter Prozessablauf mit fünf Montagetasks (Abbildung 103) definiert und die zugehörigen Arbeitsanweisungen je eingesetztem Informationsbereitstellungssystem aufbereitet (Rupprecht et al., 2020).

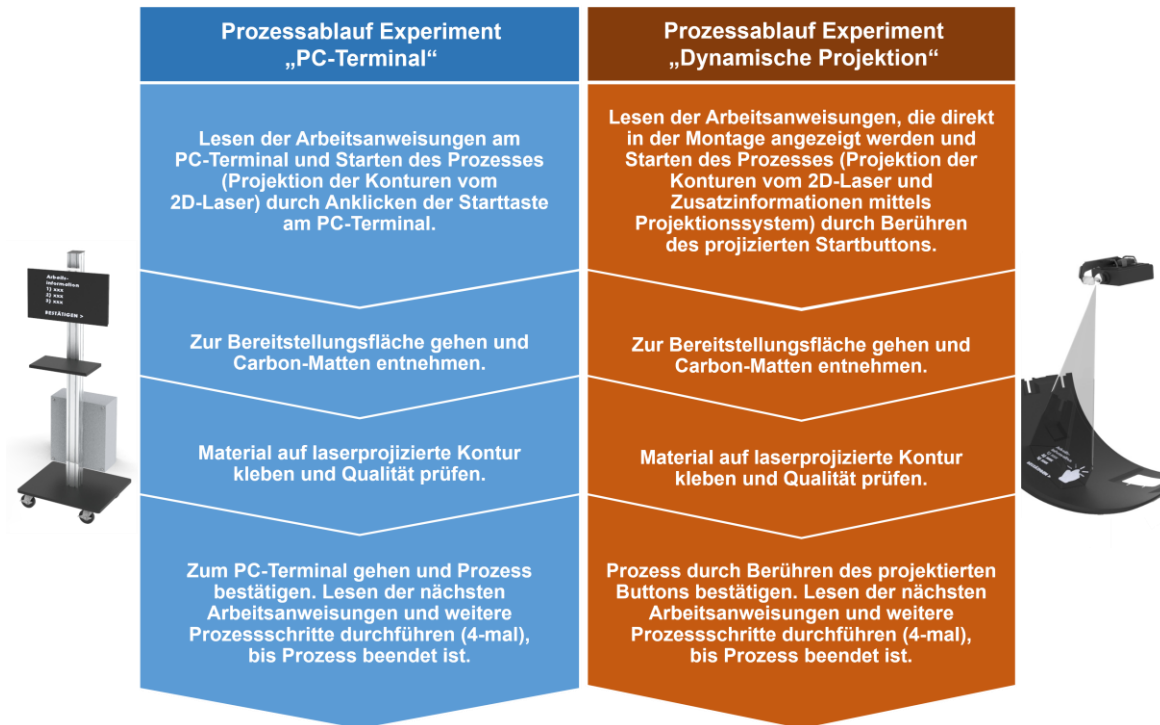


Abbildung 103 | Prozessablauf von Experiment 1, in Anlehnung an Rupprecht et al. (2020)

Der erste Task des Prozessablaufes besteht aus Lesen und anschließender Bestätigung der Start-Arbeitsanweisung, welche entweder direkt projiziert in der Montage oder am PC-Terminal angezeigt wird. Die Arbeitsanweisung gibt dabei Informationen über die jeweiligen Montageschritte und Montagepositionen. Im zweiten Task wird die jeweilige Carbon-Matte aus der angrenzenden Materialbereitstellungsfläche geholt und auf die jeweilige mit 2D-Laser angezeigte Montageposition montiert. Dabei wird beim 2D-Laser ausschließlich die Kontur der Matte angezeigt und bei der Variante „Dynamische Projektion“ noch weitere Zusatzinformationen (Pfeile) projiziert, welche beim schnellen Auffinden der Montageposition unterstützen. Nach Montage der Matte wird die Qualität überprüft und der Task durch Interaktion mit einem projizierten Button oder durch Klick am PC-Terminal bestätigt. Der Vorgang wird solange durchgeführt, bis der Prozess beendet wird (Rupprecht et al., 2020).

Ergebnisse der Evaluierung

Beim Experiment 1 werden zur Evaluierung die Methoden „System Usability Scale“, „Technology Acceptance Model“ (vgl. Kapitel 5.1.1) sowie die Gesamtprozesszeit des Montageprozesses ermittelt. Dabei wurden die Unterschiede der beiden Systeme sowie der Mehrwert der dynamischen Projektion mit 16 Teilnehmer*innen (12 männlich, 4 weiblich) evaluiert. Die Teilnehmer*innen waren zwischen 20 und 39 Jahre alt und Studierende und interessierte Laien mit technischem Background. In Tabelle 22 wird ein Überblick über die Evaluierungsergebnisse aus Experiment 1 gegeben (Rupprecht et al., 2020).

Tabelle 22 | Evaluierungsergebnisse von Experiment 1, in Anlehnung an Rupprecht et al. (2020)

Evaluierungsmethoden	PC-Terminal-Experiment (Informationsbereitstellung/ Interaktion über PC-Terminal)	Dynamisches Projektionssystem-Experiment (Bereitstellung/Interaktion direkt am Werkzeug über dynamische Projektion)
Prozesszeit	02:25 min	02:19 min
Total SUS Score	71,1	85,3
Adjective rating - SUS	OK	Excellent
TAM Score	4 +	5 ++
Präferenzen der Teilnehmer*innen	#2/#16	#14/#16

Die Evaluierungsergebnisse zeigen, dass bei der Variante „Dynamische Projektion“ eine kleine Verbesserung der Gesamtprozesszeit von rund 4 % ermittelt werden konnte (Rupprecht et al., 2020). Hierbei ist anzumerken, dass in den folgenden Experimenten die „Task Completion Time“ zur Evaluierung der Leistung und Effizienz des Montageprozesses sowie der Qualität der eingesetzten Informationsbereitstellung verwendet wird, welche sich nach Funk, Kosch et al. (2015) besser für die Analyse von Montageprozessen eignet als die reine Ermittlung der Gesamtmontagezeit (vgl. Kapitel 5.1.2).

Außerdem konnte festgestellt werden, dass der SUS Score beim Projektionssystem mit 85,3 und „excellent“ und der TAM Score mit 5 ++ wesentlich besser bewertet wurde als die Vergleichsvariante. Dadurch, dass alle Teilnehmer*innen beide Varianten durchgeführt haben, konnten die Präferenzen für ein jeweiliges Informationssystem ermittelt werden. Dementsprechend präferierten 14 von 16 Teilnehmer*innen den Einsatz des dynamischen Projektionssystems. Durch die Kurzinterviews konnten weitere essenzielle Erkenntnisse (Tabelle 23) gewonnen werden, welche anschließend in den weiteren Gestaltungen integriert wurden. (Rupprecht et al., 2020).

Tabelle 23 | Kernaussagen von Experiment 1, in Anlehnung an Rupprecht et al. (2020)

Erkenntnisse zum Experiment mit dem PC-Terminal	Erkenntnisse zum Experiment mit dem dynamischen Projektionssystem
Lange Wege zum PC-Terminal zur Bestätigung des Prozessschrittes sind umständlich.	Information und Interaktion direkt am Werkstück ist komfortabel.
Wichtige Informationen wurden in der Menge des Textes (z. B. „Prozess abgeschlossen“ laut sagen) am Ende des Experiments übersehen.	Keine langen Wege, um an Information zu gelangen.
Keine Flexibilität durch feste Position des PC-Terminals.	Einfach und intuitiv zu bedienen und zu verstehen.
Orientierung der angezeigten Informationen zu einzelnen Positionen auf dem PC-Terminal ist schwer auf das reale Werkstück zu transferieren.	Zu Beginn wird eine höhere Menge an Informationen benötigt, als später im Prozess (Vereinfachung von gleichen Informationen durch z. B. Pfeile oder weniger und konkretere Informationen).
Erhebliche Menge an Informationen ist schwer zu merken und erfordert häufige Rückkehr zum PC-Terminal.	Es wäre sinnvoll, wenn sich die Informationen am Mensch ausrichten würden.
Bestätigen nach jedem Arbeitsschritt am PC-Terminal ist umständlich.	Projektionen auf einzelne Elemente (Wabenkerne) sind schwer zu lesen.

Eine(r) der Teilnehmer*innen merkte an, dass er/sie sich mit dem PC-Terminal wohler fühle, weil Erfahrung mit der Bedienung des PCs vorhanden war.	Leicht verständlich und in unmittelbarer Nähe der Bedienung/des Montageortes.
	Die Information „Prozess abgeschlossen“ wurde nicht übersehen, da sie direkt am Bauteil angezeigt wurde.
	Für geschulte Anwender*innen ist es jedoch notwendig, die angezeigten Informationen zu optimieren, z. B. weniger Informationen über Änderungen etc.

Dabei konnte festgestellt werden, dass bei der Variante mit PC-Terminal die langen Wege und die textuellen Arbeitsanweisungen als nicht komfortabel wahrgenommen wurden und die Information mit hohem Aufwand auf die realen Montageobjekte transferierbar sind. Des Weiteren konnte ermittelt werden, dass der Einsatz des Projektionssystem sowie die Anzeige der Informationen direkt in der Montage eine komfortable und intuitiv verständliche Möglichkeit bietet, Arbeitsanweisungen bereitzustellen. Hier ist anzumerken, dass die Informationen noch besser aufbereitet und sich automatisch dem Menschen ausrichten sollten sowie dass die Informationen im Laufe der Prozesses bzw. je nach Qualifizierungsgrad des Menschen individuell angepasst werden können (Rupprecht et al., 2020).

Zwischenergebnis und Resümee Experiment 1

Als Zwischenergebnis des ersten Experiments kann zusammengefasst werden, dass der Einsatz eines dynamischen Projektionssystems bereits eine vielversprechende Lösung zur direkten Anzeige von Arbeitsinformationen und Interaktion in der Montage zeigt und bereits eine gute Usability aufweist und die Technologie akzeptiert wird. Durch die qualitativen Interviews konnte ermittelt werden, dass das System intuitiv und komfortabel bedienbar ist und Wege zur Information im Vergleich zum PC-Terminal-Experiment eingespart werden können. In folgenden Evaluierungen ist es noch notwendig, Überlegungen zu treffen, wie die intuitive und projektionsgerechte Informationsgestaltung umgesetzt und wie eine dynamische Führung mittels Projektionssystem gestaltet werden kann. Ebenso ist es notwendig die projizierten Buttons durch eine funktionierende Mensch-Maschine-Interaktion zu ersetzen, welche ebenso direkt in der Montage und im großen Arbeitsraum verwendet werden kann. Die Erkenntnisse aus Experiment 1 waren die Grundlagen für weitere Entwicklungen des Spatial Augmented Reality Systems und gaben einen guten Einblick über die Potenziale des dynamischen Projektionssystems in der industriellen Großgeräte- und Baustellenmontage.

5.2.2 Experiment 2 – Fan-Cowl-Montage: Dynamische In-situ-Projektion mit Informationsführung zur Montageposition, Interaktion mittels Gestensteuerung

Im Experiment 2 wird als Informationsbereitstellungssystem ausschließlich das dynamische Projektionssystem verwendet und zwei Varianten der Informationsführung (In-view- und In-situ-Projektion) evaluiert (Abbildung 104). Darüber hinaus wird die Mensch-Maschine-

Interaktion zur Steuerung des Projektionssystems mithilfe einer visuellen Gestensteuerung mit Kamera und eines manuell gelabelten Datensatzes umgesetzt (Rupprecht et al., 2021).



Abbildung 104 | Experiment 2: In-view- und In-situ-Projektion, in Anlehnung an Rupprecht et al. (2021)

Bei der dynamischen In-situ-Projektion mit Führung zur Montageposition wird mittels Projektionssystem dynamisch zur jeweiligen Montageposition im Arbeitsraum geführt und gleichzeitig die exakte Position (grüne Kontur) und Orientierung (roter Punkt) (vgl. Kapitel 3.3.2) der zu montierenden Bauteile sowie Informationen zu den Montagetätigkeiten mittels projizierten Arbeitsinformationen angezeigt. Für die Steuerung des dynamischen Projektionssystems (Weiter- und Zurückschalten der Arbeitsinformationen) wird eine Gesteninteraktion mit zwei Gesten („offene Hand“ und „Faust“) eingesetzt, welche mit einem eigens für die Anforderungen der Baustellenmontage entwickelten Datensatz entwickelt und trainiert wurden (vgl. Kapitel 3.3.3). Mit der natürlichen Gesteninteraktion kann intuitiv die Steuerung des Projektionssystems über einen großen Arbeitsraum und direkt an der jeweiligen Montageposition durchgeführt werden. Die Informationen über die jeweiligen Möglichkeiten der Gesteninteraktion sowie visuelles Gestenfeedback nach Ausführung einer Geste werden mittels Projektionssystem angezeigt, sodass zu jedem Zeitpunkt im Prozess der Status der Informationsbereitstellung bekannt ist. Nach Bestätigung des jeweiligen Tasks mittels Geste „offene Hand“ wird mit dem Projektionssystem zur nächsten Montageposition geleitet, solange bis der Prozess abgeschlossen ist. Mittels der Geste „Faust“ kann zur vorherigen Arbeitsanweisung zurückgeschaltet werden (Rupprecht et al., 2021).

Als Baseline-Variante des Experiments dient die statische In-view-Projektion, bei der die Informationen ausschließlich statisch in der Mitte der Montagestation angezeigt werden und mittels Gesteninteraktion der Inhalt der Projektion wechselt. Bei dieser Variante wird rein mittels textlicher bzw. grafischer Information angezeigt, wo und welche Bauteile zu montieren sind. Dabei ist jede der 15 verfügbaren Montageposition mit einer Nummer (Position 001 bis 015) versehen, welche dementsprechend erst gefunden werden muss (Rupprecht et al., 2021). Die Ziele des Experiments sind (Rupprecht et al., 2021):

- Evaluierung der verbesserten Informationsgestaltung der Arbeitsinformationen
- Evaluierung der Usability und Prozesseffizienz bei der dynamischen In-situ-Projektion mit Führung zur Montageposition im Vergleich zur statischen In-view-Projektion
- Evaluierung der Gesteninteraktion mit eigens aufbereitetem Datensatz
- Verbesserung der Zeit $t_{\text{lokalisieren Montageposition}}$ durch eine dynamische In-situ-Projektion und Führung mittels Projektionssystem

Des Weiteren soll durch die Evaluierung folgende Frage beantwortet werden (Rupprecht et al., 2021): „Was sind die Vorteile der Informationsbereitstellung mittels dynamischer In-situ-Projektion und Informationsführung zur Montageposition im Vergleich zur statischen In-view-Projektion bei Einsatz eines Spatial Augmented Reality Systems in der industriellen Baustellenmontage?“

Prozessablauf Experiment 2

Beim Experiment 2 wurde für beide Systeme (In-view- und In-situ-Projektion) ein einheitlicher Montageprozess mit sechs Montagetasks zum Montieren von Carbon-Matten definiert (Abbildung 105). Die Gestaltung der Arbeitsinformationen wurden dementsprechend auf die beiden Systeme adaptiert (Rupprecht et al., 2021).

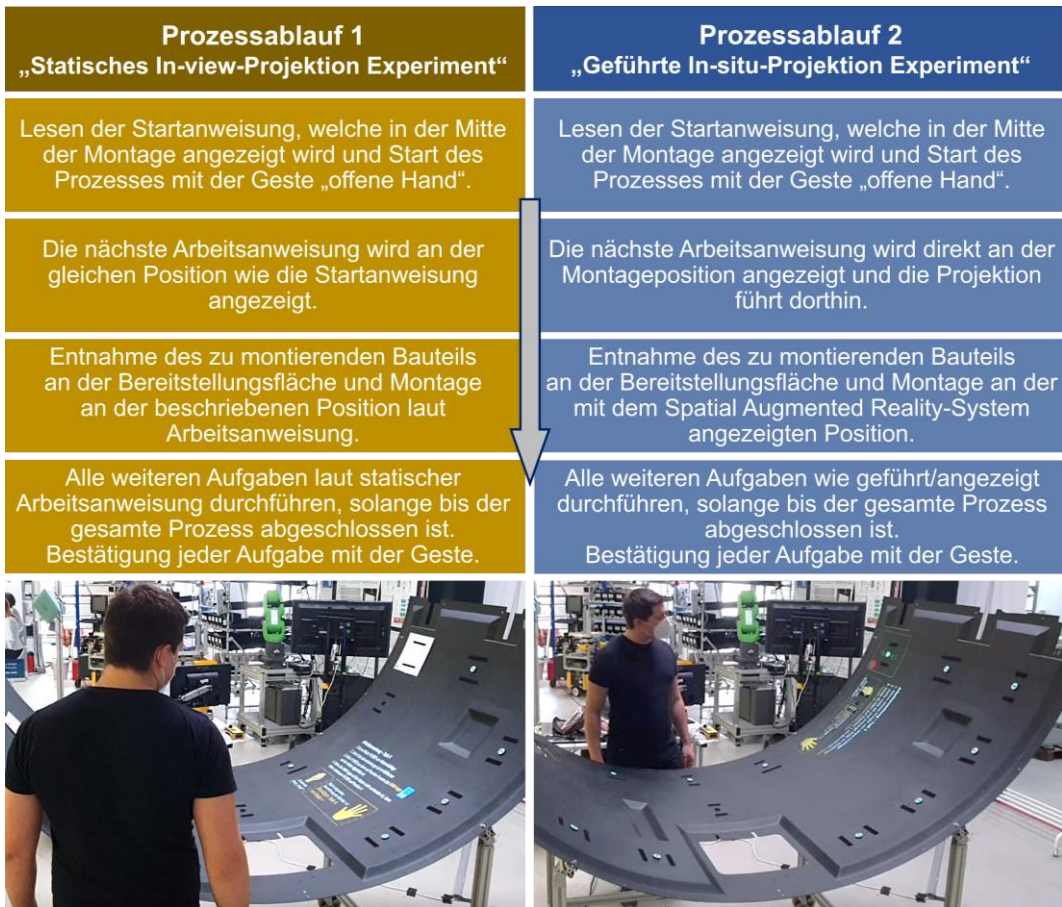


Abbildung 105 | Prozessablauf von Experiment 2, in Anlehnung an Rupprecht et al. (2021)

Der erste Task des Prozessablaufes besteht aus dem Lesen der Start-Arbeitsanweisung, welche bei beiden Varianten in der Mitte der Montagestation projiziert wurde. Durch Bestätigung mittels Geste „offene Hand“ wird zur nächsten Arbeitsinformation geschaltet. Dabei wird bei der dynamischen In-situ-Variante direkt zum jeweiligen zweiten Task hingeführt, wobei bei der statischen In-view-Projektion die Arbeitsanweisung an derselben Stelle wie die Start-Anweisung projiziert wird. Die jeweilige Arbeitsanweisung gibt bei beiden Varianten textliche Informationen über die Entnahme der jeweiligen Carbon-Matte an, welche mit Nummern versehen in der Bereitstellungsfläche neben der Montagestation zur Verfügung stehen. Außerdem gibt die Arbeitsanweisung Informationen über die jeweilige Montage-

position: bei der In-situ-Projektion mit grafischer Kontur und Anzeige der Orientierung an der exakten Position, bei der In-view-Projektion rein als textliche Information. Die Matte wird dabei von der Bereitstellungsfläche entnommen und entweder an der jeweiligen projizierten Kontur oder den mithilfe von Nummern markierten Positionen montiert. Nach jedem Arbeitsschritt wird mit der Geste „offene Hand“ der Montagefortschritt bestätigt, die Feedbackinformation zur Gestensteuerung projiziert und die nächste Anweisung angezeigt. Der Vorgang wird solange durchgeführt, bis der Prozess zu Ende ist (Rupprecht et al., 2021).

Ergebnisse der Evaluierung

Beim Experiment 2 werden zur Evaluierung der Usability und kognitiven Arbeitsbelastung die Methoden „System Usability Scale“ und „NASA-RTLX“ verwendet. Für die Ermittlung der Prozesseffizienz wurde die „Task Completion Time“ nach GATM (vgl. Kapitel 5.1.2) verwendet, um nähere Aussagen über den Mehrwert der Informationsführung treffen zu können. Ebenso wird zusätzlich die durchschnittliche Taskzeit als Summe der Zeiten der GATM-Phasen sowie die Zeit für die Gesteninteraktion ermittelt. Das Experiment wurde in der TU Wien-Pilotfabrik mit 12 Teilnehmer*innen (8 männlich, 4 weiblich) im Alter zwischen 19 und 29 Jahren durchgeführt. Die Teilnehmer*innen waren Studierende und interessierte Laien mit technischem Background und die meisten davon waren mit dem Demonstrator „Fan-Cowl-Montage“ bereits vertraut. Die Ergebnisse der Evaluierung der kognitiven Arbeitsbelastung mit der NASA-RTLX-Methode werden in Abbildung 106 dargestellt (Rupprecht et al., 2021).

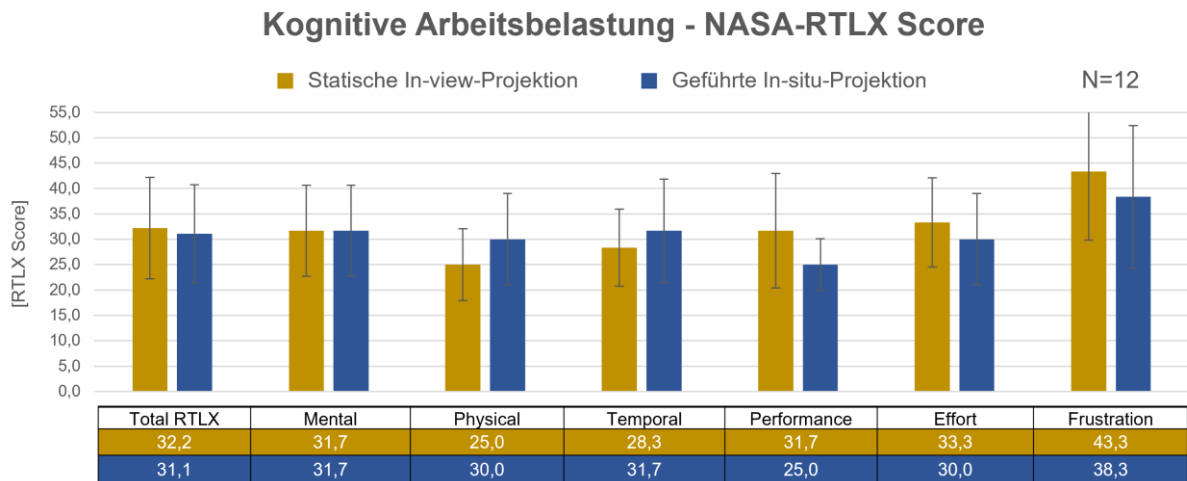


Abbildung 106 | NASA-RTLX – Experiment 2, in Anlehnung an Rupprecht et al. (2021)

Dabei konnte festgestellt werden, dass die Gesamtbewertung der kognitiven Arbeitsbelastung (Total RTLX Score) bei der dynamischen In-situ-Projektion mit 31,1 Punkten im Vergleich zur statischen In-view-Projektion mit 32,2 Punkten besser abschneidet. Hervorzuheben ist, dass vor allem bei der detaillierten Einzelauswertung der NASA-RTLX Scores die In-situ-Projektion bei der Leistungsfähigkeit (Performance) mit 25,0 Punkten (in view 31,7 Punkte), beim Frustrationslevel mit 38,3 (in view: 43,3 Punkte) sowie bei Bemühung/Aufwand „Effort“ mit 30,0 (in view: 33,3 Punkten) vorne liegt. Dies kann darauf zurückgeführt werden, dass bei der In-situ-Projektion dynamisch zur Montageposition geführt wird und nicht erst die Positionen gesucht werden müssen. Ebenso wird hier mittels des Assistenzsystems bei der Platzierung

und Orientierung der Carbon-Matten unterstützt, was sich auf die Leistung und auch Vermeidung von Frustration positiv auswirkt. Die mentale Belastung (Mental) wurden mit 31,7 Punkten bei beiden Varianten gleich bewertet, die physische (Physical) und temporäre Belastung (Temporal) liegt bei der In-view-Variante vorne (Rupprecht et al., 2021).

Zusammengefasst kann ermittelt werden, dass durch die gezielte Anzeige der Montageposition und Orientierung sowie der dynamischen Führung zu den jeweiligen Positionen eine Reduzierung der kognitiven Arbeitsbelastung im Gesamten sowie im Speziellen eine Verbesserung des Frustrationslevel, der Leistungsfähigkeit und des Aufwands erreicht wurde.

Für die praktische Anwendung des Assistenzsystems ist die Evaluierung Benutzerfreundlichkeit (Usability) ein entscheidender Faktor, dementsprechend wurde der „System Usability Scale“ herangezogen und die Ergebnisse mittels Total SUS Score angegeben (Abbildung 107) (Rupprecht et al., 2021).

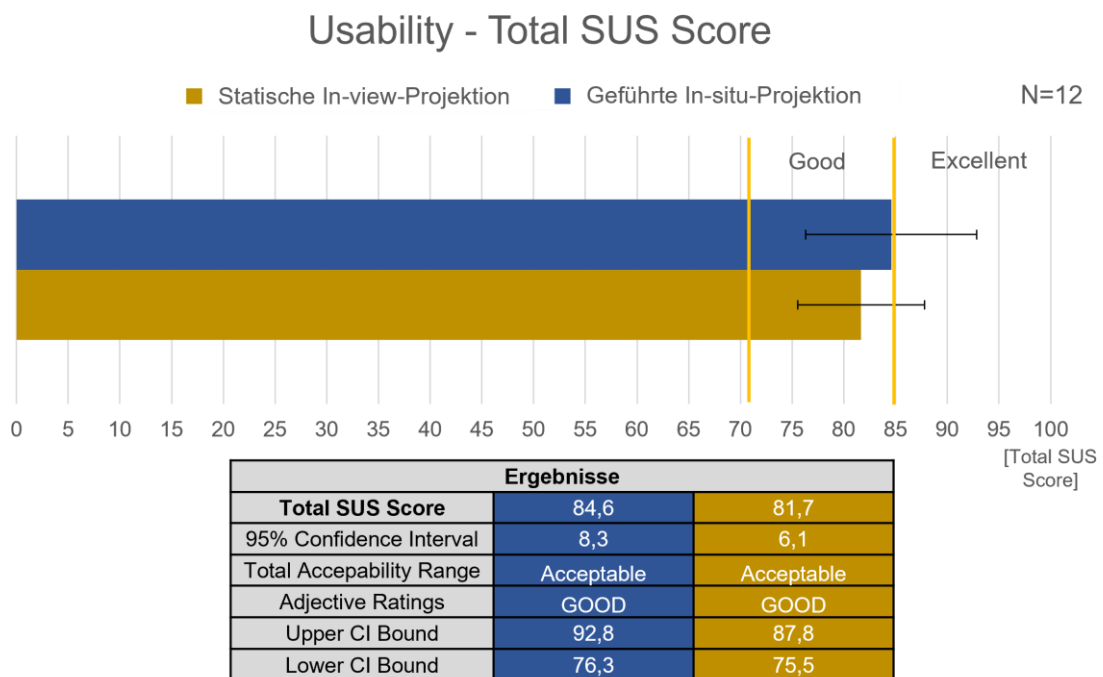


Abbildung 107 | SUS Score – Experiment 2, in Anlehnung an Rupprecht et al. (2021)

Dabei wurde festgestellt, dass die In-situ-Projektion mit einem Total SUS Score von 84,6 Punkten im Vergleich zur statischen In-view-Projektion mit 81,7 Punkten eine bessere Usability aufweist, beide Varianten jedoch im Bereich der positiven Akzeptanz liegen und mit dem „Adjective Rating GOOD“ bewertet werden. Zusammengefasst haben beide Varianten eine gute Benutzerfreundlichkeit, wobei die In-situ-Projektion in den Einzelattributen etwas höhere Werte hält (Rupprecht et al., 2021).

Zur Bewertung der Prozesseffizienz und der einzelnen Montagezeiten wurde die „Task Completion Time“ nach dem GATM (vgl. Kapitel 5.1.2) verwendet. Für die exakte Analyse der Zeiten ist es notwendig, die einzelnen Phasen in Bezug auf das Experiment und

den Montageprozess genau zu definieren und abzugrenzen (Tabelle 24) (Rupprecht et al., 2021).

Tabelle 24 | Beschreibung der GATM-Phasen – Experiment 2, in Anlehnung an Rupprecht et al. (2021)

GATM-Phasen	Definition der Phasen (Beginn/Ende)
t_lokalisieren Bauteil	Beginn des Lesens der Arbeitsanweisung bis zum Austrecken des Armes, um das Montageteil aus der Bereitstellungsfläche aufnehmen zu können. Diese Phase beinhaltet auch den Weg zur Bereitstellungsfläche.
t_aufnehmen	Zeit zum Aufnehmen des Bauteils von der Bereitstellungsfläche. Beginn des Aufnehmens nach ausgestrecktem Arm bis zum Heranführen des Bauteils an die Körpermitte.
t_lokalisieren Montageposition	Suchen der Montageposition an der Montage, Hinführen des Bauteils zur exakten Montageposition (dabei schwebt das Teil in geringem Abstand über der Position). Der Weg von der Bereitstellungsfläche zur jeweiligen Montageposition ist hier integriert.
t_montieren	Montieren des Bauteils an der Montageposition gemäß Arbeitsanweisung. Dabei sind alle weiteren Aufgaben der Anweisung ebenfalls in dieser Phase enthalten, wie z. B. das Entfernen der Folie von den Matten, das Entsorgen der Folie im Müll oder die Durchführung der Qualitätsprüfung. Sie umfasst auch die Zeit für die Ausführung der Gesten und die Projektion des visuellen Gestenfeedbacks.

Dabei wurde der zugrunde liegende Montageprozess inklusive der einzelnen Montagetasks in die vier GATM-Phasen zerlegt und dementsprechend beschrieben. Die GATM-Zeitphasen *t_lokalisieren Bauteil* und *t_lokalisieren Montageposition* sind dabei aufgabenunabhängige Zeiten und ein Maß für die Qualität der Arbeitsanweisungen sowie für die Evaluierung des Assistenzsystems zur Unterstützung beim Auffinden und Positionieren von Bauteilen. Beim In-situ-Experiment wurde speziell durch den Einsatz des Spatial Augmented Reality Assistenzsystems mit dynamischer In-situ-Projektion die Führung zur Montageposition verbessert, was sich dementsprechend positiv auf die Zeit *t_lokalisieren Montageposition* auswirken soll. Die Zeiten *t_aufnehmen* und *t_montieren* sind aufgabenabhängige Zeiten und dementsprechend stark von den Aufgaben und Umfängen der jeweiligen Montagetasks abhängig. Um die beiden Experimente vergleichen zu können, wurden die jeweiligen Montagetasks mit gleichem Komplexitätsgrad gestaltet. Zusätzlich zu den detaillierten Zeiten der GATM-Phasen wird die durchschnittliche Taskzeit zur Interpretation der jeweiligen Montagetasks und die Zeit der Gesteninteraktion (*t_Geste*) zur Evaluierung der Gestensteuerung aufgenommen (Rupprecht et al., 2021).

Um die Taskzeiten ermitteln zu können, wurden die Teilnehmer*innen während des Experiments durch eine zusätzliche Kamera gefilmt. Die Aufnahmen wurden anschließend mit dem Open-Source-Video-Annotationstool „ANVIL2“ (www.anvil-software.org) ausgewertet und die Zeitwerte als CSV-Format nach MS Excel exportiert. Anschließend wurden aus den Einzelauswertungen statistische Parameter wie Mittelwert, oberes und unteres Quartil (Q1, Q3) und der Interquartilsabstand (IQR) berechnet und die Ausreißer, nach der Formel $Q3 + IQR \cdot 1,5$ und $Q1 - IQR \cdot 1,5$, aus dem Datensatz eliminiert. Für die Darstellung des statistischen Fehlers wird das 95-%-Konfidenzintervall verwendet. Die Daten wurden zur weiteren Interpretation und des Vergleichs der beiden Experimente grafisch ausgewertet (Abbildung 108) (Rupprecht et al., 2021).

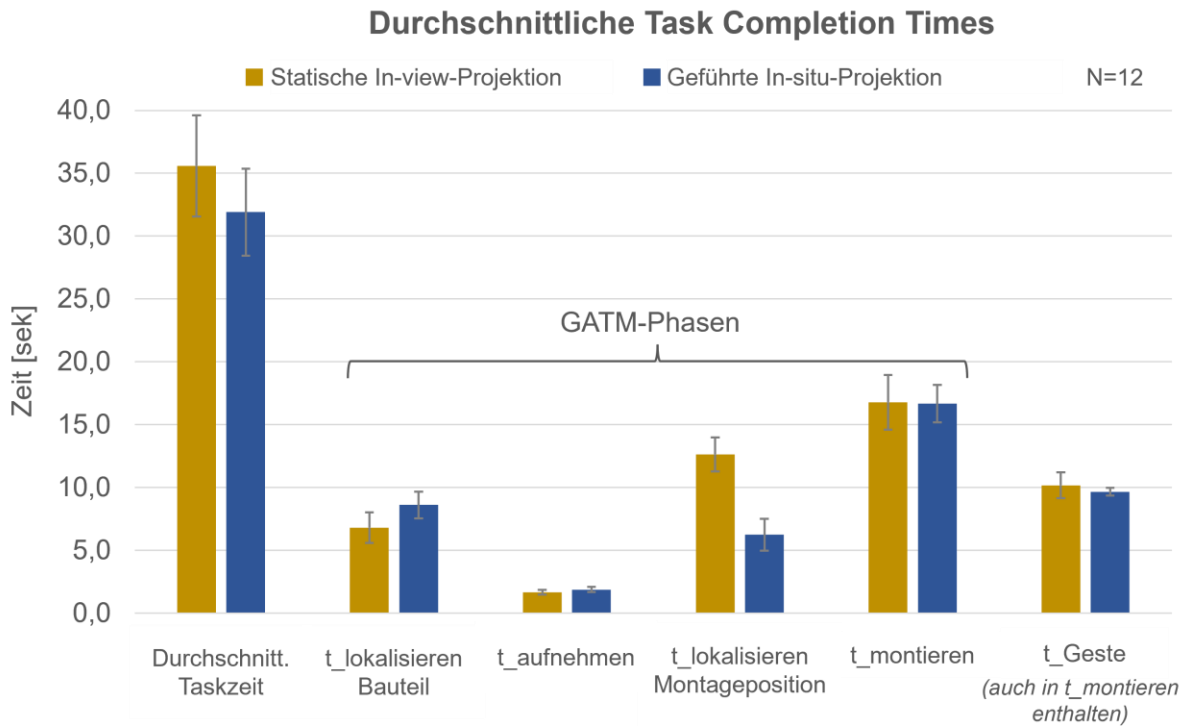


Abbildung 108 | GATM – Experiment 2, in Anlehnung an Rupprecht et al. (2021)

Im Allgemeinen kann aus der Analyse der Taskzeiten festgestellt werden, dass die dynamische In-situ-Projektion bei einer durchschnittlichen Taskzeit von 31,9 Sekunden im Vergleich zur In-view-Projektion mit 35,6 Sekunden um rund 10 % besser abschneidet. Die Analyse der durchschnittlichen Taskzeit ist jedoch nur bedingt für die Analyse der Evaluierung des eingesetzten Assistenzsystems aussagekräftig. Durch Analyse der Zeiten der GATM-Phasen kann festgestellt werden, dass die Zeiten *t_aufnehmen* und *t_montieren* nahezu ident sind, was auf die gleiche Komplexität der Arbeitsaufgaben zurückzuführen ist. Die Zeit *t_gesture* ist sowohl in *t_montieren* enthalten und wurde zusätzlich getrennt ausgewertet. Dabei konnte eine durchschnittliche Zeit der Gesteninteraktion *t_gesture* von rund 10 Sekunden erreicht werden. Unterschiede zeigen sich vor allem in den aufgabenunabhängigen Zeiten *t_lokalisieren Bauteil* und *t_lokalisieren Montageposition*. Die Zeit *t_lokalisieren Bauteil* wurde bei der In-view-Projektion mit durchschnittlich 2 Sekunden schneller analysiert. Der Grund liegt darin, dass die Wegzeiten zur Entnahmeposition bei der In-situ-Projektion etwas länger waren, weil zuerst zur geführten Arbeitsanweisung hingegangen werden musste, um Informationen über das zu entnehmende Bauteil zu erhalten, bevor es aus der Bereitstellungsfläche entnommen werden konnte. Bei der statischen In-view-Projektion entfällt diese Wegstrecke, weil die Projektion immer an der gleichen Stelle in der Montage angezeigt wird. **Das Kernergebnis der Analyse ist, dass die Zeit *t_lokalisieren Montageposition* bei der In-situ-Projektion mit Führung zur Montageposition um den Faktor 2 niedriger ist, als bei der In-view-Projektion** (Rupprecht et al., 2021).

Das Ergebnis resultiert vor allem daraus, dass mithilfe des dynamischen Projektionssystems zur Montageposition hingeführt und gleichzeitig die exakte Position und Orientierung des montierten Teils angezeigt wird (Rupprecht et al., 2021).

Zur Analyse der Fehler wurden die Ausreißer der jeweiligen Zeitwerte analysiert und anschließend die Fehlergründe aus dem aufgenommenen Video ermittelt. Insbesondere konnte festgestellt werden, dass von 144 der durchgeführten Gesteninteraktionen 14 (10 %) länger als der Durchschnitt dauerten. Diese wurden dementsprechend als Fehler interpretiert. Der Grund dafür war, dass die Hand nicht sofort erkannt wurde und die Geste mehrmals ausgeführt werden musste, woraus höhere Zeiten resultieren. Weiterhin konnten etwa 10 Werte in der Zeit $t_{\text{montieren}}$ als Fehler identifiziert werden, was auf eine falsche Handhabung oder Wiederholung der Montageaufgaben zurückzuführen ist (Rupprecht et al., 2021).

Qualitative Ergebnisse

Mittels Kurzinterviews nach jedem Experiment und der Auswertung der offenen Fragen des Fragebogens konnten folgende qualitative Aussagen ermittelt werden (Tabelle 25) (Rupprecht et al., 2021).

Tabelle 25 | Qualitative Ergebnisse – Experiment 2, in Anlehnung an Rupprecht et al. (2021)

Erkenntnisse In-view-Experiment
▪ Langes Suchen nach Montagepositionen war notwendig, was lästig war. ▪ Es war gut, dass die Arbeitsinformationen immer im Blickfeld des Werkers waren.
Erkenntnisse In-situ-Experiment
▪ Die angezeigte Montageposition war großartig und unterstützte bei der Durchführung des Vorgangs. ▪ Irritierend war, dass sich die Projektion zur nächsten Position auf die andere Seite der Montage bewegte, so dass der Projektion gefolgt werden musste. ▪ Die angezeigte Information für die Entnahme des Bauteils sollte sich an der gleichen Stelle wie die vorherige befinden und nur die Informationen zur projizierten Montageposition sollte zur nächsten Position führen.
Allgemeines Feedback
▪ Die Gestensteuerung war großartig, könnte aber in Bezug auf Schnelligkeit noch verbessert werden. ▪ Die Feedbackfolien zur Bestätigung der Aufgaben könnten schneller erscheinen, um schneller mit den nächsten Aufgaben fortzufahren. ▪ Die Piktogramme waren verständlich und sollten zukünftig noch öfters verwendet werden. ▪ Nach ein paar Aufgaben können die Informationen vereinfacht werden, z. B. nur noch Piktogramme. ▪ Am besten wäre eine Mischung aus beiden Versuchen, d.h. die Position und Orientierung mit SAR und die Arbeitsanweisungen immer im Blickfeld zu zeigen, evtl. relativ zum Menschen adaptiv ausgerichtet.

Aus dem qualitativen Feedback der Teilnehmer*innen kann abgeleitet werden, dass das lange Suchen der Montagepositionen bei der In-view-Projektion negativ wahrgenommen wurde. Es wurde jedoch positiv vermerkt, dass die In-view-Projektion immer im Blickfeld zu sehen war (Rupprecht et al., 2021). Bei der In-situ-Projektion wurde positiv angemerkt, dass die Führung zur nächsten Montageposition und die direkte Anzeige von Montagepositionen und Orientierung eine gute Unterstützung bei der Durchführung der Arbeitstätigkeiten war. Eine Anmerkung zur Verbesserung war, dass die Information über die Entnahme der Bauteile am gleichen Ort wie die vorherige Anweisung angezeigt werden sollte und erst nach der Entnahme an der Bereitstellungsfläche zur Montageposition geführt werden soll. Dies würde Weg und Zeit (vgl. $t_{\text{lokalisieren Bauteil}}$) einsparen und die Effizienz erhöhen (Rupprecht et al., 2021). Aus dem allgemeinen Feedback, welches für beide Varianten gilt, konnte festgestellt werden, dass die Gesteninteraktion bereits annehmbar funktionierte, jedoch noch etwas schneller reagieren könnte. Außerdem wurde die sinnvolle Nutzung der Feedback-Informationsanzeige

nach jeder Gesteninteraktion hinterfragt und vorgeschlagen, dass schneller zur nächsten Arbeitsinformation weitergeschaltet werden sollte. Die Verwendung von intuitiven Piktogrammen wurden als besonders angenehm und effektiv empfunden und könnte sogar noch häufiger eingesetzt werden. Zu guter Letzt wurde angemerkt, dass bei gleichen Arbeitsschritten im Laufe des Arbeitsprozesses der Detailgrad der Arbeitsinformationen reduziert oder durch vereinfachte Anzeige dargestellt werden könnte (Rupprecht et al., 2021).

Zwischenergebnis und Resümee Experiment 2

Aus der Evaluierung des zweiten Experiments kann ermittelt werden, dass sich der Einsatz eines Spatial Augmented Reality Assistenzsystems mit dynamischer In-situ-Projektion besonders für die direkte Informationsanzeige in der Baustellenmontage eignet und die Interaktion mittels Gesten eine gute Möglichkeit zur Steuerung des Projektionssystems bietet. Im Speziellen konnte festgestellt werden, dass die dynamische Führung zur Montageprojektion eine Verbesserung der Prozesszeit (*t_{lokalisieren Montageposition}*) sowie Reduzierung der kognitiven Arbeitsbelastung bringt, was sich ebenso in den Aussagen des qualitativen Feedbacks wiederfindet. Verbesserungspotenzial besteht noch darin, dass die jeweiligen Entnahme-Informationen an der gleichen Position wie die vorherige Arbeitsanweisung angezeigt werden sollte bzw. dass das Projektionssystem gleich direkt zu den jeweiligen Entnahmepositionen hinführt und anzeigt, welches Bauteil zu entnehmen ist. Dabei besteht noch Potenzial in der Vereinfachung der Gestaltung der Arbeitsanweisungen im Zeitverlauf und dem vermehrten Einsatz von Piktogrammen. Bei der Gesteninteraktion besteht noch Bedarf, die Aufbereitung des Datensatzes zu vereinfachen und die Gestensteuerung auch für Montagebauteile ohne dunklen Hintergrund zu ermöglichen sowie die Erkennungsstabilität und -geschwindigkeit zu verbessern. Zusammengefasst kann gesagt werden, dass durch die Durchführung des Experiments essenzielle Erkenntnisse für die Entwicklung des Spatial Augmented Reality Assistenzsystems gefunden wurden, dass die In-situ-Projektion einen Mehrwert für die Anzeige von Arbeitsinformationen bietet und sich die Gesteninteraktion für die Steuerung des Assistenzsystems in der Baustellenmontage eignet. Im folgenden Experiment werden die gewonnenen Erkenntnisse miteinbezogen und zusätzlich zur dynamischen In-situ-Projektion mit Informationsführung zur Montageposition, auch die Führung zur jeweiligen Entnahmeposition integriert.

5.2.3 Experiment 3 – Spielzeugbaustein-Montage: Dynamische In-situ-Projektion mit Informationsführung zur Entnahme- und Montageposition, Interaktion mittels Gestensteuerung

Das Experiment 3 dient zur Verallgemeinerung der Ergebnisse und Erkenntnisse aus den ersten beiden Experimenten und der Anwendung des Spatial Augmented Reality Assistenzsystems im weiteren Use Case „Spielzeugbaustein-Montage“ (Abbildung 109).

Baseline:

Informationsbereitstellung über statisches Tablet

Dynamische In-situ-Projektion mit Informationsführung zur Entnahme- und Montageposition | Interaktion mittels Gestensteuerung

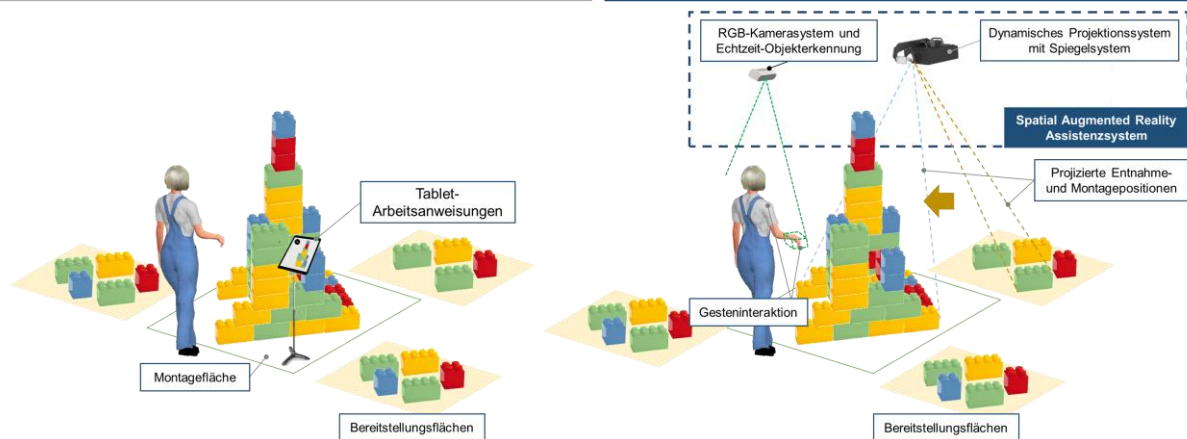


Abbildung 109 | Experiment 3: Spielzeugbaustein-Montage, dynamische In-situ-Projektion mit Informationsführung zur Entnahme- und Montageposition (Eigendarstellung)

Für das Experiment wurde, im Vergleich zu den ersten beiden Experimenten, ein komplexerer Montageprozess mit 18 Montagetaasks, drei Bereitstellungsflächen und 64 Bauteilen in zwei verschiedenen Größen und verschiedenen Farben gestaltet, bei dem unter Einsatz des Spatial Augmented Reality Assistenzsystems bzw. den Anweisungen des Tablets ein Spielzeugbausteinurm gebaut wurde.

Als Weiterentwicklung des vorherigen Experiments wird zusätzlich zur dynamischen In-situ-Projektion mit Informationsführung zur Montageposition mit dem Spatial Augmented Reality Assistenzsystem angezeigt, welches Bauteil an welcher Bereitstellungsfläche zu entnehmen ist. Die Informationsführung ist dabei so gestaltet, dass sie den Menschen dynamisch zur jeweiligen Bereitstellungsfläche und anschließend zur jeweiligen Montageposition führt. Zur besseren und schnelleren Erkennung im Raum wurde dabei die Information der Arbeitsanweisung als dynamische Animation umgesetzt (vgl. Kapitel 3.3.2). Zur Bestätigung eines Arbeitsschrittes sowie zur Weiterschaltung zur nächsten geführten Arbeitsinformation dient wiederum die Gesteninteraktion mit der Geste „offene Hand“ fürs Weiterschalten und der Geste „Faust“ fürs Zurückschalten. Die Datengrundlage der Gestensteuerung bzw. Bilderkennung wurde hierbei mit einem synthetisch generierten Datensatz umgesetzt. Der Datensatz hat die Besonderheit, dass die aufgenommenen Gesten durch den Algorithmus an 15 000 verschiedenen Positionen automatisiert platziert werden (vgl. Kapitel 3.3.3). Außerdem ist es möglich, die Gesten auf verschiedenen Hintergründen – auch auf hellem Hintergrund (z. B. dem Hallenboden) – auszuführen.

Als Baseline des Experiments dient die Anzeige der Arbeitsinformationen über ein statisches Tablet, welches auf einem Ständer in 1,5 Meter Entfernung zur Montagefläche bereitgestellt wurde. Die Bestätigung des Arbeitsschrittes sowie das Weiterschalten zur nächsten Information wurde mit interaktiven Buttons umgesetzt.

Folgend werden die eingesetzten Evaluierungsmethoden sowie die einzelnen Experimente beschrieben und anschließend die Ergebnisse präsentiert.

Die Ziele des Experiments sind folgende:

- Verallgemeinerung der Ergebnisse aus Experiment 1 und 2 und Ermittlung der Skalierbarkeit des Assistenzsystems auf weitere Use Cases
- Evaluierung der Usability und Prozesseffizienz bei der dynamischen In-situ-Projektion mit Informationsführung zur Entnahme- und Montageposition im Vergleich zur Tablet-Anweisung
- Evaluierung der Gesteninteraktion mit synthetisch aufbereitetem Datensatz auf hellem Hintergrund
- Verbesserung der Zeit $t_{\text{lokalisieren Bauteil}}$ und $t_{\text{lokalisieren Montageposition}}$ durch dynamische In-situ-Projektion und Informationsführung mittels Projektionssystem

Prozessablauf Experiment 3

Beim Experiment 3 wurde für die Informationsbereitstellung mit dem Spatial Augmented Reality Assistenzsystem sowie für die Anzeige der Informationen mit statischem Tablet ein einheitlicher Montageprozess mit 18 Montagetasks zum Bauen eines Spielzeugbausteinbaus definiert (Abbildung 110). Die Gestaltung der Arbeitsinformationen wurden dementsprechend auf das dynamische Projektionssystem und das Tablet angepasst.

Prozessablauf – Baseline Informationsbereitstellung über statisches Tablet	Prozessablauf Dynamische In-situ-Projektion mit Informations- führung zur Entnahme- und Montageposition
Lesen der Startanweisung, welche am Tablet angezeigt wird und Start des Prozesses durch Klick auf den Start-Button am Tablet.	Lesen der Startanweisung, welche in der Mitte der Montage am Hallenboden projiziert wird und Start des Prozesses mit der Geste „offene Hand“.
Textliche und grafische Beschreibung des zu entnehmenden Bauteils, der jeweiligen Bereitstellungsfläche und der Montagepositionen am Tablet.	Die Projektion führt dynamisch zur jeweiligen Bereitstellungsfläche und zeigt die zu entnehmenden Bauteile für 5 Sekunden an. Entnahme der Bauteile laut Projektion.
Ermittlung der Entnahmeposition und Entnahme des zu montierenden Bauteils an der Bereitstellungsfläche. Ermittlung der Montagepositionen laut Arbeitsanweisung.	Die Projektion führt nach Zeitablauf dynamisch zur jeweiligen Montageposition und zeigt mittels farbigen oder weißen Lichtflächen die exakten Montagepositionen und Orientierungen an.
Montage der Bauteile laut Tablet-Anweisung und Bestätigung des jeweiligen Montagetasks mit interaktivem Button am Tablet.	Montage der Bauteile laut projizierten Informationen und Bestätigung des jeweiligen Montagetasks mit der Geste „offene Hand“.
Alle weiteren Montagetasks laut statischer Tablet-Anweisung durchführen, solange bis der gesamte Prozess abgeschlossen ist. Bestätigung jedes Tasks mit jeweiligem Button.	Alle weiteren Montagetasks wie geführt/angezeigt durchführen, solange bis der gesamte Prozess abgeschlossen ist. Bestätigung jedes Tasks mit der Geste.

Abbildung 110 | Prozessablauf von Experiment 3 (Eigendarstellung)

Bei der dynamischen In-situ-Projektion wird die Startanweisung direkt in der Mitte der Montagestation auf den Hallenboden projiziert. Mittels Ausführung der Geste „offene Hand“ wird durch Bilderkennung der Prozess gestartet und die dynamische Projektion ausgelöst. Mittels animierter Arbeitsanweisung wird der Mensch zur richtigen Bereitstellungsfläche hingeführt, die zu entnehmenden Bauteile mittels Projektion angezeigt und nach Zeitablauf von 5 Sekunden (ab Ausführung der Geste) wird mittels Projektion zur Montageposition geführt und die richtigen Montagepositionen mittels Licht- und Farbflächen angezeigt. Der Mensch

folgt der Projektion zur Bereitstellungsfläche, entnimmt die jeweiligen Bauteile und verbaut diese auf den angezeigten Positionen. Nach Montage der Bauteile wird mittels Geste „offene Hand“ der Arbeitsschritt bestätigt und gleichzeitig die Führung und Anzeige der nächsten Schritte gestartet, bis der Prozess abgeschlossen ist (Abbildung 111).



Abbildung 111 | In-situ-Projektion und Informationsführung im Experiment 3 (Eigendarstellung)

Bei der Informationsbereitstellung mit statischem Tablet werden die Arbeitsanweisungen ausschließlich am Tablet angezeigt und die Bestätigung bzw. das Zurückschalten mit interaktiven Buttons umgesetzt. Die Arbeitsanweisung wird als textliche Information und als grafische Darstellung der Entnahme- und Montagepositionen umgesetzt (Abbildung 112).

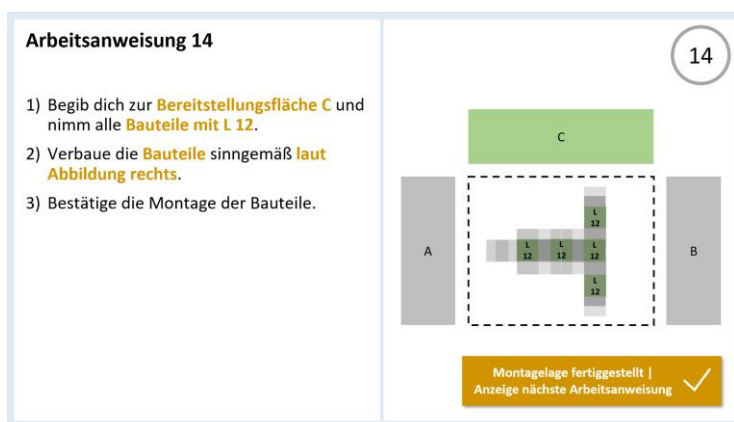


Abbildung 112 | Arbeitsanweisung via Tablet (Eigendarstellung)

Zur korrekten Identifizierung der jeweiligen Bauteile sind alle mit einer eindeutigen ID-Nummer versehen und die Bereitstellungsflächen mit jeweils den Buchstaben A, B, C markiert. Mittels Arbeitsanweisung wird angezeigt, welche Bauteile mit welcher ID-Nummer von welcher Bereitstellungsfläche zu entnehmen sind. Des Weiteren wird mittels grafischer Anzeige die exakte Montageposition angezeigt. Der Mensch begibt sich demnach zur Entnahmeposition, entnimmt die Bauteile und verbaut diese auf die jeweiligen Positionen, wobei er den Montageschritt mit Klick auf den Button am Tablet bestätigt. Der Vorgang wird so lange durchgeführt, bis der komplette Prozess zu Ende ist.

Ergebnisse der Evaluierung

Beim Experiment 3 wird analog zum vorherigen Experiment die Evaluierung der Usability mit dem „System Usability Scale“ und die Evaluierung der kognitiven Arbeitsbelastung mit der NASA-RTLX-Methode umgesetzt. Für die Ermittlung der Prozesseffizienz sowie um detaillierte Aussagen über den Effekt des Assistenzsystem bei der In-situ-Projektion mit Informationsführung zur Entnahme- und Montageposition treffen zu können, wird die „Task Completion Time“ nach dem GATM verwendet. Ebenso wird die durchschnittliche Taskzeit als Summe der Zeiten der GATM-Phasen sowie die Zeit für die Gesteninteraktion ermittelt. Beim Experiment 3 mit der Variante der In-situ-Projektion wird die Gestenzeit exklusive ermittelt und ist nicht in $t_{montieren}$ integriert. Dies hat den Grund, dass die Interaktion der Varianten einmal mittels Geste und einmal durch Klick auf einen Button umgesetzt wurde, was demnach nicht vergleichbar ist. Dadurch kann die reine Montagezeit, also das Auffinden der Bauteile sowie das eigentliche Montieren der beiden Varianten, verglichen und Aussagen über die Unterschiede ermittelt werden. Das Experiment am Use Case „Spielzeugbaustein-Montage“ wurde in der TU Wien-Pilotfabrik durchgeführt. Das Experiment wurde von 30 Teilnehmer*innen (63,33 % männlich, 23,33 % weiblich, 13,33 % keine Angabe) mit einem Alter zwischen 19 und 53 Jahren durchgeführt. Bei den Personen handelte es sich um Studierende und interessierten Laien mit technischem Background, einige davon waren jedoch auch Personen ohne technische Ausbildung. Die Ergebnisse der Evaluierung der kognitiven Arbeitsbelastung mit der NASA-RTLX-Methode werden in Abbildung 113 dargestellt.

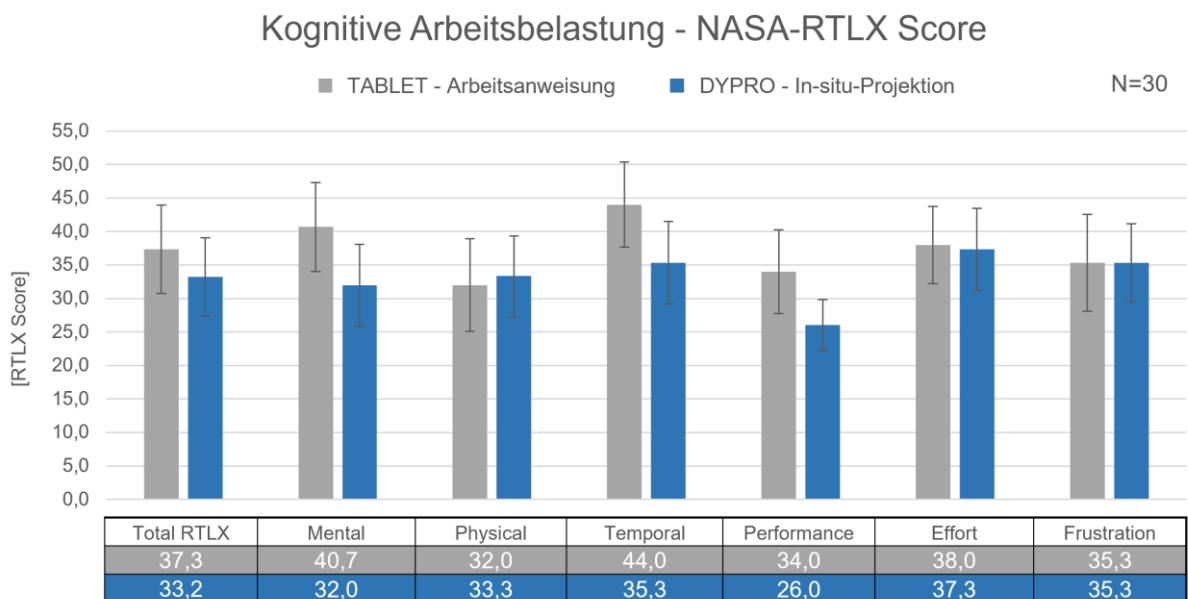


Abbildung 113 | Ergebnisse „Kognitive Arbeitsbelastung“ – Experiment 3 (Eigendarstellung)

Aus der Analyse der kognitiven Arbeitsbelastung konnte festgestellt werden, dass der Total RTLX Score bei der dynamischen In-situ-Projektion mit 33,2 Punkten im Vergleich zur Tablet-Anweisung mit 37,3 Punkten besser bewertet wurde. Des Weiteren konnte die In-situ-Projektion bei den Einzelauswertungen „Mentale Belastung“ (DYPRO: 32,0 P.; Tablet: 40,7 P.), „Temporäre Belastung“ (DYPRO: 35,3 P.; Tablet: 44,0 P.) sowie Leistungsfähigkeit (DYPRO: 26,0 P.; Tablet: 34,0 P.), Effort (DYPRO: 37,3 P.; Tablet: 38,0 P.) besser

abschneiden. Der Einzelwert „Frustrationslevel“ (DYPRO: 35,3 P.; Tablet: 35,3 P.) wurde mit gleicher Punktzahl bewertet. Bei der physischen Belastung liegt die Tablet-Anweisung mit 32,0 zu 33,3 Punkten vorne. Zusammengefasst kann aus den Ergebnissen des NASA-RTLX Score analysiert werden, dass durch die direkte In-situ-Projektion von Entnahme- und Montageposition sowie durch die dynamische Führung zu den jeweiligen Positionen eine Reduzierung der kognitiven Arbeitsbelastung im Gesamten erreicht wurde.

Wie bereits bei den vorherigen Experimenten wird auch bei Experiment 3 die Evaluierung der Usability mit dem „System Usability Scale“ durchgeführt. Die Ergebnisse der Auswertung werden mittels Total SUS Score angegeben (Abbildung 114).

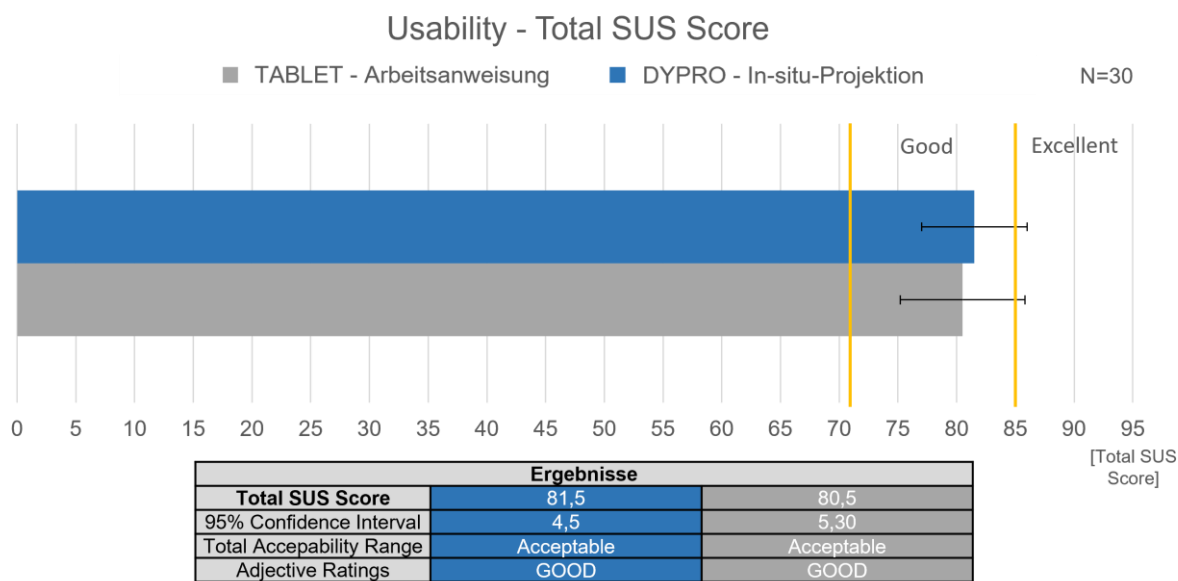


Abbildung 114 | Ergebnisse „Usability“ – Experiment 3 (Eigendarstellung)

Die Ergebnisse der Evaluierung zeigen, dass die In-situ-Projektion mit einem Total SUS Score von 81,5 Punkten im Vergleich zur statischen Tablet-Anweisung mit 80,5 Punkten eine bessere Usability ausweist. Beide Informationsbereitstellungsvarianten liegen im Bereich der positiven Akzeptanz („Acceptable“) und werden mit dem „Adjective Rating GOOD“ bewertet. Demnach ergibt sich für beide Varianten eine gute Benutzerfreundlichkeit, wobei die In-situ-Projektion etwas höhere Werte aufweist.

Die Evaluierung der „Task Completion Time“ nach dem GATM dient auch im Experiment 3 zur Bewertung der Prozesseffizienz sowie zur detaillierten Ermittlung der einzelnen Montagezeiten. Die genaue Definition der einzelnen GATM-Phasen in Bezug auf das Experiment 3 wird in Tabelle 26 dargestellt.

Tabelle 26 | GATM-Phasen und Gesteninteraktion – Experiment 3

GATM-Phasen	Definition der Phasen im Experiment (Beginn/Ende)
t_lokalisieren Bauteil	DYPRO – In-situ-Projektion: Zeit ab Start der In-situ-Projektion (ausgelöst durch Geste), Folgen der Projektion zur jeweiligen Bereitstellungsfläche bis zur Ermittlung der jeweiligen Entnahmeposition der Bauteile. Tablet-Arbeitsanweisung: Zeit ab der Bewegung zur jeweiligen Bereitstellungsfläche bis zur Ermittlung der jeweiligen Entnahmeposition der Bauteile.

t_aufnehmen	DYPRO – In-situ-Projektion/Tablet-Arbeitsanweisung: Zeit zum Aufnehmen des Bauteils von der Bereitstellungsfläche, Start mit Beginn des Bückens/Vorbeugens zum Aufnehmen bis zum Heranführen des Bauteils an die Körpermitte.
t_lokalisieren Montageposition	DYPRO – In-situ-Projektion: Zeit ab Start der Bewegung zum Folgen der In-situ-Projektion zur Montagestation und initiale Erfassung der projizierten Montagepositionen. Tablet-Arbeitsanweisung: Start mit Bewegung zur Montagestation und Ermittlung der jeweiligen Montagepositionen. Die Wegstrecke bzw. das Nachschauen der exakten Positionen am Tablet ist ebenso inkludiert.
t_montieren	DYPRO – In-situ-Projektion: Zeit zum Montieren der jeweiligen Montagebauteile gemäß In-situ-Projektion bis zum Heben der Hand für die Ausführung der Geste zur Bestätigung des Montageschrittes. Die Zeit für die Gesteninteraktion ist nicht integriert und wird getrennt aufgenommen. Tablet-Arbeitsanweisung: Zeit zum Montieren der jeweiligen Montagebauteile gemäß Tablet-Arbeitsanweisung bis zur Bestätigung des Montageschrittes mittels Button am Tablet.
t_Geste (Interaktionszeit)	DYPRO – In-situ-Projektion: Zeit ab Heben der Hand zur Ausführung der Geste bis zum Start der In-situ- Projektion. Info: Wenn die Geste neu angesetzt wird (z. B. Hand runter und wieder hoch), startet die Zeit neu. Tablet-Arbeitsanweisung: Interaktionszeit durch Klicken auf den Button am Tablet. Info: Zeit wird aufgrund zu kurzer Zeitspanne von < 1 Sekunde vernachlässigt und nicht extra ermittelt.

Um die beiden Varianten des Experiments vergleichbar zu machen, wurden die jeweiligen Montagetaasks sowie auch die Arbeitsanweisungen bezüglich Inhalt und Komplexitätsgrad gleich gestaltet. Dementsprechend ist zu erwarten, dass die aufgabenabhängigen Zeiten *t_aufnehmen* und *t_montieren* bei beiden Varianten annähernd gleich ausgewertet werden. Des Weiteren wird durch den Einsatz des Spatial Augmented Reality Assistenzsystems mit dynamischer In-situ-Projektion und Führung zur Entnahme- und Montageposition eine Verbesserung der Zeiten der GATM-Phasen *t_lokalisieren Bauteil* und *t_lokalisieren Montageposition* erwartet. Zusätzlich zu den Zeiten der GATM-Phasen, wird die durchschnittliche Taskzeit (Summe der GATM-Phasen) ausgewertet, um einen Gesamtüberblick der einzelnen Montagetaasks zu erhalten.

Um die Prozesseffizienz der Gestensteuerung bzw. Interaktionszeit mit dem Informationssystem näher zu analysieren, wird zusätzlich zu den GATM-Phasen die Zeit der Gesteninteraktion (*t_Geste*) ermittelt. Die Interaktionszeit mit Betätigen des Buttons am Tablet wird aufgrund eines zu kleinen Zeitanteils nicht extra analysiert.

Beim Experiment 3 wurden die Teilnehmer*innen während des Experiments mit einer Kamera gefilmt, die Aufnahmen anschließend mit dem Open-Source-Video-Annotationstool „ANVIL2“ ausgewertet und die Zeitwerte in MS Excel analysiert. Nach Eliminierung der Ausreißer nach der Formel $Q3 + IQR \cdot 1,5$ und $Q1 - IQR \cdot 1,5$ wurden die Mittelwerte und die statistischen Fehler ausgewertet und zur Interpretation grafisch aufbereitet (Abbildung 115).

Task Completion Time (GATM) inkl. Gestenzeit

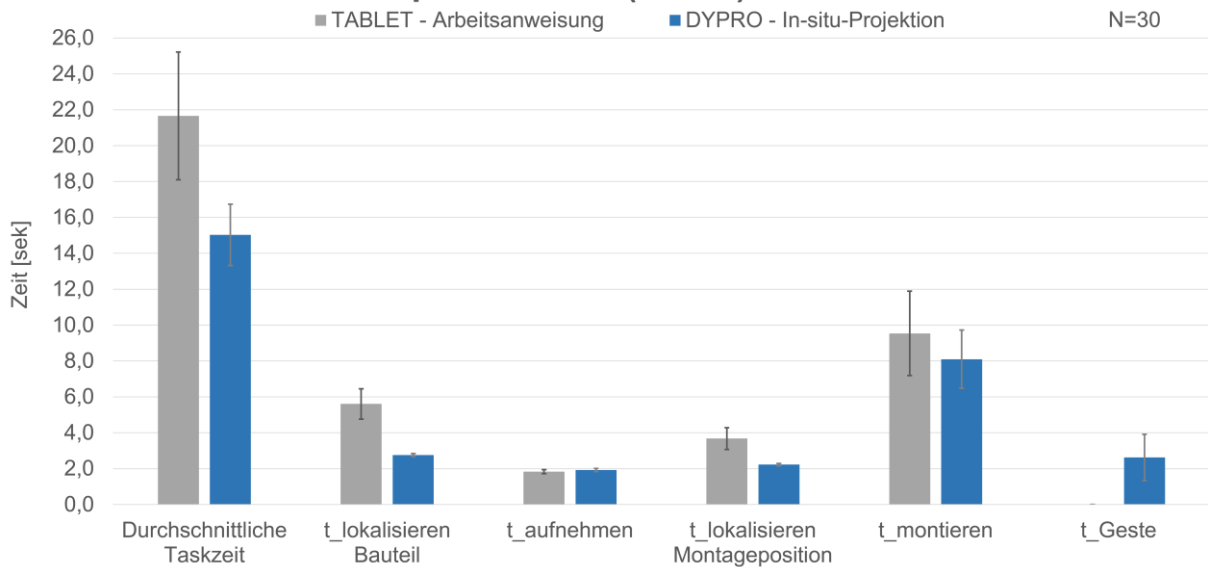


Abbildung 115 | Ergebnisse der „Task Completion Time“ – Experiment 3 (Eigendarstellung)

Die Analyse der Taskzeiten der GATM-Phasen geben einen guten Überblick über die Dauer des Montageprozesses sowie auch über die Verbesserung durch den Einsatz des Assistenzsystems. Dabei konnte festgestellt werden, dass die dynamische In-situ-Projektion bei einer durchschnittlichen Taskzeit von 15 Sekunden im Vergleich zur Anweisung mit dem Tablet mit einer durchschnittlichen Taskzeit von 21,7 Sekunden um 30,6 % besser abschneidet. Es wurde analysiert, dass die Zeit $t_{\text{aufnehmen}}$ bei beiden Varianten mit 1,8–1,9 Sekunden nahezu ident war und die Montagezeit $t_{\text{montieren}}$ bei der In-situ-Projektion um 15,1 % (DYPRO: 8,1 Sek.; Tablet: 9,5 Sek.) schneller bewertet wurde.

Eine starke Verbesserung wurde in den Zeiten $t_{\text{lokalisieren Bauteil}}$ und $t_{\text{lokalisieren Montageposition}}$ festgestellt. Dabei konnte bei der In-situ-Projektion die Lokalisierungszeit des Bauteils um 50,9 % (DYPRO: 2,7 Sek.; Tablet: 5,6 Sek.) und die Lokalisierungszeit der Montageposition um 39,5 % (DYPRO: 2,2 Sek.; Tablet: 3,7 Sek.) verbessert werden. Der Grund der Verbesserung ist auf die dynamische Führung zu den jeweiligen Positionen sowie die direkte Anzeige der Positionen durch das Projektionssystem zurückzuführen, dadurch konnten die Positionen schneller ermittelt werden.

Die Interaktionszeit wurde beim Einsatz der In-situ-Projektion extra analysiert. Die durchschnittliche Zeit für die Gesteninteraktion t_{gesture} lag bei 2,6 und einer Standardabweichung von 1,3 Sekunden, was im Vergleich zur manuellen Weiterschaltung der Projektion („Wizard of Oz“) mit einer durchschnittlichen Interaktionszeit von 1,5 Sekunden und dem Feedback der Teilnehmer*innen als angemessen interpretiert wurde. Die Gestensteuerung wurde in diesem Experiment mit einem synthetisch erzeugten Gestendatensatz umgesetzt (vgl. Kapitel 3.3.3).

Zusammengefasst kann aus der Analyse der Prozesseffizienz mit der „Task Completion Time“ ermittelt werden, dass durch den Einsatz der In-situ-Projektion und dynamischer Führung zur Entnahme- und Montageposition eine erhebliche Verbesserung der

Lokalisierungszeit der Bauteilpositionen an den Bereitstellungsflächen sowie an den exakten Montagepositionen erreicht wurde und damit der Montageprozess durch das schnellere Auffinden der Positionen im Gesamten effizienter gestaltet wurde. Des Weiteren eignet sich die Gestensteuerung für den Einsatz in der Großgeräte- und Baustellenmontage, wobei die durchschnittliche Gestenerkennungszeit im Vergleich zu Experiment 2 um den Faktor 4 reduziert werden konnte.

Qualitative Ergebnisse

Um die subjektiven Eindrücke der Teilnehmer*innen zu ermitteln, wurden auch hier nach dem Experimentdurchgang Kurzinterviews durchgeführt und die Kernaussagen aufgenommen. Ebenso wurden in einem ausgehändigten Fragebogen nach den Herausforderungen sowie positiven und negativen Aspekten gefragt. Die qualitativen Ergebnisse und Kernaussagen werden in Tabelle 27 präsentiert.

Tabelle 27 | Qualitative Ergebnisse – Experiment 3

Erkenntnisse DYPRO – In-situ-Projektion
- Die dynamische Führung zur Entnahme- und Montageposition war effektiv und angenehm. - Durch die Projektion war es einfach, die Bauteile und Montagepositionen zu finden. - Der Einsatz von dynamischen In-situ-Anweisungen, Lichtflächen, Konturen etc. war intuitiv verständlich. - Die Gestensteuerung hat gut funktioniert und war einfach zu erlernen. - Informationsführung ist top! - Gestenerkennung könnte noch schneller und stabiler stattfinden. - Es war teilweise anstrengend, jedes Mal den Arm zur Gesteninteraktion auszustrecken. - Es muss klar sein, wo die Geste überall im Raum ausgeführt werden kann.
Erkenntnisse Tablet-Arbeitsanweisung
- Grundsätzlich gut und verständlich aufbereitete Arbeitsinformationen. - Es musste mehrmals am Tablet nachgesehen werden, wo sich die exakten Montagepositionen befinden. - Es wird gutes räumliches Denken benötigt, um die grafische Information auf die reale Montagestation zu übertragen. - Es benötigt Zeit, um die Bauteile mit den jeweiligen ID-Nummern an den Bereitstellungsflächen zu finden. - Abstände und exakte Positionen im Raum können nur schwer eingeschätzt werden.

Aus den qualitativen Ergebnissen kann ermittelt werden, dass die In-situ-Projektion sowie die dynamische Informationsführung für effektiv und angenehm empfunden wurde und dadurch die Bereitstellungsflächen, die Bauteile sowie Positionen intuitiv gefunden werden konnten. Dabei war der Einsatz der Arbeitsinformationen mit Lichtflächen, Farben und Konturen intuitiv verständlich. Ebenso eignete sich der Einsatz der Gesteninteraktion besonders für die Steuerung der Projektion. Die Erkennung der Gesten könnte laut Aussagen der Testpersonen jedoch noch stabiler und schneller umgesetzt werden. Ebenso wurde angemerkt, dass klar sein müsse, wo im Raum die Gesten ausgeführt werden können, und dass es auf Dauer anstrengend sei, den Arm nach vorne auszustrecken, um die Geste auszuführen.

Aus dem qualitativen Feedback bei der Variante mit der Tablet-Arbeitsanweisung konnte festgestellt werden, dass sie grundsätzlich textlich und grafisch gut aufbereitet sind, jedoch ein gutes räumliche Verständnis erforderlich ist, um die Montagepositionen ermitteln zu können. Des Weiteren kostete es Zeit, die Bauteile mit den jeweiligen ID-Nummern zu finden und anschließend die exakte Montageposition zu ermitteln. Dabei musste mehrmals am Tablet

nachgesehen werden, was wiederum erhöhten Aufwand bedeutete. Ebenso war die Herausforderung, die exakten Positionen und Abstände im Raum einschätzen zu können. Zusammengefasst kann aus dem qualitativen Feedback ermittelt werden, dass die dynamische Informationsführung erhebliche Vorteile für die schnelle Lokalisierung von Bauteilen bringt und die Anzeige von Informationen mittels Projektionssystem als besonders effektiv empfunden wurde. Die Tablet-Anweisung war zwar gut aufbereitet, kostete jedoch mehr Aufwand, um die Bauteile zu finden und die exakten Positionen zu ermitteln.

Zwischenergebnis und Resümee Experiment 3

Aus der Durchführung und Evaluierung des Experiments 3 konnte festgestellt werden, dass die In-situ-Projektion und dynamische Informationsführung auch in einem weiteren und komplexeren Use Case erhebliche Verbesserungen im Bereich der Lokalisierungszeiten sowie der Reduktion der kognitiven Arbeitsbelastung mit sich bringt. Neben der Verbesserung des Zeitanteils $t_{\text{lokalisieren Montageposition}}$ um 39,5 % konnte zusätzlich durch die dynamische Führung zur Entnahmeposition der Zeitanteil $t_{\text{lokalisieren Bauteil}}$ um 50,9 % im Vergleich zur Tablet-Arbeitsanweisung verbessert werden.

Die Gestensteuerung (Gestenerkennungszeit) konnte im Vergleich zum vorherigen Experiment um den Faktor 4 verbessert werden und eignet sich demnach mit einer Erkennungszeit von rund 2,6 Sekunden für den Einsatz in der Großgeräte- und Baustellenmontage. Potenzial besteht noch in der weiteren Verbesserung der Gestenerkennungszeit, der Erweiterung der Interaktionsmöglichkeiten sowie des Designs, wo und in welcher Form Gesten im großen Arbeitsraum am effektivsten eingesetzt werden können.

Durch das Experiment 3 wurde gezeigt, dass sich ein Spatial Augmented Reality Assistenzsystem mit In-situ-Projektion und nutzeradaptiver Gesteninteraktion besonders für den Einsatz in der Großgeräte- und Baustellenmontage eignet und dementsprechend auch auf weitere Use Cases skaliert werden kann.

5.3 Zusammenfassung der Ergebnisse der Evaluierung

Ein zentrales Ergebnis des Design-Prozesses sind die Erkenntnisse aus den Evaluierungsexperimenten. Vor allem wurde durch die Evaluierung des Spatial Augmented Reality Assistenzsystems der Mehrwert für den Menschen und für den Prozess quantifiziert und aufgezeigt. Dabei stellten die Experimente 1 und 2 am Demonstrator der Fan-Cowl-Montage die ersten Iterationen des Designprozesses dar und dienten zum ersten Einsatz des dynamischen Projektionssystems in der Baustellenmontage (Experiment 1) und zur Evaluierung der dynamischen In-situ-Projektion mit Informationsführung (Experiment 2). Die Erkenntnisse aus den ersten zwei Iterationen wurden bei der Gestaltung des Experiments 3 integriert, welches mit dem Demonstrator „Spielzeugbaustein-Montage“ zur Verallgemeinerung der Ergebnisse und finalen Evaluierung des Assistenzsystems diente. Die Ergebnisse aus den Experimenten 1 und 2 werden in Tabelle 28 zusammengefasst. Anschließend werden die finalen Evaluierungsergebnisse des Experiments 3 präsentiert.

Tabelle 28 | Zusammenfassung der Ergebnisse aus Experiment 1 und 2

Evaluierungsergebnisse Experiment 1 Fan-Cowl-Montage: Informationsbereitstellung über ein Projektionssystem, Interaktion mittels projizierten Buttons („Wizard of Oz“) Baseline: Informationsbereitstellung und Interaktion über PC-Terminal und statischem 2D-Laser
Prozesszeit: Verbesserung der Prozesszeit um rund 4 % Usability: Total SUS Score von 85,3 „Excellent“ (Baseline: 71,1 „OK“) Technologieakzeptanz: TAM Score mit 5 ++ (Baseline: 4+) Präferenz der Teilnehmer*innen für das Projektionssystem: #14/#16 Qualitative Ergebnisse Projektionssystem: - Information und Interaktion direkt am Werkstück und in unmittelbarer Nähe des Montageortes ist komfortabel; dadurch Vermeidung langer Wege, um an Information zu gelangen; Informationen werden nicht übersehen. - Intuitiv verständlich zu bedienen. - Vereinfachung von gleichen Informationen im Zeitverlauf des Montageprozesses wäre sinnvoll (weniger textliche Informationen, Anzeige über Piktogramme, nur Informationen über Änderungen). Kernergebnisse des Experiments Der Einsatz des dynamischen Projektionssystems zur Anzeige von Arbeitsinformationen über einem großen Arbeitsraum direkt in der Baustellenmontage bietet eine potenzielle Lösung für den Einsatz in der Großgeräte- und Baustellenmontage und vor allem einen Mehrwert für den Menschen in Bezug zur Usability, Technologie-Akzeptanz und Präferenz der Versuchsteilnehmer*innen. Eine weitere Erkenntnis war, dass die Prozesszeit zwar etwas verbessert werden konnte, jedoch zur Evaluierung der Prozesszeit eine detailliertere Evaluierungsmethode eingesetzt werden muss, um nähere Details über die Prozesseffizienz ermitteln zu können. Durch die Evaluierung wurden relevante Erkenntnisse für die weitere Gestaltung des Assistenzsystems ermittelt, welche vor allem im Bereich der Informationsgestaltung, In-situ-Projektion und Informationsführung sowie der nutzeradaptiven Interaktion wertvolle Inputs lieferte. Damit kann durch die Evaluierungsergebnisse die Forschungsfrage des Experiments „Welchen Mehrwert bietet die dynamische Informationsbereitstellung und Interaktion direkt am Montageplatz gegenüber der Bereitstellung und Interaktion mit PC-Terminal und statischer Laserprojektion?“ beantwortet werden. Weiterer Forschungsbedarf Das Experiment 1 diene als Grundlage für den Einsatz des dynamischen Projektionssystems als Informationsassistent in der Großgeräte- und Baustellenmontage und lieferte relevante Inputs für den weiteren, im Folgenden angeführten Forschungsbedarf: - Gestaltung und Evaluierung der Usability und Prozesseffizienz der In-situ-Projektion - Gestaltung und Evaluierung von verbesserten Arbeitsinformationen - Gestaltung und Evaluierung einer Gesteninteraktion für die Anforderungen der Baustellenmontage - Verbesserung der Prozesszeit und Anwendung einer detaillierteren Evaluierung der Prozesseffizienz
Evaluierungsergebnisse Experiment 2 Fan-Cowl-Montage: Dynamische In-situ-Projektion mit Informationsführung zur Montageposition, Interaktion mittels Gestensteuerung Baseline: Anzeige von Informationen mit statischer In-view-Projektion
Prozesszeit und Gestenzeit: - Verwendung der „Task Completion Time“ nach dem GATM zur detaillierten Evaluierung der Prozesseffizienz (vgl. Kapitel 5.1.2) - Verbesserung der durchschnittlichen Taskzeit um rund 10 % $t_{\text{aufnehmen}}$ und $t_{\text{montieren}}$ sind nahezu identisch, was auf die gleiche Komplexität der Arbeitsaufgaben zurückzuführen ist. - Durchschnittliche Zeit der Gesteninteraktion t_{gesture} von rund 10 sek. 14 von 144 (10 %) Gesteninteraktionen dauerten länger als der Durchschnitt und wurden als Fehler interpretiert. $t_{\text{lokalisieren Bauteil}}$ war bei der Baseline-Variante um durchschn. 2 Sekunden schneller.

Kernergebnis:

***t_lokalisieren Montageposition* war bei der In-situ-Projektion mit Führung zur Montageposition um den Faktor 2 niedriger ist als bei der Baseline-Variante.**

Usability: Total SUS Score von 84,6 „GOOD“ (Baseline: 81,7 „GOOD“)

Kognitive Arbeitsbelastung:

- Total RTLX Score 31,1 Punkten (Baseline: 32,2 Punkte)
 - Leistungsfähigkeit (Performance) mit 25,0 Punkten (Baseline: 31,7 Punkte)
 - Frustrationslevel mit 38,3 Punkten (Baseline: 43,3 Punkte)
 - Bemühung/Aufwand „Effort“ mit 30,0 Punkten (Baseline: 33,3 Punkte)
 - Mentale Belastung (Mental) mit 31,7 Punkten (Baseline: 31,7 Punkte)
 - Bezüglich der physischen (Physical) und temporären Belastung (Temporal) liegt die Baseline-Variante etwas vorne.

Qualitative Ergebnisse des Projektionssystems:

- Die angezeigte Montageposition und -ausrichtung unterstützte bei der Durchführung des Vorgangs.
- Dass sich die Projektion zur nächsten Position auf die andere Seite der Montage bewegte, sodass der Projektion gefolgt werden musste, bevor das Bauteil aus der Bereitstellungsfläche entnommen werden konnte, war irritierend. Dies zeichnete sich in der Zeit *t_lokalisieren Bauteil* ab.
- Die Gestensteuerung war großartig, könnte aber in Bezug auf Schnelligkeit und Genauigkeit noch verbessert werden.
- Die Feedback-Folien zur Bestätigung der Aufgaben könnten schneller erscheinen, um schneller mit den nächsten Aufgaben fortzufahren.
- Die Piktogramme waren verständlich und sollten noch öfters und anstatt von Text verwendet werden.
- Nach ein paar Aufgaben können die Informationen vereinfacht werden, d. h. nur noch Piktogramme Verwendung finden.
- Am besten wäre eine Mischung aus beiden Versuchen, d. h. die Position und Orientierung mit SAR und die Arbeitsanweisungen immer im Blickfeld zu zeigen.
- Baseline: Langes Suchen nach Montagepositionen war nicht komfortabel.
- Baseline: Es war gut, dass die Arbeitsinformationen immer im Blickfeld des Menschen waren.

Kernergebnisse des Experiments

Die Kernergebnisse von Experiment 2 sind, dass sich die dynamische In-situ-Projektion mit Informationsführung besonders für die schnelle Ermittlung von Positionen (hier der Montageposition) im großen Arbeitsraum eignet, was durch die Verbesserung der Lokalisierungszeit um den Faktor 2 bei gleichzeitiger Reduktion der kognitiven Arbeitsbelastung bestätigt wurde. Demnach eignet sich die dynamische In-situ-Projektion und Informationsführung zur Montageposition besonders für den Einsatz in der Großgeräte- und Baustellenmontage. Ebenso war die Erkenntnis, dass sich die Gesteninteraktion mit natürlichen Gesten für den Einsatz im großen Raum eignet und zur Steuerung des Spatial Augmented Reality Assistenzsystems eingesetzt werden kann. Durch die Durchführung des Experiments konnte die Forschungsfrage „Was sind die Vorteile der Informationsbereitstellung mittels dynamischer In-situ-Projektion und Informationsführung zur Montageposition im Vergleich zur statischen In-view-Projektion, bei Einsatz eines Spatial Augmented Reality Systems in der industriellen Baustellenmontage?“ beantwortet werden.

Weiterer Forschungsbedarf

Weiterer Forschungsbedarf entsteht vor allem in der Verbesserung der Gestenerkennung und des Prozesses der Erstellung des Gestendatensatzes. Ebenso besteht der Bedarf, die dynamische In-situ-Projektion mit Informationsführung auf weitere Positionen wie die Anzeige der Montagebauteile an den Bereitstellungsflächen zu erweitern und die Arbeitsinformationen noch weiter zu verbessern. Vor allem zeigt sich die Notwendigkeit, die gewonnenen Erkenntnisse der Informationsbereitstellung auf einen weiteren Use Case mit Charakteristika der Großgeräte- und Baustellenmontage anzuwenden und dort den Mehrwert zu validieren.

Die Erkenntnisse aus den Experimenten 1 und 2 wurden in die finale Gestaltung des Assistenzsystems in Experiment 3 integriert. Die Ergebnisse der finalen Evaluierung werden in Tabelle 29 zusammengefasst.

Tabelle 29 | Ergebnisse und Erkenntnisse aus der finalen Evaluierung (Experiment 3)

Evaluierungsergebnisse Experiment 3 Spielzeugbaustein-Montage: Dynamische In-situ-Projektion mit Informationsführung zur Entnahme- und Montageposition, Interaktion mittels Gestensteuerung
Baseline: Informationsbereitstellung über statisches Tablet

Prozesszeit und Gestenzeit:

- **Verbesserung der durchschnittlichen Taskzeit um 30,6 %**
- *t_{aufnehmen}* war beiden Varianten mit 1,8–1,9 Sekunden nahezu ident.
- **Verbesserung der Montagezeit *t_{montieren}* um 15,1 %**
- **Starke Verbesserung der Lokalisierungszeit des Bauteils *t_{lokalisieren Bauteil}* um 50,9 %**
- **Verbesserung der Lokalisierungszeit der Montageposition *t_{lokalisieren Montageposition}* um 39,5 %**
- **Verbesserung der durchschnittlichen Zeit für die Gesteninteraktion *t_{gesture}* auf 2,6 Sekunden**

Usability: Total SUS Score von 81,5 „GOOD“ (Baseline: 80,5 „GOOD“)

Kognitive Arbeitsbelastung: Total RTLX Score mit 33,2 Punkten (Baseline: 37,3 Punkte)

- **Leistungsfähigkeit (Performance) mit 26,0 Punkten (Baseline: 34,0 Punkte)**
- Frustrationslevel mit 35,3 Punkten (Baseline: 35,3 Punkte)
- Bemühung/Aufwand „Effort“ mit 37,3 Punkten (Baseline: 38,0 Punkten)
- **Mentale Belastung (Mental) mit 32,0 Punkten (Baseline: 40,7 Punkte)**
- Physische (Physical) mit 33,3 Punkten (Baseline: 32,0 Punkte)
- **Temporäre Belastung (Temporal) mit 35,3 Punkten (Baseline: 44,0 Punkte)**

Qualitative Ergebnisse des Projektionssystems:

- Die dynamische Führung zur Entnahme- und Montageposition war effektiv und angenehm.
- Durch die Projektion war es unkompliziert, die Bauteile und Montagepositionen zu finden.
- Der Einsatz von Lichtflächen, Konturen und Farben war intuitiv verständlich.
- Die Gestensteuerung hat gut funktioniert und war intuitiv zu bedienen.
- Informationsführung war top!
- Es war teilweise anstrengend, jedes Mal den Arm zur Gesteninteraktion auszustrecken.
- Es muss klar sein, wo die Geste überall im Raum ausgeführt werden kann.

Kernergebnisse des Experiments

Die Kernergebnisse von Experiment 3 sind die Anwendung und Evaluierung der Spatial Augmented Reality in einem weiteren sowie komplexeren Use Case der Großgeräte- und Baustellenmontage und vor allem die Bestätigung des Mehrwerts für Mensch und Prozess durch erhebliche Verbesserung der Lokalisierungszeit des Bauteiles und der Montageposition bei gleichzeitiger Reduktion der Arbeitsbelastung. Ebenso war die Erkenntnis, dass die gestaltete Gesteninteraktion mit einer akzeptablen Erkennungszeit für die Baustellenmontage eingesetzt werden kann.

Weiterer Forschungsbedarf

Weiterer Forschungsbedarf besteht bei der Verbesserung der Ergonomie der Gesteninteraktion und Erweiterung der Interaktionsmöglichkeiten. Demnach besteht der Bedarf, mit dem Projektionssystem darzustellen, wo überall Gesten im Arbeitsraum ausgeführt werden können. Für die Informationsbereitstellung könnte noch eine prozessintegrierte Funktion berücksichtigt werden, bei der erkannt wird, ob das Bauteil entnommen bzw. montiert wurde, um damit auch die Häufigkeit der Interaktionen zu reduzieren.

Die Zusammenfassung zeigt die Ergebnisse der Evaluierungsexperimente, welche vor allem den Mehrwert des gestalteten Assistenzsystems für den Menschen und den Prozess bestätigen. Im folgenden Kapitel werden die Ergebnisse des Designprozesses sowie der Evaluierung diskutiert und anschließend die Forschungsfragen beantwortet.

6. Diskussion der Ergebnisse

Das Ergebnis des iterativen Design-Prozesses ist die Gestaltung, Demonstration sowie Evaluierung des **Spatial Augmented Reality Assistenzsystems mit dynamischer In-situ-Projektion und nutzeradaptiver Gesteninteraktion** als zentrales Artefakt zur **Verbesserung der Informationsbereitstellung und Interaktion** in der Großgeräte- und Baustellenmontage. Neben dem technischen Aufbau der Artefakte ist das zentrale Ergebnis die Gestaltung und Integration der dynamischen In-situ-Projektion mit Informationsführung sowie Interaktion mit natürlichen Gesten im großen Arbeitsraum, welche im Laufe der Iterationen stetig verbessert wurden. Die Erkenntnisse aus den Ergebnissen des Designprozesses sowie die kritische Diskussion der Evaluierung des Experiments stellen den folgenden Teil der Arbeit dar. Anschließend wird auf die Limitationen der Forschungsergebnisse eingegangen und die Forschungsfragen beantwortet.

6.1 Erkenntnisse aus dem Design der Artefakte

Durch den iterativen Designprozess wurden die Erkenntnisse der Evaluierungsergebnisse und der Aufbau der Demonstration stetig in die Gestaltung der Artefakte integriert. Tabelle 30 stellt eine Zusammenfassung der entwickelten Artefakte dar und präsentiert die Optimierungen und Erweiterungen des Spatial Augmented Reality Assistenzsystems für die Großgerätemontage.

Tabelle 30 | Ergebnisse aus dem Design der Artefakte

Artefakte des Spatial Augmented Reality Assistenzsystems
- Integration des dynamischen Projektionssystem aus der Veranstaltungstechnik für die Anforderungen der industriellen Großgeräte- und Baustellenmontage als zentrales Informationsbereitstellungssystem (vgl. Kapitel 3.2.1) - Integration des visuellen Interaktionssystem mit hochauflösendem RGB-Videostream der RealSense-Kamera (D435i) und dem echtzeitfähigen Objekterkennungsalgorithmus YOLO (vgl. Kapitel 3.2.2)
Erweiterungen und Optimierungen der Artefakte
- Optimierung des Erstellungsprozesses des Gestendatensatzes (Synthetisch erzeugter Gestendatensatz) (vgl. Kapitel 3.3.3) - Integration „Adaptive Homographie“ – Automatisierte Einstellung der Verzerrung mit Verwendung der Stereokamera ZED-2 (vgl. Kapitel 3.2.1)

Dynamisches Projektionssystem

Das recherchierte dynamische Projektionssystem aus der Veranstaltungsbranche (vgl. Kapitel 1.2.2) wurde für die Anforderungen der industriellen Großgeräte- und Baustellenmontage als zentrales Informationsbereitstellungssystem adaptiert und in der Pilotfabrik aufgebaut. Die Einsatzmöglichkeiten, Effizienz, Akzeptanz und Usability des dynamischen Projektionssystems wurden anschließend im Vergleich zur State-of-the-Art-Bereitstellung mittels PC-Terminal evaluiert. Dabei waren vor allem die Auswahl der Spezifikationen des Projektors entscheidend, um damit laut den Gegebenheiten der Montagehalle (Deckenhöhe, Lichtverhältnisse etc.) und für die Größe des Projektionsraumes

ein optimales System auszulegen. Durch die enge Zusammenarbeit mit den Experten des Dynamic Projection Institute konnte ein geeigneter Projektor ausgewählt werden, welcher sich vor allem für die Gegebenheiten der großen Arbeitsfläche und für den Einsatz bei Tageslicht eignet. Essenziell dabei war die Verfügbarkeit von geeigneten digitalen Schnittstellen und einer steuerbaren motorbetriebenen Linse, welche von der Ferne aus bedient werden kann. Ein weiterer Punkt war, eine geeignete Authoring- und Mediensoftware zu verwenden, um die Arbeitsinformationen schnellstmöglich in das System zu integrieren und damit auch die Verzerrung je nach Position im Arbeitsraum zu ermöglichen. Dabei wurde die mit dem Projektor kompatible Software (vgl. Kapitel 3.2.1) verwendet, welche sich für die Einstellung der manuellen Verzerrung eignet. Um den Aufwand der manuellen Einstellung der Verzerrung zu optimieren, wurde dieser Schritt versucht zu automatisieren und das Feature „Adaptive Homographie“ implementiert und mit dem Projektionssystem verbunden. Dieses ermöglicht die automatisierte Einstellung der Verzerrung zu vorab definierten und kalibrierten Projektionspositionen (vgl. Kapitel 3.2.1). Durch die intuitive Integration der Arbeitsanweisungen in das Projektionssystem und die Möglichkeit der Erstellung von Projektionspositionen inklusive deren Verzerrungen, können über die Bewegung des Spiegelsystems, Arbeitsinformationen zum Montageprozess über einen großen Arbeitsraum bereitgestellt und der Mensch gezielt durch den Prozess geführt werden.

Visuelles Interaktionssystem

Ein weiterer Gestaltungsaspekt war die Integration einer Mensch-Maschine-Interaktion mit der Verwendung von natürlichen Gesten (vgl. Kapitel 1.2.3), um damit das Projektionssystem überall im Arbeitsraum bzw. an der jeweiligen Montageposition steuern zu können. Dabei war die Entwicklung des visuellen Interaktionssystems, welches die Gesten des Menschen erkennt und damit eine Möglichkeit zur Steuerung des Projektionssystems bietet, das zentrale Gestaltungsartefakt. Das erste Konzept mit projizierten interaktiven Buttons („Wizard-of-Oz“) wurde in Experiment 1 getestet und das Potenzial der direkten Interaktion in der Montage ermittelt. Das Interaktionssystem mit Gestenerkennung wurde in Experiment 2 für die Verwendung in der Fan-Cowl-Montagestation integriert und die Interaktion dabei mit zwei Gesten evaluiert. Dabei wurde das RGB-Videomodul der RealSense-Kamera zur Aufnahme der Montagestation an der Decke neben dem Projektionssystem platziert und mit dem echtzeitfähigen Objekterkennungsalgorithmus YOLO verknüpft (vgl. Kapitel 3.2.2). Hierbei war die Herausforderung, einen für industrielle Anwendungen geeigneten und bereits getesteten Algorithmus zu finden, welcher unkompliziert zu implementieren ist. Über den Zugang zu Hochleistungscomputer der TU Wien (vgl. Kapitel 3.3.3) konnte der Deep-Learning Algorithmus trainiert werden.

Der Einsatz des visuellen Interaktionssystems ermöglicht die nutzeradaptive Gesteninteraktion über einen großen Raum und direkt an der Montagestation. Das integrierte Kamerasystem und Algorithmen zur Gestenerkennung sowie der eigens erstellte Gestendatensatz eigneten sich vor allem in Experiment 3 zur intuitiven und schnellen Interaktion mit dem Assistenzsystem. Auch die Optimierungen in Bezug auf die Gestendatensatzerstellung und Entwicklung eines synthetisch erzeugten Gestendatensatzes

ermöglichen die schnelle Erweiterung auf neue Gestenarten sowie die Verwendung der Gesteninteraktion in anderen Montagesettings. Die flexible Gestaltung hat einen positiven Einfluss auf die Skalierbarkeit und vielfältige Verwendung im industriellen Umfeld.

Spatial Augmented Reality System: Kostenoptimierung und Return on Investment (ROI)

Durch die Gestaltung der Artefakte wurde ein Assistenzsystem mit Gesamtanschaffungskosten von 15 300,10 Euro (vgl. Tabelle 14) entwickelt, welches neue Möglichkeiten zur Informationsbereitstellung und Interaktion in der industriellen Großgeräte- und Baustellenmontage bietet, die Prozesseffizienz (vgl. Abbildung 115) verbessert und die kognitive Arbeitsbelastung (vgl. Abbildung 113) reduziert. Das Spatial Augmented Reality System bietet im Vergleich zu kommerziellen Aufbauten mit mehreren Projektoren (vgl. Kapitel 3.2.1 und Abbildung 45) eine kostengünstige Variante und bringt im Weiteren, durch die Verbesserung der durchschnittlichen Taskzeit, eine Kostenoptimierung des Montageprozesses. In Tabelle 31 werden die Ergebnisse aus Experiment 3 (vgl. Kapitel 5.2.3) herangezogen und die Kostenoptimierung, resultierend aus der Verbesserung der durchschnittlichen Taskzeit, beispielhaft am Use Case „Spielzeugbaustein-Montage“ aufgezeigt. Außerdem werden die Gesamtanschaffungskosten des SAR-Systems der Kostenoptimierung gegenübergestellt und eine vereinfachte ROI-Berechnung durchgeführt.

Tabelle 31 | Berechnung der Kostenoptimierung und des Return on Investments

Grunddaten			
Stundensatz für eine/n Montagemitarbeiter*in	31,56 Euro ²⁸		
Stundenanzahl pro Jahr	1720 Stunden ²⁹		
Montage Use Case	Spielzeugbausteinturm Montage mit 18 Montagetasks		
Berechnung der Verbesserung			
	Tablet- Anweisung	Spatial Augmented Reality	Verbesserung durch SAR
durchschnittliche Taskzeit in Sekunden ³⁰	21,7	15	6,7
Gesamtmontagezeit für 18 Montagetasks in Sekunden	390,6	270	120,6
Gesamtmontagezeit für 18 Montagetasks in Stunden	0,109	0,075	0,034
Mögliche Spielzeugbausteintürme pro Stunde	9,17	13,33	4,16
Mögliche Spielzeugbausteintürme pro Jahr (1720h)	15 772	22 928	7156
Kosten pro Stück (Personalkosten)	3,44 Euro	2,37 Euro	Verbesserung von 1,07 Euro
Return on Investment (ROI) Berechnung			
Gesamtanschaffungskosten Spatial Augmented Reality System	15 300,10 Euro		
Break-Even-Stückzahl	14 260 Türme		
Return-on-Investment (Jahre)	0,62 Jahre		

²⁸ Ein Jahresgehalt von 44 800,00 Euro brutto (3200 Euro brutto / Monat) und jährliche Lohnnebenkosten von 9479,04 Euro ergibt eine Summe von 54 279,04 Euro Personalkosten pro Jahr. Resultierender Stundensatz von 31,56 Euro (54 279,04 Euro dividiert durch 1720 Jahresstunden). Online-Rechner verfügbar unter: <https://onlinerechner.haude.at/Brutto-Netto-Rechner>.

²⁹ pauschalierter Jahresstundenteiler für Vollzeitbeschäftigte in industriellen Forschungsprojekten; vgl. <https://www.ffg.at/recht-finanzen/kostenleitfaden/version-21>

³⁰ vgl. Tabelle 29 und Ergebnisse des Experiments 3

Aus der Berechnung konnte ermittelt werden, dass durch den Einsatz des Spatial Augmented Reality Systems 4,16 Spielzeugbausteintürme pro Stunde und 7156 Türme pro Jahr zusätzlich hergestellt werden können. Des Weiteren ergibt sich eine Kostenoptimierung von 1,07 Euro pro hergestelltem Spielzeugbausteinurm, wenn eine Montageperson mit einem Stundensatz von 31,56 Euro dafür eingesetzt wird. Bei Gesamtanschaffungskosten von 15 300,10 Euro wurde eine Break-Even-Stückzahl von 14 260 Türmen ermittelt, woraus sich ein Return on Investment von 0,62 Jahren, bei einer Jahresstückzahl von 22 928 Türmen, ergibt. Bei der Berechnung wurden ausschließlich die Gesamtanschaffungskosten des Spatial Augmented Reality Systems laut Tabelle 14 herangezogen. Anfallende Schulungs-, Installations-, sowie Arbeitsvorbereitungskosten wurden in der Berechnung nicht berücksichtigt, sollten jedoch je nach Anwendungsfall im Unternehmen miteinbezogen werden.

6.2 Erkenntnisse aus der Gestaltung der dynamischen Informationsbereitstellung

Die Gestaltung und Verbesserung der dynamischen Informationsbereitstellung (vgl. Kapitel 3.3.1 und 3.3.2) durch den Einsatz des Spatial Augmented Reality Assistenzsystems stellt einen wesentlichen Mehrwert für die Großgeräte- und Baustellenmontage dar (vgl. Kapitel 1.2.1). Dabei konnte durch die Evaluierungen festgestellt werden, dass sich die intuitive Gestaltung der Arbeitsinformationen sowie die dynamische In-situ-Projektion positiv auf die Prozesseffizienz sowie auf die Reduktion der kognitiven Arbeitsbelastung des Menschen auswirkt. Folgend wird ein Überblick über die integrierten und evaluierten Aspekte der dynamischen Informationsbereitstellung dargestellt (Tabelle 32).

Tabelle 32 | Erkenntnisse der Gestaltung der dynamischen Informationsbereitstellung

Entwicklungen für die Verbesserung der Informationsbereitstellung in der Baustellenmontage
▪ Intuitive und projektionsoptimierte Gestaltung der Arbeitsanweisungen und Informationen zum Montageprozess (vgl. Kapitel 3.3.1): ○ intuitives Design mit Piktogrammen, Symbolen und Zeichen ○ Farb- und Kontrastgestaltung ○ dynamische Animation mit bewegten Pfeilen ▪ Verzerrungsfreie Projektion durch manuelle oder automatisierte Einstellung der Verzerrung ▪ Software-Mapping der Projektionspositionen und Projektionswege im großen Arbeitsraum und auf amorphen Oberflächen ▪ Integration und Evaluierung der dynamischen In-situ-Projektion und Informationsführung für die Großgeräte- und Baustellenmontage (vgl. Kapitel 3.3.2): ○ dynamische In-situ-Projektion von Montage- und Bauteilpositionen im zwei- und zweieinhalbdimensionalen Arbeitsraum ○ dynamische Informationsführung zur Montage- und Entnahmeposition ○ Projektion exakter Positionen, Orientierung und Lage ○ Multi-Point-Projektion mit gleichzeitiger Anzeige der Montageposition und Informationen zur Gesteninteraktion ○ Anzeige der Wege durch einen kontinuierlichen Projektionsstrahl

Verbesserung der Informationsbereitstellung in Bezug zur Prozesseffizienz
Verbesserung der durchschnittlichen Taskzeit durch Einsatz des Spatial Augmented Reality Assistenzsystems ▪ Verbesserung der durchschnittlichen Taskzeit um 30,6 % Verbesserung der aufgabenabhängigen Zeiten durch intuitive und projektionsoptimierte Informationsgestaltung ▪ Verbesserung der Zeit zum Montieren der Bauteile um 15,1 % Verbesserung der aufgabenunabhängigen Zeiten durch Einsatz der dynamischen In-situ-Projektion und Informationsführung: ▪ Starke Verbesserung der Lokalisierungszeit des Bauteils um 50,9 % ▪ Verbesserung der Lokalisierungszeit der Montageposition um 39,5 %
Mehrwert der Informationsbereitstellung für den Menschen
▪ Akzeptable und gute Usability des eingesetzten Assistenzsystems ▪ Reduktion der kognitiven Arbeitsbelastung - Total RTLX Score von 33,2 Punkten (Baseline: 37,3 Punkte) ○ Leistungsfähigkeit (Performance) mit 26,0 Punkten (Baseline: 34,0 Punkte) ○ Bemühung/Aufwand (Effort) mit 37,3 Punkten (Baseline: 38,0 Punkte) ○ Mentale Belastung (Mental) mit 32,0 Punkten (Baseline: 40,7 Punkte) ○ Temporäre Belastung (Temporal) mit 35,3 Punkten (Baseline: 44,0 Punkte)

Durch die Gestaltung der intuitiven und projektionsoptimierten Informationsgestaltung konnten die Inhalte der Arbeitsanweisungen verständlich auf die Umsetzung der Montagetätigkeiten transferiert werden, was sich vor allem in der Verbesserung der Montagezeit und der durchschnittlichen Taskzeit zeigte. Ebenso konnte aufgrund der Gestaltung der Arbeitsinformationen eine gute Usability sowie Reduktion der kognitiven Arbeitsbelastung umgesetzt werden. Aus dem qualitativen Feedback der Experimente, den Demonstrationen sowie aus der Literatur konnten dazu Inputs gewonnen werden, wie Informationen intuitiv und vor allem für den großen Arbeitsraum projektionsoptimiert gestaltet werden können. Dabei wurden kontrastreiche Farben, eine vereinheitlichte Farbgestaltung, eine klare Struktur der Arbeitsanweisungen sowie ein gezielter Einsatz von Piktogrammen, Symbolen und Animationen verwendet. Die Ergebnisse und das Feedback aus der Evaluierung zeigte, dass die finale Gestaltung der Arbeitsinformationen intuitiv verständlich war und sich optimal für die Anwendung in der Großgeräte- und Baustellenmontage eignet. Bei der Gestaltung der Informationsbereitstellung sind vor allem die Gegebenheiten des Arbeitsraumes, des Montageobjektes und dessen Projektionsoberfläche zu berücksichtigen. Allen voran ist auf die Komplexität des Montageprozesses zu achten. Dementsprechend sind die Anforderungen der Montagestation zu analysieren und die Informationsgestaltung nach den Erkenntnissen (vgl. Kapitel 3.3.1) anzupassen.

Durch die dynamische In-situ-Projektion und Informationsführung zur Entnahme- und Montageposition konnten vor allem die aufgabenunabhängigen Zeiten (vgl. Kapitel 5.1.2), welche stark von der Art der Informationsbereitstellung und dem Einsatz eines Assistenzsystems abhängen, deutlich verbessert werden. Durch die gezielte Projektion wurde der Mensch einerseits beim Auffinden der Bauteile an den Bereitstellungsflächen und andererseits bei der Lokalisierung der exakten Montagepositionen unterstützt, was sich als deutliche Reduktion der Lokalisierungszeiten abzeichnete. Bei der Gestaltung der

In-situ-Projektion ist die Hauptherausforderung, dass eine verzerrungsfreie Projektion überall im Arbeitsraum und auch am dreidimensionalen Montageobjekt gewährleistet werden muss. Durch die Möglichkeiten der Authoring- und Medienbearbeitungssoftware (vgl. Kapitel 3.2.1) war eine intuitive Einstellung der Verzerrung je nach Position der Projektion und Oberflächenform möglich. Die weitere Entwicklung der adaptiven Homographie unterstützte bei der automatisierten Einstellung der Verzerrung, benötigt jedoch Vorwissen und erhöhten Kalibrierungsaufwand als bei der manuellen Verzerrung. Eine besondere Herausforderung liegt auch im schnellen Auffinden von Projektionen im Raum sowie im Nachfolgen der Informationsführung. Dabei liegt ein besonderes Augenmerk beim Einsatz der dynamischen Animation (vgl. Abbildung 69) und der kontinuierlichen und dynamischen Führung durch den Projektionsstrahl von Position zu Position, welche vor allem im großen Raum beim schnellen Auffinden der Projektion unterstützte.

Ein weiterer zentraler Aspekt bei der Gestaltung der Informationsbereitstellung war, eine gute Usability des Assistenzsystems zu gewährleisten und eine Verbesserung der kognitiven Arbeitsbelastung zu erreichen. Dabei konnte das Assistenzsystem durch eine gute Usability überzeugen. Vor allem zeichnete sich eine Reduktion der kognitiven Arbeitsbelastung ab, welche vor allem bei den Aspekten Leistungsfähigkeit, Aufwand und mentale und temporäre Belastung eine starke Reduktion erzielen konnte. Der Grund dafür war vor allem die intuitive Informationsgestaltung sowie die dynamische Informationsführung durch den Montageprozess. Hier konnten die Bauteile im großen Arbeitsraum ohne viel Aufwand gefunden und auf den jeweilig angezeigten projizierten Positionen montiert werden. Die Verbesserung der Informationsgestaltung konnte vor allem durch die Inputs der Experimentteilnehmer*innen der Experimente 1 und 2 erreicht werden, welche in das System integriert wurden und sich in der finalen Evaluierung in Experiment 3 ebenso positiv auf die kognitive Arbeitsbelastung und Usability auswirkten.

Als Erkenntnis aus der Gestaltung der dynamischen Informationsbereitstellung kann festgestellt werden, dass sich das Spatial Augmented Reality Assistenzsystem für die Informationsbereitstellung eignet, das volle Potenzial für den Einsatz in der Großgeräte- und Baustellenmontage jedoch erst durch die iterative Gestaltung der dynamischen Informationsbereitstellung, die intuitiv und projektionsoptimierte Arbeitsinformationen und eine durchdachte In-situ-Projektion und Informationsführung erreicht wurde. Eine potenzielle Weiterentwicklung der Informationsbereitstellung könnte eine prozessintegrierte Steuerung des Informationsbereitstellungssystems sein, welche erkennt, ob ein Bauteil entnommen bzw. montiert wurde, um damit die Anzeige der nächsten Arbeitsinformation einzuleiten. Des Weiteren besteht Potenzial zur Weiterentwicklung des Authoringprozesses, bei dem die Arbeitsinformationen, Projektionspositionen und deren Verzerrung direkt in der Montagestation intuitiv eingestellt werden kann.

6.3 Erkenntnisse aus der Gestaltung der Gesteninteraktion

Neben der Gestaltung der dynamischen Informationsbereitstellung sowie der technischen Entwicklung und Integration des visuellen Interaktionssystems war die Gestaltung der

Gesteninteraktion mit natürlichen Gesten zur Steuerung des dynamischen Projektionssystems ein essenzieller Teil des Designprozesses (vgl. Kapitel 3.3.3). Folgend werden die Erkenntnisse und Gestaltungsaspekte der nutzeradaptiven Gesteninteraktion zusammengefasst und die Verbesserungen und Entwicklungen dargestellt (Tabelle 33).

Tabelle 33 | Erkenntnisse der Gestaltung der Gersteninteraktion

Entwicklungen zur Verbesserung der Mensch-Maschine-Interaktion	
▪ Integration einer nutzeradaptiven Gesteninteraktion mit natürlichen Gesten im großen Arbeitsraum zur Steuerung des dynamischen Projektionssystems (vgl. Kapitel 3.3.3) ○ Auswahl geeigneter und intuitiver Gesten für die Verwendung im großen Arbeitsraum ○ Eigens erstellter Gestensatz für die Anforderungen der Großgeräte- und Baustellenmontage (Erkennung von oben) ○ Manuell gelabelter Datensatz mit 50 000 Gestenvariationen ○ Optimierter Prozess zur Gestendatensatzerstellung (Synthetisch erzeugter Datensatz) ○ Synthetisch erzeugter Gestendatensatz mit 15 000 Gestenvariationen ▪ Gestaltung der projizierten Informationen zur Anzeige der Interaktionsmöglichkeiten sowie Gestaltung des Interaktionsfeedbacks (vgl. Kapitel 3.3.2)	
Verbesserung zur Gestenerkennungszeit und Gestendatensatzerstellung	
▪ Verbesserte Gestenerkennungszeit mit durchschnittlich 2,6 Sekunden und einer Standardabweichung von 1,3 Sekunden im Experiment 3 ▪ Reduktion der Anzahl an benötigten Gestenbilder auf 15 000 Bilder ▪ Verbesserung des Zeitaufwandes zum Labeln von durchschnittlich 2 Sekunden pro Gestenbild (manuelles Labeln) auf 0,6 Sekunden (automatisiertes Labeln) ▪ Reduktion des Zeitaufwandes um 73 %, im Vergleich zum manuellen Labeln (15 000 Gesten)	

Aus dem ersten Experiment konnte festgestellt werden, dass eine Interaktion direkt in der Montage bereits erhebliches Potenzial bringt. Um eine Interaktion im großen Arbeitsraum zu ermöglichen, wurde auf Handgesten gesetzt, wobei intuitive Gesten ausgewählt wurden (vgl. Kapitel 2.4.2), welche von einem Kamerasystem von der Decke aus erkannt werden konnten. Dabei war die Herausforderung, dass die Handgesten in Relation zum Kameraabstand klein waren und demnach simple Gesten (offene Hand und Faust) verwendet werden musste. Im Zuge der ersten Demonstration mit Handgestendatensätzen wurde festgestellt, dass sich herkömmliche Gestendatensätze für die Anwendung in der Großgeräte- und Baustellenmontage aufgrund des großen Abstandes zur Kamera nicht eigneten. Demnach wurde ein eigener Gestendatensatz mit dem Setting der Fan-Cowl-Montage aus Experiment 2 erstellt und eigene Gesten manuell gelabelt und mit dem Algorithmus trainiert. Die erste Verwendung der Gesteninteraktion wurde für das Experiment 2 entwickelt und war mit einer durchschnittlichen Gestenzeit von rund 10 Sekunden und einer Stabilität der Erkennung von 90 % gut verwendbar, die Gestenerkennungszeit war jedoch noch ausbaufähig (vgl. Kapitel 5.2.2). Für die weitere Optimierung war es essenziell, die Erkennungszeit so zu gestalten, dass die Auslösung nicht unbeabsichtigt stattfindet und auch nicht zu lange benötigt, sodass sie sich nicht negativ auf die Ergonomie der Interaktion auswirkt. Die Erstellung des manuellen Datensatzes war aufwendig, weshalb nach Optimierungspotenzial in der Gestendatensatzerstellung gesucht wurde. Durch die Optimierung des Prozesses zur Gestendatensatzerstellung und Automatisierung des Labelns (vgl. Kapitel 3.3.3), konnte eine

Zeitersparnis von rund 73 %³¹ am Beispiel von 15 000 Gestenbildern erreicht werden. Um die Erkennungszeit und Stabilität der Erkennung zu verbessern, wurden die Trainingsdaten beim synthetisch erzeugten Datensatz optimiert, sodass für den Abstand der Kamera zur Erkennungsebene eine korrekte Größenskalierung der Gesten eingestellt wurde, welche eine Erkennungszeit von rund 2,6 Sekunden ermöglichte. Des Weiteren konnte dadurch die Anzahl der benötigten Gesten auf 15 000 Gestenbilder reduziert werden.

Durch die zusätzliche Projektion der benötigten Informationen zur Gesteninteraktion war dem Menschen stets klar, mit welcher Geste welche Funktion ausgeführt werden konnte. Nach erfolgreicher Ausführung der Geste wurde mit dem Projektionssystem ein geeignetes Gestenfeedback projiziert, um damit den aktuellen Status des Systems anzuzeigen (vgl. Kapitel 3.3.2 und Abbildung 61).

Weiteres Verbesserungspotenzial der Gesteninteraktion steckt vor allem im Bereich der Erweiterung der Interaktionsmöglichkeiten, der diversitätskonformen Datensatzgestaltung sowie der Verbesserung der Ergonomie der Gestenausführung. Demnach besteht der Bedarf, mit dem Projektionssystem anzuzeigen, wo überall Gesten im Arbeitsraum ausgeführt werden können. Damit kann die Erkennung auf einen kleineren Bereich fokussiert werden, was die Erkennungsgeschwindigkeit erhöhen und unabsichtliche Gesteninteraktionen vermeiden würde. Jedoch besteht dann nicht mehr die Möglichkeit, überall im Arbeitsraum bzw. an jeder Montageposition eine Interaktion durchzuführen.

Die gestaltete Gesteninteraktion mit natürlichen Gesten eignet sich besonders für die Anwendung im großen Arbeitsraum, für die direkte Interaktion an den jeweiligen Montagepositionen sowie für die Erkennung von intuitiven Gesten mit an der Decke montierten Kameras. Die Gesteninteraktion zur Steuerung des dynamischen Projektionssystems stellt damit gemeinsam mit der dynamischen Informationsbereitstellung die essenziellen Gestaltungsmerkmale dar, welche das Spatial Augmented Reality Assistenzsystem für die Großgeräte- und Baustellenmontage effektiv einsetzbar machen.

6.4 Limitationen der Ergebnisse und weiterer Forschungsbedarf

Das Spatial Augmented Reality Assistenzsystem wurde für die Anwendung der Großgeräte- und Baustellenmontage für die Use Cases und Demonstratoren entwickelt und relevantes Feedback der User wurde eingeholt. Trotz des iterativen Designprozesses und der Integration von verschiedenen Inputs aus User- und Prozesssicht stellen die Forschungsergebnisse aufgrund von räumlichen Gegebenheiten, definierten Hardwaresystemen sowie strategischer Ausrichtung der Entwicklung folgende Limitationen dar.

Informationsbereitstellung:

- Beschränkung der Deckenhöhe in der TU Wien-Pilotfabrik von 4 Metern und dementsprechende Einschränkung des effektiven Projektionsbereichs auf 8 x 3 Meter
- Manuelles Erstellen der Projektionsabfolge mit definierten Arbeitsanweisungen, Projektionspositionen und Verzerrung notwendig

³¹ vgl. Kapitel 3.3.3

- Bei automatisierter Einstellung der Verzerrung ist eine zusätzliche Kalibrierung des Projektors und der einzelnen Projektionspositionen notwendig.
- Beschattung der Projektionsfläche durch Montageobjekte oder Menschen möglich
- Keine interaktive Arbeitsanweisung bzw. interaktives projiziertes User-Interface

Aufgrund baulicher Beschränkungen und einer Deckenhöhe von 4 Metern ergibt sich ein maximaler Projektionsbereich von effektiv 8 x 3 Metern, welcher für die Anwendung der Use Cases ausreichend war. Das manuelle Erstellen der Projektionsabfolge, Definition der Projektionspositionen und zugehöriger Verzerrung ist zwar eine schnelle Möglichkeit, Arbeitsinformationen zu einem definierten Prozess zu erstellen, eine interaktivere Möglichkeit, wo direkt in der Montagestation die Positionen und Verzerrungen eingestellt werden können, wäre eine potenzielle Weiterentwicklung.

Auch besteht eine Limitation bei der automatisierten Verzerrung, bei der zwar ein Mehrwert durch die automatische Berechnung der Verzerrung bereits festgestellt werden konnte, jedoch die manuelle Kalibrierung des Projektors und der einzelnen Projektionspositionen einen zusätzlichen Aufwand darstellt. Die Einschränkung besteht aktuell auf statische und dynamische Arbeitsanweisungen, welche jedoch keine direkte Interaktion durch z. B. Touch auf einen Button ermöglichen (z. B. interaktive Buttons, interaktives User-Interface).

Gesteninteraktion:

- Gestenerkennung ist beschränkt auf ein Field of View (FOV) der Kamera (Intel RealSense) von 4 x 2 Metern.
- Beschränkt auf eine detaillierte Auflösung in Full-HD-Qualität (Intel RealSense)
- Beschränkung der Gesteninteraktion auf einen Menschen (1-zu-1-Interaktion)
- Beschränkung der Gesteninteraktionszeit von 2,6 Sekunden

Eine Limitation der Gesteninteraktion ist das beschränkte FOV von 4 x 2 Metern, welches jedoch bei den Experimenten kein Problem darstellte, da die Personen darüber informiert wurden wo eine Erkennung möglich ist. Um detailliertere Gesten zu integrieren, müsste ebenso eine Kamera mit höherer Auflösung verwendet werden. Erste Tests dazu wurden bereits mit der Kamera ZED-2 (Auflösung bis zu 2,2K) durchgeführt. Die Gestensteuerung ist für die Interaktion mit einem Menschen definiert, für mehrere Menschen müssten weitere Tests und Adaptionen der Gestaltung der Gesteninteraktion vorgenommen werden, damit nicht unbeabsichtigte Interaktionen ausgelöst werden. Des Weiteren ist die Erkennungszeit noch zu optimieren und wie oben vorgeschlagen ein definierter Bereich für die Gesteninteraktion zu evaluieren.

Aus der Analyse der Einschränkungen konnte ein zukünftiger Forschungsbedarf zur Weiterentwicklung des Spatial Augmented Reality Assistenzsystems ermittelt werden. Folgende Punkte geben einen Überblick über zukünftige Entwicklungen.

Weiterentwicklung des dynamischen Projektionssystems:

- Evaluierung eines Konzeptes zur Integration von Arbeitsanweisungen und Einstellung von Positionen und Verzerrungen direkt und interaktiv im Arbeitsraum zur Reduktion des Authoring-Aufwandes

- Weiterentwicklung der automatisierten Verzerrung in Bezug auf die Reduktion des Kalibrierungsaufwandes
- Integration und Evaluierung weiterer Gestaltung von Arbeitsanweisungen (z. B. In-situ-Montagevideos)
- Konzepte zur interaktiven Gestaltung der Arbeitsanweisungen und projizierte interaktive User-Interfaces
- Evaluierung einer prozessintegrierten Steuerung des Projektionssystems

Weiterentwicklung visuelles Interaktionssystem:

- Erweiterung auf ein größeres Field of View (FOV) und höherer Auflösung zur Vergrößerung der Interaktionsmöglichkeit und Verwendung von komplexeren Gesten
- Verbesserung des Gestendatensatzes durch Erweiterung anderer Gesten
- Evaluierung diversitätskonformer Gesten (Hautfarbe, Größe etc.)
- Verbesserung der Stabilität und Schnelligkeit der Gestenerkennung
- Tests mit analogen Objekterkennungsalgorithmen
- Erweiterung auf eine multimodale Interaktion (Gesten, Sprache, etc.)

Industrielle Anwendung in der Industrie:

- Anwendung des Spatial Augmented Reality Assistenzsystems in der Industrie
- (Langzeit-)Evaluierung in einem industriellen Montageprozess

Industrialisierung und Spin-off:

- Konzeption der Möglichkeiten zur Industrialisierung und Kommerzialisierung des Spatial Augmented Reality Systems für industrielle Anwendungen
- Ausgründung als Spin-off und kommerzielle Verwertung der Ergebnisse

Der zukünftige Forschungsbedarf stellt eine potenzielle Weiterentwicklung des Spatial Augmented Reality Assistenzsystems für Anwendungen im industriellen Umfeld dar und erweitert die Funktionen und Möglichkeiten der Informationsbereitstellung sowie Mensch-Maschine-Interaktion mit dem System.

6.5 Kritische Diskussion in Bezug auf die Forschungsfragen

Aus den dargestellten Ergebnissen und Erkenntnissen (vgl. Kapitel 6.1 bis 6.3) können die Forschungsfragen anschließend beantwortet werden.

Zentrale Forschungsfrage:

Welchen Mehrwert bietet ein Spatial Augmented Reality Assistenzsystem mit dynamischer In-situ-Projektion und nutzeradaptiver Gesteninteraktion zur Verbesserung der Informationsbereitstellung in der industriellen Großgeräte- und Baustellenmontage?

Unterfragen:

- 1) Wie kann das Spatial Augmented Reality Assistenzsystem für die Anforderungen der industriellen Großgeräte- und Baustellenmontage **gestaltet** und **demonstriert** werden?
- 2) Wie kann eine **dynamische In-situ-Projektion und Informationsführung** im Montageprozess umgesetzt werden?
- 3) Wie kann eine **nutzeradaptive Mensch-Maschine-Interaktion mit natürlichen Gesten** gestaltet werden?
- 4) Welchen **Mehrwert** bietet das Assistenzsystem für **Mitarbeiter*innen** (Usability, kognitive Arbeitsbelastung) und **Unternehmen** (Prozesszeit) im Vergleich zu bestehenden Informationssystemen in der Großgeräte- und Baustellenmontage?

Das Spatial Augmented Reality Assistenzsystem mit den Hauptkomponenten des dynamischen Projektionssystems und visuellen Interaktionssystems wurde laut den Anforderungen der industriellen Großgeräte- und Baustellenmontage mit der iterativen Design-Research-Methode (Alan R. Hevner et al., 2004) gestaltet und demonstriert (vgl. Kapitel 3 und 4). Dabei wurde die dynamische In-situ-Projektion und Informationsführung sowie eine nutzeradaptive Mensch-Maschine-Interaktion mit natürlichen Gesten ins Assistenzsystem integriert und der Mehrwert für Mensch und Unternehmen in drei praxisrelevanten Evaluierungsexperimenten ermittelt (vgl. Kapitel 5). Die Erkenntnisse wurden in Kapitel 6 diskutiert und Limitationen sowie der zukünftige Forschungsbedarf ermittelt.

Durch die Gestaltung, Demonstration und Evaluierung des „Spatial Augmented Reality Assistenzsystem mit dynamischer In-situ-Projektion und nutzeradaptiver Gesteninteraktion“ konnte die Informationsbereitstellung im Kontext der industriellen Großgeräte- und Baustellenmontage verbessert und somit die Zielsetzung der Dissertation (vgl. Kapitel 1.3) erreicht, die Forschungsfragen beantwortet und ein konkreter wissenschaftlicher und praxisrelevanter Beitrag geleistet werden.

7. Dissemination und Zusammenfassung

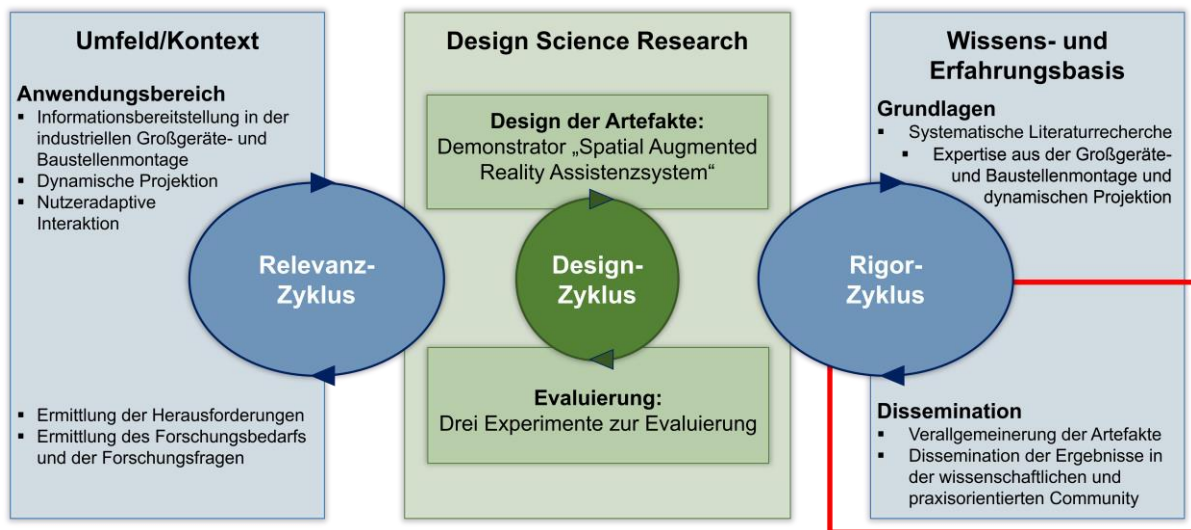


Abbildung 116 | Wissens- und Erfahrungsbasis – Dissemination, in Anlehnung an Alan R. Hevner et al. (2004)

Ein zentraler Punkt der Design-Science-Research-Methode (vgl. Abbildung 116) sowie der wirkungsvollen Forschungskommunikation/Dissemination ist, die gewonnenen Forschungsergebnisse und Erkenntnisse aus dem Design der Artefakte, der Demonstrationen sowie der Evaluierungen, der Forschungscommunity und der Community aus der Praxis zurückzuführen (Alan R. Hevner et al., 2004).

Die Forschungskommunikation wurde dazu mithilfe von Publikationen und Vorträgen in fach einschlägigen Konferenzen wie der CIRP Conference on Manufacturing Systems (CIRP CMS) (vgl. Rupprecht et al., 2020; Rupprecht et al., 2021), der International Conference on Production Research Manufacturing Innovation: Cyber Physical Manufacturing (ICPR) (vgl. Mayrhofer, Rupprecht & Schlund, 2019) sowie der International Conference on Intelligent Human Systems Integration (IHSI) (vgl. Rupprecht & Schlund, 2021) und Publikationen in Zeitschriften wie der „Zeitschrift für Arbeitswissenschaft (ZfA)“ (vgl. Schlund et al., 2018) veröffentlicht. Ebenso wurden einzelne Teilaspekte des Forschungsthemas in diversen studentischen Arbeiten (Bachelor und Diplomarbeiten) bearbeitet.

Die Ergebnisse und Erkenntnisse wurden mittels der Demonstratoren (vgl. Kapitel 4) in der TU Wien-Pilotfabrik der praxisorientierten Community präsentiert, welche dort begutachtet und getestet werden konnten. Das Spatial Augmented Reality Assistenzsystem wird in diversen Lehrveranstaltungen an verschiedenen Universitäten und Fachhochschulen als praxisrelevantes Beispiel für die Anwendung eines kognitiven Assistenzsystems in der Großgeräte- und Baustellenmontage verwendet und der Mehrwert für Unternehmen sowie für den Menschen präsentiert.

Zusammengefasst wurde ein geeignetes Spatial Augmented Reality Assistenzsystem für die Anwendung in der industriellen Großgeräte- und Baustellenmontage entwickelt, welches vor allem eine Verbesserung der Informationsbereitstellung und Interaktion und schließlich einen Mehrwert für Unternehmen und den Beschäftigten bringt. Durch die Experimente wurde vor allem relevantes Feedback für die Entwicklung von Assistenzsystemen eingeholt und in die Gestaltung integriert. Durch die Anwendung der Ergebnisse aus den Experimenten 1 und 2 konnte in der Verallgemeinerung der Ergebnisse in Experiment 3 gezeigt werden, dass das Assistenzsystem auf komplexere Montageprozesse und weitere Use Cases skalierbar ist. Das Assistenzsystem hat demnach hohes Potenzial für weitere Forschungsarbeiten und Entwicklungen sowie für die Anwendung auf weitere Montagestationen und Use Cases im industriellen Umfeld.

Durch die Adaptierung eines dynamischen Projektionssystems aus einem fachfremden Bereich und der Integration der dynamischen In-situ-Projektion und Informationsführung sowie einer nutzeradaptiven Gesteninteraktion konnte ein Spatial Augmented Reality Assistenzsystem für die Verbesserung der Informationsbereitstellung in der industriellen Großgeräte- und Baustellenmontage gestaltet werden, welches vor allem einen signifikanten Mehrwert für den Prozess und auch für die Ergonomie und Arbeitsbelastung des Menschen mit sich bringt.

Abbildungsverzeichnis

Abbildung 1 Kontext und Relevanz des Themas	4
Abbildung 2 Schematische Darstellung der Montageorganisationsart „Baustellenmontage“	5
Abbildung 3 Links: Montage einer Lokomotive, rechts: CFK-Layup-Prozess einer Flugzeugkomponente	5
Abbildung 4 Patentauszug des Projektionssystems mit der Patentnummer DE202013010365U1	7
Abbildung 5 Patentauszug des Projektionssystems mit der Patentnummer DE202013010365U1: Spiegel und Schrittmotoren	8
Abbildung 6 Forschungsfelder der vorliegenden Arbeit	10
Abbildung 7 Zyklen der Design-Science-Research-Methode	13
Abbildung 8 Design-Science-Research-Methode der Forschungsarbeit	15
Abbildung 9 Aufbau und Struktur der Forschungsarbeit	16
Abbildung 10 Wissens- und Erfahrungsbasis – Grundlagen	18
Abbildung 11 Kriterien der Systematischen Literaturrecherche	19
Abbildung 12 Detaillierte Vorgehensweise mit Inklusions- und Exklusionskriterien	19
Abbildung 13 Revolutionen von Operator 1.0 bis Operator 4.0	24
Abbildung 14 Einstellung von Systemen	27
Abbildung 15 Charakteristik der Assembly 4.0-Montagestation	31
Abbildung 16 Vision der nutzeradaptiven Gestaltung des Baustellenmontagearbeitsplatzes	32
Abbildung 17 Assistenzsysteme mit Fokus auf Informationsbereitstellung und Mensch-Maschine-Interaktion	33
Abbildung 18 Informationsbereitstellungsmöglichkeiten in der Großgeräte- und Baustellenmontage	34
Abbildung 19 Erstes Spatial Augmented Reality Assistenzsystem <i>Wellner's Digital Desk</i> ..	37
Abbildung 20 Laserprojektoren für 2D-Informationen	38
Abbildung 21 Einsatz von SAR bei ŠKODA AUTO	39
Abbildung 22 Statische Projektion in der Kleingerätemontage	39
Abbildung 23 Vergleich von Visualisierungsvarianten	43
Abbildung 24 Wahrnehmung von grüner Farbe bei unterschiedlicher Umgebungsfarbe (Simultankontrast)	44
Abbildung 25 Obere Reihe: Komplementärkontrast, untere Reihe: Hell-Dunkel-Kontrast...	45
Abbildung 26 TCT-Vergleich von verschiedenen Systemen	46
Abbildung 27 Vergleich und Evaluation der Informationsanzeige mit In-situ-Projektion, Tablet, HMD und Papierarbeitsanweisung	47
Abbildung 28 Maschinelles Lernen – Vorgehensweise	49
Abbildung 29 Aufbau eines neuronalen Netzwerkes	50
Abbildung 30 Schematischer Ablauf in einem neuronalen Netzwerk	51
Abbildung 31 Bilderkennungsprozess mit YOLO	52
Abbildung 32 Vergleich von YOLOv3 mit anderen Algorithmen	53

Abbildung 33 YOLOv4-Algorithmus	53
Abbildung 34 Mensch-Computer-Modell	55
Abbildung 35 Informationsübertragung in Mensch-Maschine-Systemen	55
Abbildung 36 Gestaltungsrichtlinien Mensch-Maschine-Interaktion	56
Abbildung 37 Überblick der Gesteninteraktion	58
Abbildung 38 Geste „Faust“ inklusive Objekterkennung	59
Abbildung 39 Design der Artefakte	64
Abbildung 40 Erste Konzeptskizze eines Assistenzsystems in der Großgeräte- und Baustellenmontage	65
Abbildung 41 Konzeptidee eines Spatial Augmented Reality Systems mit dynamischer Projektion und nutzeradaptiver Interaktion	66
Abbildung 42 Design- und Demonstrationselemente des Spatial Augmented Reality Assistenzsystems	67
Abbildung 43 Aufbau des Assistenzsystems	68
Abbildung 44 Projektor Panasonic PT-RZ660BE	69
Abbildung 45 Projektorenanordnung für den Winglet-Demonstrator	70
Abbildung 46 Mirror-Head zur dynamischen Projektion	71
Abbildung 47 Mirror-Head Typ MH11	72
Abbildung 48 Drehung und Neigung des Spiegelsystems	73
Abbildung 49 Vernetzung des Spiegelsystems zum Medienserver und Projektor	73
Abbildung 50 Links: User-Interface-Authoring- und Medienbearbeitungssoftware, rechts: PC zur Medien- und Projektorsteuerung	74
Abbildung 51 Vorgehensweise zur Erstellung von dynamischen Projektionen	75
Abbildung 52 Oben: Digitale Positionssteuerung des Spiegelsystems, unten: digitale Anpassungsmöglichkeit (Skalierung, Drehung, Verzerrung) der integrierten Daten	76
Abbildung 53 Oben: Projektor mit Stereolabs ZED-2-Kamera, unten links: Schachbrettmuster und projizierte Punkte zur Kalibrierung, unten rechts: Schematische Darstellung der Projektionsfläche und Stereokamera	78
Abbildung 54 Beispiel der automatisierten Verzerrung	78
Abbildung 55 Rendering „Intel RealSense – D435i“	80
Abbildung 56 Bilderkennung und Steuerung der dynamischen Projektion	81
Abbildung 57 Setting Baustellenmontage – Gestenerkennung von oben	82
Abbildung 58 Gestaltungsaspekte der dynamischen Informationsbereitstellung und Mensch-Maschine-Interaktion	84
Abbildung 59 Erste Arbeitsanweisung mit „Interaktive Buttons“. Links: Anzeige der Montagetätigkeiten, rechts: Anzeige der Montageposition mittels Symbolik	85
Abbildung 60 Komplexere Arbeitsanweisung mit Information über die Gesteninteraktion ..	86
Abbildung 61 Feedbackinformation und Bestätigung der erfolgreichen Task-/Gestenausführung	87
Abbildung 62 Verbesserung der intuitiven Gestaltung der Gesteninformation	88
Abbildung 63 Verringerung der Informationsdetails mit zeitlichem Montagefortschritt	88

Abbildung 64 In-view-Projektion mit Projektionssystem und In-situ-Projektion mit statischem Lasersystem	90
Abbildung 65 In-situ-Projektion mit farbigen Lichtpunkten.....	91
Abbildung 66 In-situ-Projektion mit dynamischer Informationsführung zur Montageposition	92
Abbildung 67 Anzeige der Montageposition und -orientierung	92
Abbildung 68 Experiment 3: In-situ-Projektion mit dynamischer Informationsführung.....	93
Abbildung 69 Links: Dynamische Animation zur Anzeige der Entnahmeposition.....	94
Abbildung 70 Links: Geste „offene Hand“, rechts: Geste „Faust“	95
Abbildung 71 Vorklassifizierung der Objektklasse „Mensch von oben“ mit den Labels	96
Abbildung 72 Trainingsergebnisse YOLO	97
Abbildung 73 Links: Aufnahme der Gesten vor neutralem Hintergrund, rechts: entfernter Hintergrund mittels GrabCut.....	98
Abbildung 74 Automatisierte Integration der Gesten in die Szene	98
Abbildung 75 Spatial Augmented Reality Assistenzsystem in der TU Wien-Pilotfabrik.....	100
Abbildung 76 Demonstratoren und Use Cases der Forschungsarbeit	101
Abbildung 77 Demonstrator „Winglet-Montagestation“	101
Abbildung 78 Fan-Cowl-Montagestation in der TU Wien-Pilotfabrik.....	102
Abbildung 79 Spielzeugbaustein-Montage	102
Abbildung 80 Full-HD+ Projektor mit digital steuerbarem Spiegelsystem	103
Abbildung 81 Praxisbeispiele von Projektionen auf verschiedenen Oberflächen und Materialien	106
Abbildung 82 Projektionsgerechte Aufbereitung der Arbeitsinformationen	108
Abbildung 83 Vergleich von Gesteninformationen: textlastige Information (links), Verwendung von Piktogrammen (rechts)	109
Abbildung 84 Demonstration von Arbeitsinformationen mit weißem und transparentem Hintergrund	109
Abbildung 85 Farbendemonstration in der Baustellenmontage	110
Abbildung 86 Vergleich der Farben Gold-Orange und Gelb in der Software des Projektors und am Demonstrator	111
Abbildung 87 Finale Farbgestaltung	111
Abbildung 88 In-view-Projektion in der Fan-Cowl-Montage.....	112
Abbildung 89 In-situ-Projektion von Arbeitsanweisungen in der Fan-Cowl-Montage	113
Abbildung 90 Dynamische Informationsführung und In-situ-Projektion	114
Abbildung 91 Anzeige der Montageposition durch Lichtpunkte und Lichtvorhang bei erfolgreicher Durchführung	114
Abbildung 92 In-situ-Projektion und Informationsführung zur Entnahme- und Montageposition bei der Spielbaustein-Montage.....	115
Abbildung 93 Mensch-, Positions- und Gestenerkennung.....	116
Abbildung 94 Veranschaulichung: Projektion folgt dem Menschen	117
Abbildung 95 Interaktion mittels Handgesten inklusive Feedbackinformation	118
Abbildung 96 Gesteninteraktion Spielzeugbaustein-Demonstrator	119

Abbildung 97 Evaluierung mittels drei Experimenten	120
Abbildung 98 Akzeptanzbereiche, Notenskala und Adjektivbewertungen von SUS- Ergebnissen.....	121
Abbildung 99 Video-Annotation-Tool „ANVIL2”	125
Abbildung 100 Fehlerliste des Montageprozesses	127
Abbildung 101 Evaluierungsexperimente und Use Cases.....	128
Abbildung 102 Experiment 1: Dynamische Projektion vs. PC-Terminal	129
Abbildung 103 Prozessablauf von Experiment 1	130
Abbildung 104 Experiment 2: In-view- und In-situ-Projektion	133
Abbildung 105 Prozessablauf von Experiment 2	134
Abbildung 106 NASA-RTLX – Experiment 2	135
Abbildung 107 SUS Score – Experiment 2	136
Abbildung 108 GATM – Experiment 2	138
Abbildung 109 Experiment 3: Spielzeugbaustein-Montage, dynamische In-situ-Projektion mit Informationsführung zur Entnahme- und Montageposition	141
Abbildung 110 Prozessablauf von Experiment 3	142
Abbildung 111 In-situ-Projektion und Informationsführung im Experiment 3	143
Abbildung 112 Arbeitsanweisung via Tablet	143
Abbildung 113 Ergebnisse „Kognitive Arbeitsbelastung“ – Experiment 3.....	144
Abbildung 114 Ergebnisse „Usability“ – Experiment 3.....	145
Abbildung 115 Ergebnisse der „Task Completion Time“ – Experiment 3	147
Abbildung 116 Wissens- und Erfahrungsbasis – Dissemination	164

Tabellenverzeichnis

Tabelle 1 Fokus und Zielsetzung der Forschungsarbeit	11
Tabelle 2 Umsetzung der Guidelines der Design-Research-Methode	14
Tabelle 3 Kombination der Schlüsselwörter zu konkreten Suchbegriffen.....	20
Tabelle 4 Ergebnisse der Systematischen Literaturrecherche	21
Tabelle 5 State-of-the-Art-Analyse – Kategorisierung der Literatur	21
Tabelle 6 Taxonomie der nutzeradaptiven Gestaltung	29
Tabelle 7 Vor- und Nachteile von Gesteninteraktion	62
Tabelle 8 Zusammenfassung der relevanten Elemente der Literaturrecherche	63
Tabelle 9 Spezifikationen des Projektors Panasonic PT-RZ660BE	70
Tabelle 10 Spezifikationen Spiegelsystem	71
Tabelle 11 Arbeitsaufwand zur Erstellung einer dynamischen Projektion	77
Tabelle 12 Spezifikationen der Stereokamera „ZED 2“	77
Tabelle 13 Spezifikationen der Kamera „Intel RealSense Typ D435i“	80
Tabelle 14 Komponenten, Preise und Gesamtanschaffungskosten des SAR-Systems	83
Tabelle 15 Vergleich Arbeitsaufwand der Gestendatensatzerstellung	99
Tabelle 16 Herausforderungen der Projektion und Gegenmaßnahmen	104
Tabelle 17 Erkenntnisse und Vorschläge für projektionsgerechte Medienaufbereitung	107
Tabelle 18 TAM-Fragebogen	122
Tabelle 19 NASA-RTLX zur Bewertung von Assistenzsystemen	123
Tabelle 20 Phasendefinition GATM für die Baustellenmontage	124
Tabelle 21 Interpretation der TCTs	126
Tabelle 22 Evaluierungsergebnisse von Experiment 1	131
Tabelle 23 Kernaussagen von Experiment 1	131
Tabelle 24 Beschreibung der GATM-Phasen – Experiment 2	137
Tabelle 25 Qualitative Ergebnisse – Experiment 2	139
Tabelle 26 GATM-Phasen und Gesteninteraktion – Experiment 3	145
Tabelle 27 Qualitative Ergebnisse – Experiment 3	148
Tabelle 28 Zusammenfassung der Ergebnisse aus Experiment 1 und 2	150
Tabelle 29 Ergebnisse und Erkenntnisse aus der finalen Evaluierung (Experiment 3)	152
Tabelle 30 Ergebnisse aus dem Design der Artefakte	153
Tabelle 31 Berechnung der Kostenoptimierung und des Return on Investments	155
Tabelle 32 Erkenntnisse der Gestaltung der dynamischen Informationsbereitstellung	156
Tabelle 33 Erkenntnisse der Gestaltung der Gersteninteraktion	159

Literaturverzeichnis

- Abraham, M. & Annunziata, M. (2017). *Augmented Reality is already Improving Worker Performance*<https://hbr.org/2017/03/augmented-reality-is-already-improvingworker-performance>
- acatech. (2011). *Cyber-Physical Systems: Driving force for innovation in mobility, health, energy and production. acatech POSITION PAPER, December 2011*. Springer.
<http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10650532>
<https://doi.org/10.1007/978-3-642-29090-9>
- acatech. (2019). *Themenfelder Industrie 4.0: Forschungs- und Entwicklungsbedarfe zur erfolgreichen Umsetzung von Industrie 4.0*.
- Aerospace Manufacturing Magazine - MIT Publishing. (2018). *The construction of light*.
<https://www.aero-mag.com/the-construction-of-light>
- AIM Agile IT Management GmbH. *Human in the loop und Machine Learning*.
<https://www.agile-im.de/2019/08/14/human-in-the-loop/>
- ARENA2036 e.V. (2020). *ARENA 2036*. <https://www.arena2036.de/de/>
- Assembly Solutions GmbH. (2020). *Assembly Solutions*. <https://www.assembliesolutions.de/>
- Atzori, L., Iera, A. & Morabito, G. (2010). The Internet of Things: A survey. *Computer Networks*, 54(15), 2787–2805. <https://doi.org/10.1016/j.comnet.2010.05.010>
- Aviation Media & IT GmbH. (2021). *Fertigung eines 787 Schubumkehrgehäuses (Cleanroom) im neuen Werk 4 der FACC AG*.
https://www.aero.de/content/pics/p_1988.jpg
- Ayodele, T. O. (2010). Machine Learning Overview. In Y. Zhang (Hrsg.), *New Advances in Machine Learning*. InTech. <https://doi.org/10.5772/9374>
- Azuma, R. T. (1997). A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments*, 6(4), 355–385. <https://doi.org/10.1162/pres.1997.6.4.355>
- Bangor, A., Kortum, P. & Miller, J. (2009). Determining What Individual SUS Scores Mean: Adding an Adjective Rating Scale. *Journal of Usability Studies*, 4 (3), 114–123.
- Baudel, T. & Beaudouin-Lafon, M. (1993). Charade. *Communications of the ACM*, 36(7), 28–35. <https://doi.org/10.1145/159544.159562>
- Bertram, P., Birtel, M., Quint, F. & Ruskowski, M. (2018). Intelligent Manual Working Station through Assistive Systems. *IFAC-PapersOnLine*, 51(11), 170–175.
<https://doi.org/10.1016/j.ifacol.2018.08.253>
- Bibliographisches Institut GmbH. (2021). *Geste*.
<https://www.duden.de/rechtschreibung/Geste>
- Blattgerste, J., Strenge, B., Renner, P., Pfeiffer, T. & Essig, K. (2017). Comparing Conventional and Augmented Reality Instructions for Manual Assembly Tasks. In Unknown (Hrsg.), *Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments - PETRA '17* (S. 75–82). ACM Press. <https://doi.org/10.1145/3056540.3056547>
- Bochkovskiy, A., Wang, C.-Y. & Liao, H.-Y. M. (2020, 23. April). *YOLOv4: Optimal Speed and Accuracy of Object Detection*. <https://arxiv.org/pdf/2004.10934>

- Bokranz, R. & Landau, K. (2012). *Handbuch Industrial Engineering: Produktivitätsmanagement mit MTM* (2., überarb. und erw. Aufl.). Schäffer-Poeschel.
- Bornmann, J., Kurzweg, A. & Heinrich, K. (2015). Tragbare Assistenzsysteme in der Automobilmontage - Forschung und Entwicklung innovativer orthetischer Systeme zur physischen Unterstützung während der Überkopfarbeit. In R. Weidner, T. Redlich & J. P. Wulfsberg (Hrsg.), *Technische Unterstützungssysteme* (Zweite transdisziplinäre Konferenz: Konferenzband, S. 507–516). Springer Vieweg.
- Bortolini, M., Accorsi, R., Faccio, M., Galizia, F. G [Francesco Gabriele] & Pilati, F. (2019). Toward a Real-Time Reconfiguration of Self-Adaptive Smart Assembly Systems. *Procedia Manufacturing*, 39, 90–97. <https://doi.org/10.1016/j.promfg.2020.01.232>
- Bortolini, M., Faccio, M., Gamberi, M. & Pilati, F. (2020). Motion Analysis System (MAS) for production and ergonomics assessment in the manufacturing processes. *Computers & Industrial Engineering*, 139, 105485. <https://doi.org/10.1016/j.cie.2018.10.046>
- Bortolini, M., Ferrari, E., Gamberi, M., Pilati, F. & Faccio, M. (2017). Assembly system design in the Industry 4.0 era: a general framework. *IFAC-PapersOnLine*, 50(1), 5700–5705. <https://doi.org/10.1016/j.ifacol.2017.08.1121>
- Brinzer, B., Banerjee, A. & Hauth, M. (2017). Complexity Thinking and Cyber-Physical Systems. *International Journal of Industrial Engineering*, 4(1), 14–20. <https://doi.org/10.14445/23499362/IJIE-V4I1P103>
- Brooke, J. (1996). SUS: A quick and dirty usability scale. In Jordan, P, W, B. Thomas, McClelland, I, L & B. Weerdmeester (Hrsg.), *Usability evaluation in industry*. Taylor & Francis.
- Bühler, P., Schlaich, P. & Sinner, D. (2017). *Zeichen und Grafik*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-53850-0>
- Bullinger, H.-J. (1994). *Ergonomie: Produkt- und Arbeitsplatzgestaltung. Technologiemanagement*. Teubner.
- Bullinger-Hoffmann, A. C. & Mühlstedt, J. (2016). *Homo Sapiens Digitalis - Virtuelle Ergonomie und digitale Menschmodelle*. Springer Berlin Heidelberg. <https://ebookcentral.proquest.com/lib/gbv/detail.action?docID=4832097>
- Bundesministerium der Finanzen. (2021). *AfA-Tabellen*. https://www.bundesfinanzministerium.de/Content/DE/Standardartikel/Themen/Steuer n/Weitere_Steuerthemen/Betriebspruefung/AfA-Tabellen/Ergaenzende-AfA-Tabellen/AfA-Tabelle_AV.pdf?__blob=publicationFile&v=3
- Büttner, S., Funk, M., Sand, O. & Röcker, C. (2016). Using Head-Mounted Displays and In-Situ Projection for Assistive Systems. In Unknown (Hrsg.), *Proceedings of the 9th ACM International Conference on Pervasive Technologies Related to Assistive Environments - PETRA '16* (S. 1–8). ACM Press. <https://doi.org/10.1145/2910674.2910679>
- Büttner, S., Mucha, H., Funk, M., Kosch, T., Aehnelt, M., Robert, S. & Röcker, C. (2017). The Design Space of Augmented and Virtual Reality Applications for Assistive Environments in Manufacturing. In Unknown (Hrsg.), *Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive*

- Environments - PETRA '17* (S. 433–440). ACM Press.
<https://doi.org/10.1145/3056540.3076193>
- Buxton, W., Billingham, M., Guiard, Y., Sellen, A., Zhai, S. (2011). *Gesture Based Interaction*. In *Human Input to Computer Systems: Theories, Techniques and Technology*. <http://www.billbuxton.com/inputManuscript.html>
- Cadoz, C. (1994). Le geste, canal de communication instrumental. *Techniques et sciences informatiques*. Cachan: Lavoisier, Vol 13.
- Calefato, C., Montanari, R. & Tesauri, F. (2008). The Adaptive Automation Design. In K. Asai (Hrsg.), *Human Computer Interaction: New Developments*. InTech.
<https://doi.org/10.5772/5878>
- Card, S. K., Moran, T. P. & Newell, A. (1983). *The psychology of human-computer interaction*. Erlbaum.
- Coletti, P. & Aichner, T. (2011). *Mass Customization: An Exploration of European Characteristics*. SpringerBriefs in Business. Paolo Coletti.
<http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10464257>
<https://doi.org/10.1007/978-3-642-18390-4>
- Davis, F. D., Bagozzi, R. P. & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science*, 35(8), 982–1003. <https://doi.org/10.1287/mnsc.35.8.982>
- Dhiman, H., Büttner, S., Röcker, C. & Reisch, R. (2019). Handling Work Complexity with AR/Deep Learning. In *Proceedings of the 31st Australian Conference on Human-Computer-Interaction* (S. 518–522). ACM. <https://doi.org/10.1145/3369457.3370919>
- DIN EN ISO 6385:2016-12. *DIN EN ISO 6385:2016-12, Grundsätze der Ergonomie für die Gestaltung von Arbeitssystemen (ISO_6385:2016); Deutsche Fassung EN_ISO_6385:2016*. Berlin. Beuth Verlag GmbH.
- DIN EN ISO 9241-11:2018-11. *DIN EN ISO 9241-11:2018-11, Ergonomie der Mensch-System-Interaktion_ - Teil_11: Gebrauchstauglichkeit: Begriffe und Konzepte (ISO_9241-11:2018); Deutsche Fassung EN_ISO_9241-11:2018*. Berlin. Beuth Verlag GmbH.
- DIN EN ISO 9241-110:2020. *DIN EN ISO 9241-110:2020-10, Ergonomie der Mensch-System-Interaktion_ - Teil_110: Interaktionsprinzipien (ISO_9241-110:2020); Deutsche Fassung EN_ISO_9241-110:2020*. Berlin. Beuth Verlag GmbH.
- DIN EN ISO 9241-112. *DIN EN ISO 9241-112:2017-08, Ergonomie der Mensch-System-Interaktion_ - Teil_112: Grundsätze der Informationsdarstellung (ISO_9241-112:2017); Deutsche Fassung EN_ISO_9241-112:2017*. Berlin. Beuth Verlag GmbH.
- DIN EN ISO 9241-125. *DIN EN ISO 9241-125:2018-05, Ergonomie der Mensch-System-Interaktion_ - Teil_125: Empfehlungen zur visuellen Informationsdarstellung (ISO_9241-125:2017); Deutsche Fassung EN_ISO_9241-125:2017*. Berlin. Beuth Verlag GmbH.
- DIN EN ISO 9241-171. *DIN EN ISO 9241-171 - Ergonomie der Mensch-System-Interaktion - Teil 171_Leitlinien für die Zugänglichkeit von Software*.

- DIN EN ISO 9241-210. *DIN EN ISO 9241-210:2019_Ergonomie der Mensch-System-Interaktion_ - Teil_210: Prozess zur Gestaltung gebrauchstauglicher interaktiver Systeme (ISO_9241-210:2010); Deutsche Fassung EN_ISO_9241-210:2010*. Berlin. Beuth Verlag GmbH.
- DIN EN ISO 9241-960:2018. *DIN EN ISO 9241-960:2018-01, Ergonomie der Mensch-System-Interaktion- Teil_960: Rahmen und Anleitung zur Gestensteuerung (ISO_9241-960:2017); Deutsche Fassung EN_ISO_9241-960:2017*. Berlin. Beuth Verlag GmbH.
- Dombrowski, U., Wagner, T. & Riechel, C. (2013). Concept for a Cyber Physical Assembly System. In M. Y. Wang (Hrsg.), *2013 IEEE International Symposium on Assembly and Manufacturing (ISAM): July 30, 2013 - Aug. 2, 2013, Xi'an, China* (S. 293–296). IEEE. <https://doi.org/10.1109/ISAM.2013.6643461>
- Dynamic Projection Institut GmbH. (2020a). *Dynamic Projection*. <https://www.dynamicprojection.com>
- Dynamic Projection Institut GmbH. (2020b). *MDC-X Complete Guide and Manual*. <https://www.dynamicprojection.com/transfer/MDC-Linux/MDC-X-Manual-Complete.pdf>
- Dynamic Projection Institut GmbH. (2020c). *Mirror Head Quick Installation Guide*. https://www.dynamicprojection.com/transfer/MH-manuals/MHs/MH11/DPI-QIG-Mirror_Head-MH11.pdf
- Dynamic Projection Institut GmbH. (2020d). *Tech-Sheet MH11*. https://www.dynamicprojection.com/transfer/MH-manuals/MHs/MH11/DPI-TecSheets-Mirror_Head-MH11.pdf
- Dynamic Projection Institut GmbH. (2020e). *Wandhalterung Mirror-Head*. https://www.dynamicprojection.com/transfer/MH-manuals/MHs/MH11/DPI-Accessories-Mirror_Head-MH11.pdf
- Dynamic Projection Institute, Herstellungs- und Vertriebs GmbH. (2015). Vorrichtung zum Projizieren eines Bilds(DE202013010365U1). <https://worldwide.espacenet.com/patent/search/family/052673459/publication/DE202013010365U1?q=DE202013010365U1>
- Eckert, C., Isaksson, O., Hallstedt, S., Malmqvist, J., Öhrwall Rönnbäck, A. & Panarotto, M. (2019). Industry Trends to 2040. *Proceedings of the Design Society: International Conference on Engineering Design*, 1(1), 2121–2128. <https://doi.org/10.1017/dsi.2019.218>
- Ernst, H., Schmidt, J. & Beneken, G. (2020). Maschinelles Lernen: Deep Learning mit neuronalen Netzen. In H. Ernst, J. Schmidt & G. Beneken (Hrsg.), *Grundkurs Informatik* (S. 801–833). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-30331-0_18
- Faccio, M., Ferrari, E., Galizia, F. G [Francesco G.], Gamberi, M. & Pilati, F. (2019). Real-time assistance to manual assembly through depth camera and visual feedback. *Procedia CIRP*, 81, 1254–1259. <https://doi.org/10.1016/j.procir.2019.03.303>

- Fantini, P., Leitao, P., Barbosa, J. & Taisch, M. (2019). Symbiotic Integration of Human Activities in Cyber-Physical Systems. *IFAC-PapersOnLine*, 52(19), 133–138. <https://doi.org/10.1016/j.ifacol.2019.12.124>
- FARO Europe GmbH & Co. KG. (2020a). *Faro - Laserprojektor*. <https://www.faro.com/de-de/produkte/3d-manufacturing/faro-laser-projector/>
- FARO Europe GmbH & Co. KG. (2020b). *FARO® Laser Projector*. <https://www.faro.com/de-de/produkte/3d-manufacturing/faro-laser-projector/>
- Fikkert, W. F. (2010). *Gesture interaction at a distance*. <https://doi.org/10.3990/1.9789036529730>
- Fletcher, S. R., Johnson, T., Adlon, T., Larreina, J., Casla, P., Parigot, L., Alfaro, P. J. & Otero, M. d. M. (2020). Adaptive automation assembly: Identifying system requirements for technical efficiency and worker satisfaction. *Computers & Industrial Engineering*, 139, 105772. <https://doi.org/10.1016/j.cie.2019.03.036>
- Fraunhofer IPA. (2018). *MonSiKo: Ein adaptives Assistenzsystem vereinfacht die Montage der Industrie 4.0*. <https://www.ipa.fraunhofer.de/de/referenzprojekte/MonSiKo.html>
- Fraunhofer-Institut für Arbeitswirtschaft und Organisation IAO. (2020). *Immersive Engineering Lab - Fraunhofer IAO*. <https://www.iao.fraunhofer.de/lang-de/labors-ausstattung/976-immersive-participation-lab.html>
- Funk, M. (2016). *Augmented reality at the workplace : a context-aware assistive system using in-situ projection*. <https://doi.org/10.18419/OPUS-8997>
- Funk, M., Bächler, A., Bächler, L., Korn, O., Krieger, C., Heidenreich, T. & Schmidt, A. (2015). Comparing projected in-situ feedback at the manual assembly workplace with impaired workers. In F. Makedon (Hrsg.), *Proceedings of the 8th ACM International Conference on Pervasive Technologies Related to Assistive Environments - PETRA '15* (S. 1–8). ACM Press. <https://doi.org/10.1145/2769493.2769496>
- Funk, M., Bächler, A., Bächler, L., Kosch, T., Heidenreich, T. & Schmidt, A. (2017). Working with Augmented Reality? A Long-Term Analysis of In-Situ Instructions at the Assembly Workplace. In Unknown (Hrsg.), *Proceedings of the 10th International Conference on Pervasive Technologies Related to Assistive Environments - PETRA '17* (S. 222–229). ACM Press. <https://doi.org/10.1145/3056540.3056548>
- Funk, M., Kosch, T., Greenwald, S. W. & Schmidt, A. (2015). A benchmark for interactive augmented reality instructions for assembly tasks. In C. Holzmann & R. Mayrhofer (Hrsg.), *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia* (S. 253–257). ACM. <https://doi.org/10.1145/2836041.2836067>
- Funk, M., Kosch, T. & Schmidt, A. (2016). Interactive worker assistance: comparing the effects of in-situ projection, head-mounted displays, tablet, and paper instructions. In P. Lukowicz, A. Krüger, A. Bulling, Y.-K. Lim & S. N. Patel (Hrsg.), *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing* (S. 934–939). ACM. <https://doi.org/10.1145/2971648.2971706>
- Funk, M., Kosch, T., Wolf, K., Knierim, P., Mayer, S. & Schmidt, A. (2016). Automatic projection positioning based on surface suitability. In J. Müller, N. Memarovic, T. Ojala & V. Kostakos (Hrsg.), *Proceedings of the 5th ACM International Symposium on*

- Pervasive Displays - PerDis '16* (S. 75–79). ACM Press.
<https://doi.org/10.1145/2914920.2915014>
- Funk, M., Mayer, S., Nistor, M. & Schmidt, A. (2016). Mobile In-Situ Pick-by-Vision: Order Picking Support using a Projector Helmet. In Unknown (Hrsg.), *Proceedings of the 9th ACM International Conference on Pervasive Technologies Related to Assistive Environments - PETRA '16* (S. 1–4). ACM Press.
<https://doi.org/10.1145/2910674.2910730>
- Gaham, M., Bouzouia, B. & Achour, N. (2015). Human-in-the-Loop Cyber-Physical Production Systems Control (HiLCP2sC): A Multi-objective Interactive Framework Proposal. In T. Borangiu (Hrsg.), *Studies in Computational Intelligence: Bd. 594. Service orientation in holonic and multi-agent manufacturing* (Bd. 594, S. 315–325). Springer. https://doi.org/10.1007/978-3-319-15159-5_29
- Galaske, N. & Anderl, R. (2016). Approach for the Development of an Adaptive Worker Assistance System Based on an Individualized Profile Data Model. In C. Schlick & S. Trzcieliński (Hrsg.), *Advances in Intelligent Systems and Computing: Bd. 490. Advances in Ergonomics of Manufacturing: Managing the Enterprise of the Future: Proceedings of the AHFE 2016 International Conference on Human Aspects of Advanced Manufacturing, July 27-31, 2016, Walt Disney World®, Florida, USA* (Bd. 490, S. 543–556). Springer International Publishing. https://doi.org/10.1007/978-3-319-41697-7_47
- Garcia, M. A. R., Rojas, R., Gualtieri, L., Rauch, E. & Matt, D. (2019). A human-in-the-loop cyber-physical system for collaborative assembly in smart manufacturing. *Procedia CIRP*, 81, 600–605. <https://doi.org/10.1016/j.procir.2019.03.162>
- Geiselhart, F., Otto, M. & Rukzio, E. (2016). On the Use of Multi-Depth-Camera Based Motion Tracking Systems in Production Planning Environments. *Procedia CIRP*, 41, 759–764. <https://doi.org/10.1016/j.procir.2015.12.088>
- Gesellschaft für Informatik e.V. (2020). *Definition Mensch-Maschine-Interaktion*. <https://fb-mci.gi.de/fachbereich/was-ist-mci>
- Gorecky, D., Campos, R., Chakravarthy, H., Dabelow, R., Schlick, J. & Zühlke, D. (2013). Mastering Mass Customization - A Concept for Advanced, Human- Centered Assembly. *ACADEMIC JOURNAL of MANUFACTURING ENGINEERING*, 11 (2), 62–67.
- Gorecky, D., Schmitt, M., Loskyll, M. & Zühlke, D. (2014). Human-machine-interaction in the industry 4.0 era. In C. E. Pereira (Hrsg.), *2014 12th IEEE International Conference on Industrial Informatics (INDIN 2014): Porto Alegre - RS, Brazil, 27 - 30 July 2014* (S. 289–294). IEEE. <https://doi.org/10.1109/INDIN.2014.6945523>
- Günthner, W. A., Blomeyer, N., Reif, R. & Schedlbauer, M. (2009). *Pick-by-Vision: Augmented Reality unterstützte Kommissionierung*. Lehrstuhl für Fördertechnik Materialfluß Logistik (fml) Techn. Univ. München.
- Hancock, P. A [P. A.] & Chignell, M. H. (1987). Adaptive Control in Human-Machine Systems. In *Advances in Psychology. Human Factors Psychology* (Bd. 47, S. 305–345). Elsevier. [https://doi.org/10.1016/S0166-4115\(08\)62312-2](https://doi.org/10.1016/S0166-4115(08)62312-2)

- Hancock, P. A [Peter A.], Jagacinski, R. J., Parasuraman, R., Wickens, C. D., Wilson, G. F. & Kaber, D. B. (2013). Human-Automation Interaction Research. *Ergonomics in Design: The Quarterly of Human Factors Applications*, 21(2), 9–14.
<https://doi.org/10.1177/1064804613477099>
- Hart, S. G. (2006). Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 50(9), 904–908.
<https://doi.org/10.1177/154193120605000909>
- Hart, S. G. & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology. Human Mental Workload* (Bd. 52, S. 139–183). Elsevier. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- He, K., Zhang, X., Ren, S. & Sun, J. (2016). Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (S. 770–778). IEEE. <https://doi.org/10.1109/CVPR.2016.90>
- Heinz, M., Dhiman, H. & Röcker, C. (2018). A Multi-device Assistive System for Industrial Maintenance Operations. In A. Holzinger, P. Kieseberg, A. M. Tjoa & E. Weippl (Hrsg.), *Lecture Notes in Computer Science. Machine Learning and Knowledge Extraction* (Bd. 11015, S. 239–247). Springer International Publishing.
https://doi.org/10.1007/978-3-319-99740-7_16
- Hermann, M., Pentek, T. & Otto, B. (2016). Design Principles for Industrie 4.0 Scenarios. In *2016 49th Hawaii International Conference on System Sciences (HICSS)* (S. 3928–3937). IEEE. <https://doi.org/10.1109/HICSS.2016.488>
- Hevner, A. R [A. R.] & March, S. T [S. T.] (2003). IT systems perspectives - the information systems research cycle. *Computer*, 36(11), 111–113.
<https://doi.org/10.1109/MC.2003.1244541>
- Hevner, A. R [Alan R.] (2007). A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*, 19 (2), 87–92.
- Hevner, A. R [Alan R.], March, S. T [Salvatore T.], Park, J. & Ram, S. (2004). Design Science in Information Systems Research. *MIS Quarterly Vol. 28 No. 1*, 75–105.
- Hořejší, P. (2015). Augmented Reality System for Virtual Training of Parts Assembly. *Procedia Engineering*, 100, 699–706. <https://doi.org/10.1016/j.proeng.2015.01.422>
- Hornacek, M., Küffner-McCauley, H., Trimmel, M., Rupprecht, P. & Schlund, S. (2021). A Spatial AR System for Wide-area Axis-aligned Metric Augmentation of Planar Scenes (in review). *CIRP Journal of Manufacturing Science and Technology*.
- IFR International Federation of Robotics. (2019). *World Robotics 2019 Industrial Robots*.
<https://ifr.org>
- Institut für Managementwissenschaften. (2020). *TU Wien - Mensch-Maschine-Interaktion*.
<https://www.imw.tuwien.ac.at/cps/home/>
- Intel Corporation. (2020). *Intel Realsense Camers D435*.
https://www.intelrealsense.com/depth-camera-d435/?utm_source=intelcom_website&utm_medium=button&utm_campaign=day-to-day&utm_content=D435_link

- ISO/IEC 11581-10 (2010). *ISO/IEC 11581-10:2010-12 - Informationstechnik - Icons für Benutzerschnittstellen - Teil 10: Rahmenbedingungen und allgemeine Richtlinien*. ISO/IEC TR 25060:2010. *ISO/IEC TR 25060:2010*.
- Jensen, P. L. & Alting, L. (2006). Human Factors in the Management of Production. *CIRP Annals*, 55(1), 457–460. [https://doi.org/10.1016/S0007-8506\(07\)60458-2](https://doi.org/10.1016/S0007-8506(07)60458-2)
- Joseph Redmon & Ali Farhadi (2018). YOLOv3: An Incremental Improvement. *ArXiv*, *abs/1804.02767*.
- Joshi, A. V. (2020). *Machine Learning and Artificial Intelligence*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-26622-6>
- Jost, J., Kirks, T., Mättig, B., Sinsel, A. & Trapp, T. U. (2017). Der Mensch in der Industrie – Innovative Unterstützung durch Augmented Reality. In B. Vogel-Heuser, T. Bauernhansl & M. ten Hompel (Hrsg.), *Springer Reference Technik. Handbuch Industrie 4.0: Bd. 1: Produktion* (2. Aufl., Bd. 21, S. 153–174). Springer Vieweg. https://doi.org/10.1007/978-3-662-45279-0_86
- Kaasinen, E., Aromaa, S., Heikkilä, P. & Liinasuo, M. (2019). Empowering and Engaging Solutions for Operator 4.0 – Acceptance and Foreseen Impacts by Factory Workers. In *Advances in Production Management Systems. Production Management for the Factory of the Future*. Springer. https://doi.org/10.1007/978-3-030-30000-5_75
- Kantowitz, B. H. & Sorkin, R. D. (1983). *Human factors: Understanding people-system relationships*. Wiley. <http://www.loc.gov/catdir/description/wiley031/82019967.html>
- Kendon, A. (1986). *Current issues in the study of gesture*. In *The Biological Foundations of Gestures: Motor and Semiotic Aspects*. Lawrence Erlbaum Associates, Inc.
- Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511807572>
- Khuong, B. M., Kiyokawa, K., Miller, A., La Viola, J. J., Mashita, T. & Takemura, H. (2014, 29. März – 2. April). The effectiveness of an AR-based context-aware assembly support system in object assembly. In *2014 IEEE Virtual Reality (VR)* (S. 57–62). IEEE. <https://doi.org/10.1109/VR.2014.6802051>
- Klette, R. (2014). *Concise Computer Vision*. Springer London. <https://doi.org/10.1007/978-1-4471-6320-6>
- Korn, O., Funk, M. & Schmidt, A. (2015). Assistive Systems for the Workplace. In L. B. Theng (Hrsg.), *Advances in Medical Technologies and Clinical Practice. Assistive technologies for physical and cognitive disabilities* (S. 121–135). IGI Global. <https://doi.org/10.4018/978-1-4666-7373-1.ch006>
- Kosch, T., Funk, M., Schmidt, A. & Chuang, L. L. (2018). Identifying Cognitive Assistance with Mobile Electroencephalography: A Case Study with In-Situ Projections for Manual Assembly. *Proceedings of the ACM on Human-Computer Interaction*, 2(EICS), 1–20. <https://doi.org/10.1145/3229093>
- Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. <https://doi.org/10.1145/3065386>
- Krueger, M. W. (1983). *Artificial Reality*. Addison-Wesley.

- Krueger, M. W. (1993). Environmental technology. *Communications of the ACM*, 36(7), 36–37. <https://doi.org/10.1145/159544.159563>
- La Barré, R. de, Chojecki, P., Leiner, U., Mühlbach, L. & Ruschin, D. (2009). Touchless Interaction-Novel Chances and Challenges. In J. A. Jacko (Hrsg.), *Lecture Notes in Computer Science. Human-Computer Interaction. Novel Interaction Methods and Techniques* (Bd. 5611, S. 161–169). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-02577-8_18
- LAP GmbH Laser Applikationen. (2020). *LAP Laserprojektor*. <https://www.lap-laser.com/products/cad-pro/>
- Latos, B. A., Holtkötter, C., Brinkjans, J., Kalantar, P., Przybysz, P. M. & Mütze-Niewöhner, S. (2018). Partizipatives und simulationsgestütztes Vorgehen zur Konzeption einer flexiblen und demografierobusten Montagelinie. *Zeitschrift für Arbeitswissenschaft*, 72(1), 90–98. <https://doi.org/10.1007/s41449-017-0081-8>
- Lee, J., Bagheri, B. & Kao, H.-A. (2015). A Cyber-Physical Systems architecture for Industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3, 18–23. <https://doi.org/10.1016/j.mfglet.2014.12.001>
- Lorenz, J. (2014). Steigerung der Produktivität durch effizientes Shop Floor Management – neue Tools als Schlüssel zum Erfolg im Werk Leibertingen. *Mahle Global* 06/14.
- Lotter, B. & Schilling, W. (1994). *Manuelle Montage: Planung - Rationalisierung - Wirtschaftlichkeit*.
- Lotter, B. & Wiendahl, H.-P. (2012). *Montage in der industriellen Produktion: Ein Handbuch für die Praxis* (2. Aufl.). VDI-Buch. Springer.
- Lun, R. & Zhao, W. (2015). A Survey of Applications and Human Motion Recognition with Microsoft Kinect. *International Journal of Pattern Recognition and Artificial Intelligence*, 29(05), 1555008. <https://doi.org/10.1142/S0218001415550083>
- Lušić, M., Fischer, C., Bönig, J., Hornfeck, R. & Franke, J. (2016). Worker Information Systems: State of the Art and Guideline for Selection under Consideration of Company Specific Boundary Conditions. *Procedia CIRP*, 41, 1113–1118. <https://doi.org/10.1016/j.procir.2015.12.003>
- Mach, S., Kastrau, A. & Schmalfuß, F. (2018). Information at Hand – Using Wearable Devices to Display Task Information in the Context of Industry 4.0. In C. Stephanidis (Hrsg.), *Communications in Computer and Information Science: Bd. 850. HCI International 2018 – posters' extended abstracts: 20th International Conference, HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018 : proceedings* (Bd. 850, S. 93–100). Springer. https://doi.org/10.1007/978-3-319-92270-6_13
- MacKenzie, S. (2013). *Human-Computer Interaction: An Empirical Research Perspective*. Elsevier. <https://doi.org/10.1016/C2012-0-02819-0>
- March, S. T [Salvatore T.] & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, 15(4), 251–266. [https://doi.org/10.1016/0167-9236\(94\)00041-2](https://doi.org/10.1016/0167-9236(94)00041-2)

- Masood, T. & Egger, J. (2019). Augmented reality in support of Industry 4.0—Implementation challenges and success factors. *Robotics and Computer-Integrated Manufacturing*, 58, 181–195. <https://doi.org/10.1016/j.rcim.2019.02.003>
- MathWorks, Inc. (2020). *Objekt-Detection*. <https://www.mathworks.com/discovery/object-detection.html>
- Mattsson, S., Fast-Berglund, Å., Li, D. & Thorvald, P. (2020). Forming a cognitive automation strategy for Operator 4.0 in complex assembly. *Computers & Industrial Engineering*, 139, 105360. <https://doi.org/10.1016/j.cie.2018.08.011>
- Mattsson, S., Li, D. & Fast-Berglund, Å. (2018). Application of design principles for assembly instructions – evaluation of practitioner use. *Procedia CIRP*, 76, 42–47. <https://doi.org/10.1016/j.procir.2018.02.011>
- Mayrhofer, W., Kames, D. & Schlund, S. (2019). *Made in Austria: Produktionsarbeit in Österreich 2018, Studie*. Technische Universität Wien.
- Mayrhofer, W., Rupprecht, P. & Schlund, S. (2019). One-Fits-All vs. Tailor-Made: User-Centered Workstations for Field Assembly with an Application in Aircraft Parts Manufacturing. *Procedia Manufacturing*, 39, 149–157. <https://doi.org/10.1016/j.promfg.2020.01.287>
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal about Thought*. University of Chicago Press. <https://doi.org/10.2307/1576015>
- Medina, Y. T., Nadeau, S. & Landau, K. (2019). *Assembly Guidance Systems in Aerospace Manufacturing*. <https://substance.etsmtl.ca/en/assembly-guidance-systems-aerospace-manufacturing>
- Mengoni, M., Ceccacci, S., Generosi, A. & Leopardi, A. (2018). Spatial Augmented Reality: an application for human work in smart manufacturing environment. *Procedia Manufacturing*, 17, 476–483. <https://doi.org/10.1016/j.promfg.2018.10.072>
- Monostori, L. (2014). Cyber-physical Production Systems: Roots, Expectations and R&D Challenges. *Procedia CIRP*, 17, 9–13. <https://doi.org/10.1016/j.procir.2014.03.115>
- Moser, C. (2012). *User Experience Design*. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-13363-3>
- MTM ASSOCIATION e. V. (2020). *MTM-Association - MTM-1*. <https://mtm.org/mtm-standard/mtm-prozesssprache>
- Nielsen, J. (1993). *Usability Engineering. Interactive Technologies*. Academic Press, An Imprint of Elsevier. <http://gbv.ebib.com/patron/FullRecord.aspx?p=1190977>
- Niesen, T., Houy, C., Fettke, P. & Loos, P. (2016). Towards an Integrative Big Data Analysis Framework for Data-Driven Risk Management in Industry 4.0. In *2016 49th Hawaii International Conference on System Sciences (HICSS)* (S. 5065–5074). IEEE. <https://doi.org/10.1109/HICSS.2016.627>
- Norman, D. A. (2010). Natural user interfaces are not natural. *Interactions*, 17(3), 6–10. <https://doi.org/10.1145/1744161.1744163>
- Norman, D. A. (2016). *The design of everyday things: Psychologie und Design der alltäglichen Dinge* (C. Eschenfelder, Übers.) (Überarbeitete und erweiterte Auflage). Verlag Franz Vahlen.

- Norman, D. A. & Nielsen, J. (2010). Gestural interfaces. *Interactions*, 17(5), 46–49.
<https://doi.org/10.1145/1836216.1836228>
- O'Mahony, N., Campbell, S., Carvalho, A., Harapanahalli, S., Hernandez, G. V., Krpalkova, L., Riordan, D. & Walsh, J. (2020). Deep Learning vs. Traditional Computer Vision. In K. Arai & S. Kapoor (Hrsg.), *Advances in Intelligent Systems and Computing. Advances in Computer Vision* (Bd. 943, S. 128–144). Springer International Publishing. https://doi.org/10.1007/978-3-030-17795-9_10
- OECD: *The future of work*. (2019). OECD Publishing.
- Okoli, C. & Schabram, K. (2010). A Guide to Conducting a Systematic Literature Review of Information Systems Research. *SSRN Electronic Journal*. Vorab-Onlinepublikation.
<https://doi.org/10.2139/ssrn.1954824>
- Oppermann, R. (Hrsg.). (1994). *Computers, cognition, and work. Adaptive user support: Ergonomic design of manually and automatically adaptable software*. Erlbaum.
<http://www.loc.gov/catdir/enhancements/fy0745/94021630-d.html>
- Paelke, V. (2014). Augmented reality in the smart factory: Supporting workers in an industry 4.0. environment. In *IEEE [International Conference on] Emerging Technologies and Factory Automation (ETFA), 2014: 16 - 19 Sept. 2014, Barcelona, Spain* (S. 1–4). IEEE. <https://doi.org/10.1109/ETFA.2014.7005252>
- Paelke, V., Röcker, C., Koch, N., Flatt, H. & Büttner, S. (2015). User interfaces for cyber-physical systems. *at - Automatisierungstechnik*, 63(10). <https://doi.org/10.1515/auto-2015-0016>
- Panasonic Connect Europe GmbH. *Panasonic PT-RZ660BE*.
<https://business.panasonic.de/visuelle-systeme/pt-rz660>
- Picavi GmbH. (2020). *picavi.com*. <https://picavi.com>
- Pimminger, S., Kurschl, W., Panholzer, L., Neumayr, T., Augstein, M., Altmann, J. & Heinzlreiter, J. (2020). Assembly Task Analysis Using the General Assembly Task Model (GATM) on the Shop Floor. *Procedia CIRP*, 93, 1109–1114.
<https://doi.org/10.1016/j.procir.2020.04.007>
- Posada, J., Toro, C., Barandiaran, I., Oyarzun, D., Stricker, D., Amicis, R. de, Pinto, E. B., Eisert, P., Döllner, J. & Vallarino, I. (2015). Visual computing as a key enabling technology for Industrie 4.0 and Industrial Internet. *IEEE computer graphics and applications*, 35(2), 26–40. <https://doi.org/10.1109/MCG.2015.45>
- Preim, B. & Dachsel, R. (2015). *Interaktive Systeme*. Springer Berlin Heidelberg.
<https://doi.org/10.1007/978-3-642-45247-5>
- Quek, F., McNeill, D., Bryll, R., Duncan, S., Ma, X.-F., Kirbas, C., McCullough, K. E. & Ansari, R. (2002). Multimodal human discourse. *ACM Transactions on Computer-Human Interaction*, 9(3), 171–193. <https://doi.org/10.1145/568513.568514>
- Ramdhani, A., Ramdhani, M. A. & Amin, A. S. (2014). Writing a Literature Review Research Paper: A step-by-step approach. *International Journal of Basics and Applied Sciences*.
- Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. In *2016 IEEE Conference on Computer Vision and*

- Pattern Recognition (CVPR)* (S. 779–788). IEEE.
<https://doi.org/10.1109/CVPR.2016.91>
- Redmon, J. & Farhadi, A. (2017). YOLO9000: Better, Faster, Stronger. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (S. 6517–6525). IEEE. <https://doi.org/10.1109/CVPR.2017.690>
- Reinhart, G. (2017). *Handbuch Industrie 4.0: Geschäftsmodelle, Prozesse, Technik*. Hanser Verlag. <http://dx.doi.org/10.3139/9783446449893>
<https://doi.org/10.3139/9783446449893>
- Rodriguez, L., Quint, F., Gorecky, D., Romero, D. & Siller, H. R. (2015). Developing a Mixed Reality Assistance System Based on Projection Mapping Technology for Manual Operations at Assembly Workstations. *Procedia Computer Science*, 75, 327–333.
<https://doi.org/10.1016/j.procs.2015.12.254>
- Romero, D., Bernus, P., Noran, O., Stahre, J. & Fast-Berglund, Å. (2016). The Operator 4.0: Human Cyber-Physical Systems & Adaptive Automation Towards Human-Automation Symbiosis Work Systems. In I. Nääs, O. Vendrametto, J. Mendes Reis, R. F. Gonçalves, M. T. Silva, G. von Cieminski & D. Kiritsis (Hrsg.), *IFIP Advances in Information and Communication Technology. Advances in Production Management Systems. Initiatives for a Sustainable World* (Bd. 488, S. 677–686). Springer International Publishing. https://doi.org/10.1007/978-3-319-51133-7_80
- Romero, D., Stahre, J. & Taisch, M. (2020). The Operator 4.0: Towards socially sustainable factories of the future. *Computers & Industrial Engineering*, 139, 106128.
<https://doi.org/10.1016/j.cie.2019.106128>
- Romero David, Johan Stahre, Thorsten Wuest, Ovidiu Sever Noran, Peter Bernus, Åsa Fasth Berglund & Dominic Gorecky. TOWARDS AN OPERATOR 4.0 TYPOLOGY: A HUMAN-CENTRIC PERSPECTIVE ON THE FOURTH INDUSTRIAL REVOLUTION TECHNOLOGIES. In *International Conference on Computers & Industrial Engineering CIE 46*.
- Rother, C., Kolmogorov, V. & Blake, A. (2004). "GrabCut". In J. Marks (Hrsg.), *ACM SIGGRAPH 2004 Papers on - SIGGRAPH '04* (S. 309). ACM Press.
<https://doi.org/10.1145/1186562.1015720>
- Rupprecht, P., Kueffner-McCauley, H. & Schlund, S. (2020). Information provision utilizing a dynamic projection system in industrial site assembly. *Procedia CIRP*, 93, 1182–1187. <https://doi.org/10.1016/j.procir.2020.04.053>
- Rupprecht, P., Kueffner-McCauley, H., Trimmel, M. & Schlund, S. (2021). Adaptive Spatial Augmented Reality for Industrial Site Assembly. *Procedia CIRP*, 104, 405–410.
<https://doi.org/10.1016/j.procir.2021.11.068>
- Rupprecht, P. & Schlund, S. (2021). Taxonomy for Individualized and Adaptive Human-Centered Workplace Design in Industrial Site Assembly. In D. Russo, T. Ahram, W. Karwowski, G. Di Bucchianico & R. Taiar (Hrsg.), *Advances in Intelligent Systems and Computing. Intelligent Human Systems Integration 2021* (Bd. 1322, S. 119–126). Springer International Publishing. https://doi.org/10.1007/978-3-030-68017-6_18

- Ryerson University. (2020). *Deep-Learning Netzwerk - Ryerson University Toronto*.
<https://www.cs.ryerson.ca/~aharley/vis/conv/>
- Saffer, D. (2008). *Designing gestural interfaces* (1. Aufl.). *Safari Books Online*. O'Reilly Media. <http://proquest.safaribooksonline.com/9780596156756>
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3), 210–229.
<https://doi.org/10.1147/rd.33.0210>
- Sand, O., Büttner, S., Paelke, V. & Röcker, C. (2016). smARt.Assembly – Projection-Based Augmented Reality for Supporting Assembly Workers. In S. Lackey & R. P. Shumaker (Hrsg.), *Lecture notes in computer science Information systems and applications, incl. Internet/Web and HCI: Bd. 9740. Virtual, augmented and mixed reality: 8th international conference, VAMR 2016, held as part of HCI International 2016, Toronto, Canada, July 17-22, 2016 : proceedings* (Bd. 9740, S. 643–652). Springer.
https://doi.org/10.1007/978-3-319-39907-2_61
- Schäfer, P. (2019). *Skoda testet Videomapping-Projektionen in der Logistik*.
<https://www.springerprofessional.de/automobilproduktion/industrie-4-0/skoda-testet-videomapping-projektionen-in-der-logistik/16559734>
- Schembera-Kneifel, T. & Keil, M. (2016). Future ergonomics tools – From the prototype to the serial product by comprehensive product optimization. In M. Bargende, H.-C. Reuss & J. Wiedemann (Hrsg.), *Proceedings. 16. Internationales Stuttgarter Symposium: Automobil- und Motorentechnik* (S. 807–816). Springer Vieweg.
https://doi.org/10.1007/978-3-658-13255-2_59
- Schlick, C., Bruder, R. & Luczak, H. (2018). *Arbeitswissenschaft*. Springer Berlin Heidelberg.
<https://doi.org/10.1007/978-3-662-56037-2>
- Schlund, S. & Baaij, F. (2018). Describing the technological scope of Industry 4.0-a review of survey publications. *Logforum*, 14(3), 341–353.
<https://doi.org/10.17270/J.LOG.2018.289>
- Schlund, S., Mayrhofer, W. & Rupperecht, P. (2018). Möglichkeiten der Gestaltung individualisierbarer Montagearbeitsplätze vor dem Hintergrund aktueller technologischer Entwicklungen. *Zeitschrift für Arbeitswissenschaft*, 72(4), 276–286.
<https://doi.org/10.1007/s41449-018-0128-5>
- Schmidt, R., Möhring, M., Härting, R.-C., Reichstein, C., Neumaier, P. & Jozinović, P. (2015). Industry 4.0 - Potentials for Creating Smart Products: Empirical Research Results. In W. Abramowicz (Hrsg.), *Lecture Notes in Business Information Processing. Business Information Systems* (Bd. 208, S. 16–27). Springer International Publishing.
https://doi.org/10.1007/978-3-319-19027-3_2
- Scott, A. (2017). *Boeing forms venture group, invests in two tech startups*. <http://www.reuters.com/article/us-boeing-venture-idUSKBN1772BE>
- Segura, Á., Diez, H. V., Barandiaran, I., Arbelaiz, A., Álvarez, H., Simões, B., Posada, J., García-Alonso, A. & Ugarte, R. (2020). Visual computing technologies to support the Operator 4.0. *Computers & Industrial Engineering*, 139, 105550.
<https://doi.org/10.1016/j.cie.2018.11.060>

- Siemens AG. (2021). *Endmontage der Lokomotive Vectron im Werk München (Siemens)*.
<https://press.siemens.com/global/de/feature/vectron-das-fahrzeugkonzept>
- Sirajum Munir, John A. Stankovic, Chieh-Jan Mike Liang & Shan Lin (2013). Cyber Physical System Challenges for Human-in-the-Loop Control. In *Feedback Computing*.
- ŠKODA AUTO Deutschland GmbH. (2020). *ŠKODA AUTO testet Videomapping: Augmented Reality hilft in der Logistik beim Bestücken von Paletten*. <https://www.skoda-media.de/press/detail/3060/>
- SL-Laser GmbH. (2020a). *Positionierlaser & Projektionslaser für Composite Anwendungen*.
<https://www.sl-laser.com/index.php/anwendungen/composite>
- SL-Laser GmbH. (2020b). *SL-Laser Laserprojektor*. <https://www.sl-laser.com/index.php/support/installation-et-formation-contrats-de-maintenance/2-sl-laser-systeme/6-prodirector-6>
- Sochor, R., Kraus, L., Merkel, L., Braunreuther, S. & Reinhart, G. (2019). Approach to Increase Worker Acceptance of Cognitive Assistance Systems in Manual Assembly. *Procedia CIRP*, 81, 926–931. <https://doi.org/10.1016/j.procir.2019.03.229>
- Spath, D. (Hrsg.). (2013). *Produktionsarbeit der Zukunft - Industrie 4.0: Studie*. Fraunhofer Verlag.
- Stankovic, J. A. (2014). Research Directions for the Internet of Things. *IEEE Internet of Things Journal*, 1(1), 3–9. <https://doi.org/10.1109/JIOT.2014.2312291>
- Stapelkamp, T. (2007). *Screen- und Interfacedesign*. Springer Berlin Heidelberg.
<https://doi.org/10.1007/978-3-540-32950-3>
- Statistik Austria. (2019). *WKO Statistik: Industriebeschäftigte nach Branchen*.
<http://wko.at/statistik/jahrbuch/sach-ibesch-branchen.pdf>
- Stereolabs Inc. (2021). *Stereolabs - ZED 2*. <https://www.stereolabs.com/zed-2>
- Szeliski, R. (2011). *Computer vision: Algorithms and applications. Texts in computer science*. Springer. <http://site.ebrary.com/lib/alltitles/docDetail.action?docID=10421311>
<https://doi.org/10.1007/978-1-84882-935-0>
- Tang, A., Owen, C., Biocca, F. & Mou, W. (2003). Comparative effectiveness of augmented reality in object assembly. In G. Cockton & P. Korhonen (Hrsg.), *Proceedings of the conference on Human factors in computing systems - CHI '03* (S. 73). ACM Press.
<https://doi.org/10.1145/642611.642626>
- Taylor, M. P., Boxall, P., Chen, J. J., Xu, X., Liew, A. & Adeniji, A. (2020). Operator 4.0 or Maker 1.0? Exploring the implications of Industrie 4.0 for innovation, safety and quality of work in small economies and enterprises. *Computers & Industrial Engineering*, 139, 105486. <https://doi.org/10.1016/j.cie.2018.10.047>
- Technische Universität Wien - Pilotfabrik für Industrie 4.0. (2020). *pilotfabrik.tuwien.ac.at*.
pilotfabrik.tuwien.ac.at
- Theel, S. (2015). *Kommissionierung im 21. Jahrhundert: Von Pick-by-Voice bis RFID*. Diplomica Verlag GmbH.
<http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&AN=1075508>

- van Aken, J. E. (2004). Management Research Based on the Paradigm of the Design Sciences: The Quest for Field-Tested and Grounded Technological Rules. *Journal of Management Studies*, 41(2), 219–246. <https://doi.org/10.1111/j.1467-6486.2004.00430.x>
- VDI. (1978). *VDI 2815 - Begriffe für die Produktionsplanung und -steuerung: Terms of production engineering and control* (VDI-Richtlinien VDI 2815). Berlin. Verein Deutscher Ingenieure.
- VDI. (1990). *VDI-Richtlinie 2860 - Montage- und Handhabungstechnik: Handhabungsfunktionen, Handhabungseinrichtungen; Begriffe, Definitionen, Symbole. VDI-Richtlinien*.
- Vogel-Heuser, B., Bauernhansl, T. & Hompel, M. ten (Hrsg.). (2017). *Springer Reference Technik. Handbuch Industrie 4.0: Bd. 1: Produktion* (2. Aufl.). Springer Vieweg. <http://dx.doi.org/10.1007/978-3-662-45279-0> <https://doi.org/10.1007/978-3-662-45279-0>
- Voulodimos, A., Doulamis, N., Doulamis, A. & Protopapadakis, E. (2018). Deep Learning for Computer Vision: A Brief Review. *Computational intelligence and neuroscience*, 2018, 7068349. <https://doi.org/10.1155/2018/7068349>
- Wachs, J. P., Kölsch, M., Stern, H. & Edan, Y. (2011). Vision-based hand-gesture applications. *Communications of the ACM*, 54(2), 60–71. <https://doi.org/10.1145/1897816.1897838>
- Walls, J. G., Widmeyer, G. R. & El Sawy, O. A. (1992). Building an Information System Design Theory for Vigilant EIS. *Information Systems Research*, 3(1), 36–59. <https://doi.org/10.1287/isre.3.1.36>
- Wang, Y., Zhang, S., Yang, S., He, W. & Bai, X. (2018). Mechanical assembly assistance using marker-less augmented reality system. *Assembly Automation*, 38(1), 77–87. <https://doi.org/10.1108/AA-11-2016-152>
- Warnecke, H.-J., Löhr, H.-G. & Kiener, W. (1975). *Montagetechnik: Schwerpunkt der Rationalisierung. Buchreihe Produktionstechnik heute: Bd. 7*. Krausskopf.
- Weidner, R., Redlich, T. & Wulfsberg, J. P. (Hrsg.). (2015). *Technische Unterstützungssysteme*. Springer Vieweg. <http://gbv.ebib.com/patron/FullRecord.aspx?p=4182841> <https://doi.org/10.1007/978-3-662-48383-1>
- Weisner, K., Knittel, M., Enderlein, H., Wischniewski, S., Jaitner, T., Kuhlang, P. & Deuse, J. (2016). Assistenzsystem zur Individualisierung der Arbeitsgestaltung. *ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb*, 111(10), 598–601. <https://doi.org/10.3139/104.111609>
- Weiss, A., Bernhaupt, R., Schwaiger, D., Altmaninger, M., Buchner, R. & Tscheligi, M. (2009). User experience evaluation with a Wizard of Oz approach: Technical and methodological considerations. In *2009 9th IEEE-RAS International Conference on Humanoid Robots* (S. 303–308). IEEE. <https://doi.org/10.1109/ICHR.2009.5379559>
- Wellner, P. (1993). Interacting with paper on the DigitalDesk. *Communications of the ACM*, 36(7), 87–96. <https://doi.org/10.1145/159544.159630>

- Wichmann, R. L., Eisenbart, B. & Gericke, K. (2019). The Direction of Industry: A Literature Review on Industry 4.0. *Proceedings of the Design Society: International Conference on Engineering Design*, 1(1), 2129–2138. <https://doi.org/10.1017/dsi.2019.219>
- Wiendahl, H. P., Elmaraghy, H., Nyhuis, P., Zäh, M. F., Wiendahl, H. H., Duffie, N. & Brieke, M. (2007). Changeable manufacturing classification. *Design and Operation. Annals of CIRP*, 55, 783–809.
- Wright, I. (2017). *Airbus Uses Smart Glasses to Improve Manufacturing Efficiency*. <http://www.engineering.com/AdvancedManufacturing/ArticleID/14634/Airbus-Uses-Smart-Glasses-to-Improve-Manufacturing-Efficiency.aspx>
- Zhao, Z.-Q., Zheng, P., Xu, S.-T. & Wu, X. (2019). Object Detection With Deep Learning: A Review. *IEEE transactions on neural networks and learning systems*, 30(11), 3212–3232. <https://doi.org/10.1109/TNNLS.2018.2876865>
- Zolotová, I., Papcun, P., Kajáti, E., Miškuf, M. & Mocnej, J. (2020). Smart and cognitive solutions for Operator 4.0: Laboratory H-CPPS case studies. *Computers & Industrial Engineering*, 139, 105471. <https://doi.org/10.1016/j.cie.2018.10.032>

Anhang

Anhang 1: Angebot 18-0449 vom 14.11.2018 von Dynamic Projection Institute

Anhang 2: Fragebogen Experiment 1 – PC-Terminal

Anhang 3: Fragebogen Experiment 1 – Dynamische Projektion

Anhang 4: Fragebogen Experiment 2 – In-view-/In-situ-Projektion

Anhang 5: Fragebogen Experiment 3 – TABLET-Anweisung/DYPRO-In-situ-Projektion

Anhang 1: Angebot 18-0449 vom 14.11.2018 von Dynamic Projection Institute



Abs.: Dynamic Projection Institute GmbH, Renngasse 4/1/1, 1010 Wien

Technische Universität Wien - Institut
 für Managementwissenschaften
 Univ. Prof. Dr.-Ing. Sebastian Schlund
 Theresianumgasse 27
 1040 Wien
 Österreich

Datum: 14.11.2018
 Kudenr.: C00289
 Telefonnr.: +43 (1) 58801 - 33054
 Mobil: +43 664 605882401
 E-Mail: sebastian.schlund@tuwien.ac.at
 UID Nr.: ATU37675002

Angebot: 18-0449 / Projekt: TU Wien Pilotfabrik Industrie 4.0

Ansprechpartner: Thomas Kühne

Wir danken für Ihr Interesse an unseren Produkten und bieten freibleibend an:

Pos.	Artikelbezeichnung	Menge	Einzelpreis	Rabatt	Gesamtpreis
1	[MH11-L] Mirror Head MH11 Mirror Head für mittelgrosse 1 Chip DLP Projektoren mit bis zu 20Kg Eigengewicht, ArtNet und DMX, RS232 Projektorsteuerung, AC Durchschliff für den Projektor bis 8A / 240V, RGB LED 18W, Gewicht 10,8Kg, EU Tarifnummer: 8529.90 97 10	1,00 Stk.	2.700,00 €	15,00 %	2.295,00 €
2	[MHZ-BW] Montagebügel Wand MHZ-BW Montagebügel Wand für Mirror Head MH09, MH10, MH11, MH12 und MH13 mit MDC VESA Aufnahme , Gewicht 2,8Kg	1,00 Stk.	220,00 €	15,00 %	187,00 €
3	[PT-RZ660BE] Panasonic PT-RZ660BE PT-RZ660BE laser Projektor schwarz, inklusive Standard Objektiv T.R.: 1.7 - 2.4:1, aus der Solid Shine Laserprojektorserie, Flexible 360-Grad-Installation und Geometric Manager Pro-Funktion, Helligkeit 6200 ANSI Lumen, Auflösung: 1920 x 1200 WUXGA, Kontrast 10.000:1, Gewicht 22,4 kg, Garantie: 36 Monate BringIn-Service	1,00 Stk.	13.880,00 €	33,00 %	9.299,60 €

Dynamic Projection Institute GmbH
 Herstellungs und Vertriebs GmbH
 Renngasse 4 / Stiege 1 / Top 1
 A-1010 Wien / Vienna
 UID: ATU68086812, FN 400806 x

Telefon: +43 1 9962028
 Fax: +43 1 9962028 88
<http://www.dynamicprojection.com>
office@dynamicprojection.com
 ARA: 20261

Bankverbindung: Erste Bank der
 Österreichischen Sparkassen AG
 IBAN: AT10 2011 1822 3833 9000
 BIC: GIBAATWWXXX
 EORI: ATEOS1000065271



DYNAMIC PROJECTION
INSTITUTE

4	[MDCX2-DH170] MDC-X2/DH170 Media Server - Academic license MDC-X2/DH170 - Media on Demand Control, speziell für den Mirror Head entwickelt, 2x Display Port, 1x HDMI, 2x RJ45, Betriebssystem Linux, Wlan IEEE 802.11b/g/n Standard, 300 MBit/s, INTEL Core i5-6600, 3,30GHz, 120 GB SSD HD, Gewicht 1,4Kg, EU Tarifnummer: 8529.90 97 10 Mit der Inbetriebnahme stimmen Sie der EULA zu.	1,00 Stk.	2.170,00 €	25,00 %	1.627,50 €
5	[OSC-X2T] MDC Touch X2 - Academic license Webinterface für die MDC-X2 / Media Server; erlaubt die Fernsteuerung von Basisfunktionen bis hin zu detaillierten Showabläufen, Thumbnails der Medien, Echtzeit update, Bedienung über jeden Webbrowser möglich und somit auch über Smartphone und Tablet.	1,00 Stk.	341,00 €	25,00 %	255,75 €
6	[JCD-ST-v1.9] JCD Standard Edition - Academic license JCD - box control over the projector via web interface, Art-Net (DMX512), OSC or MQTT for projectors with motor driven lenses. Control possibilities include, but are not limited to absolute focus and zoom, lens memory, shutter and brightness control for laser projectors.	1,00 Stk.	650,00 €	25,00 %	487,50 €

Gesamtbetrag netto: 14.152,35 €

MWSt. 20% 2.830,47 €

Gesamtbetrag brutto: 16.982,82 €

Gültigkeitsdauer des Angebots: 2 Wochen
 Lieferzeit: on stock
 Lieferung: EX WORKS
 Zahlungsmodalitäten: 50% Anzahlung, 50% bei Auslieferung
 Angebot firmenmäßig unterfertigt, gilt als Bestellung.

Die Waren bleiben bis zur vollständigen Bezahlung unser Eigentum, insbesondere gilt auch verlängerter Eigentumsvorbehalt. Als Gerichtsstand gilt Wien als vereinbart. Mit der Bestellung bestätige ich, dass ich die Allgemeinen Geschäftsbedingungen Verkauf Dynamic Projection Institute, Herstellungs und Vertriebs GmbH gelesen, verstanden und akzeptiert habe.
<http://www.dynamicprojection.com/agb/>

Dynamic Projection Institute GmbH
 Herstellungs und Vertriebs GmbH
 Renngasse 4 / Stiege 1 / Top 1
 A-1010 Wien / Vienna
 UID: ATU68086812, FN 400806 x

Telefon: +43 1 9962028
 Fax: +43 1 9962028 88
<http://www.dynamicprojection.com>
 office@dynamicprojection.com
 ARA: 20261

Bankverbindung: Erste Bank der
 Österreichischen Sparkassen AG
 IBAN: AT10 2011 1822 3833 9000
 BIC: GIBAATWWXXX
 EORI: ATEOS1000065271

Anhang 2: Fragebogen Experiment 1 – PC-Terminal

Seite 1 von 3

Evaluation von Assistenzsystemen – FanCowl – Experiment

TERMINAL State of the Art – Informationen und Interaktion über Terminal-PC

Abschnitt	Inhalt
A	Datenschutzinformation
B	Angaben zur Person
C	Angaben zum Use Case
D	Fragen zur Usability mittels SUS (System Usability Scale)
E	Fragen zur Nutzerakzeptanz mittels TAM (Technology Acceptance Model)
F	Fragen zu Qualifikation, mentale Belastung und Flexibilität
G	Feedback

A. Datenschutzinformation

Die Befragung läuft anonym. Falls Sie für Rückfragen freiwillig zur Verfügung stehen möchten, ergänzen Sie bitte Ihre Kontaktdaten auf der letzten Seite. Wenn die Signifikanz der Daten nicht gegeben ist und daher eine Rückführung auf eine Person ermöglichen würde, werden die Daten entfernt. Die Daten werden für die Auswertung von Nutzerakzeptanz und Usability und für dieses Projekt verwendet. Interessante Zusammenhänge werden unter Einhaltung der oben angegebenen Punkte wissenschaftlich publiziert.

Unterschrift: _____

B. Angaben zur Person

Allgemeine Daten			
Geschlecht	☐ Männlich(m)	☐ Weiblich(w)	☐ Divers(div)
Altersgruppe	☐ Unter 20	☐ 20 bis 29	☐ 30 bis 39
	☐ 40 bis 49	☐ 50 bis 59	☐ Über 59
Tätigkeit/Position	☐ Führungskraft (z.B. FACC)	☐ MitarbeiterIn (z.B. FACC)	☐ StudentIn TU Wien
	☐ PraktikantIn (z.B. FACC)	☐ BesucherIn	☐ StudentIn Uni/FH:
BrillenträgerIn	☐ ja	☐ nein	☐ keine Angabe

[B.1] Wie häufig nutzen Sie folgende Technologien?

	Täglich	2-9x/Monat	6-12x/Jahr	Einmalig	Nie
Smartphone	☐	☐	☐	☐	☐
Tablet	☐	☐	☐	☐	☐
Augmented Reality AR am Smartphone/Tablet	☐	☐	☐	☐	☐
AR-Datenbrillen	☐	☐	☐	☐	☐
Virtual Reality VR Board	☐	☐	☐	☐	☐
VR Datenbrillen	☐	☐	☐	☐	☐

[B.2] Wie häufig spielen Sie aktuell auf folgenden Geräten?

	Täglich	2-9x/Monat	6-12x/Jahr	Einmalig	Nie
Brett-/Kartenspiele	☐	☐	☐	☐	☐
Gaming am Smartphone/Tablet	☐	☐	☐	☐	☐
Gaming am PC	☐	☐	☐	☐	☐
Gaming mit AR	☐	☐	☐	☐	☐
Gaming in VR	☐	☐	☐	☐	☐

C. Angaben zum Use Case

[C.1] Getesteter Use Case: **Fan Cowl – Industrielle Baustellenmontage – Pilotfabrik TU Wien**

D. Fragen zur Usability mittels SUS (System Usability Scale)

Aussage	Stimme überhaupt nicht zu 1	Stimme eher nicht zu 2	Weder noch 3	Stimme eher zu 4	Stimme voll zu 5
Ich denke, dass ich das System gerne häufig benutzen würde.	0	0	0	0	0
Ich fand das System unnötig komplex.	0	0	0	0	0
Ich fand das System einfach zu benutzen.	0	0	0	0	0
Ich glaube, ich würde die Hilfe einer technisch versierten Person benötigen, um das System benutzen zu können.	0	0	0	0	0
Ich fand, die verschiedenen Funktionen in diesem System waren gut integriert.	0	0	0	0	0
Ich denke, das System enthielt zu viele Inkonsistenzen/Widersprüche.	0	0	0	0	0
Ich kann mir vorstellen, dass die meisten Menschen den Umgang mit diesem System sehr schnell lernen.	0	0	0	0	0
Ich fand das System sehr umständlich zu nutzen.	0	0	0	0	0
Ich fühlte mich bei der Benutzung des Systems sehr sicher.	0	0	0	0	0
Ich musste eine Menge lernen, bevor ich anfangen konnte das System zu verwenden.	0	0	0	0	0

E. Fragen zur Nutzerakzeptanz

Das Assistenzsystem... (Informationen und Interaktion am Terminal-PC)	---	--	-	+	++	+++	Das Assistenzsystem... (Informationen und Interaktion am Terminal-PC)
bietet mir schlechtere Unterstützung über die auszuführende Tätigkeit	0	0	0	0	0	0	bietet mir bessere Unterstützung über die auszuführende Tätigkeit
verschlechtert meine Arbeitsleistung	0	0	0	0	0	0	verbessert meine Arbeitsleistung.
geht nicht auf meine Bedürfnisse ein.	0	0	0	0	0	0	geht auf meine Bedürfnisse ein.
kostet mir mehr Zeit im Prozessablauf.	0	0	0	0	0	0	spart mir Zeit im Prozessablauf.
bremst mich beim Erledigen meiner Aufgaben.	0	0	0	0	0	0	ermöglicht mir meine Aufgaben schneller zu erledigen.
schränkt mich ein, wodurch ich weniger Arbeit leisten kann als sonst.	0	0	0	0	0	0	ermöglicht mir mehr Arbeit zu leisten, als es sonst möglich wäre.
erhöht die Zeit, die ich für unproduktive Aktivitäten aufwende.	0	0	0	0	0	0	reduziert die Zeit, die ich für unproduktive Aktivitäten aufwende.
verschlechtert meine Effektivität bei der Arbeit.	0	0	0	0	0	0	verbessert meine Effektivität bei der Arbeit.
verschlechtert die Qualität meiner Arbeit.	0	0	0	0	0	0	verbessert die Qualität meiner Arbeit.
verschlechtert meine Produktivität am Arbeitsplatz.	0	0	0	0	0	0	erhöht meine Produktivität am Arbeitsplatz.
erschwert mir meine Arbeit.	0	0	0	0	0	0	erleichtert mir meine Arbeit.
finde ich für meinen Job nicht nützlich.	0	0	0	0	0	0	finde ich für meinen Job nützlich.
könnte mir in Zukunft meinen Job erschweren.	0	0	0	0	0	0	könnte mir in Zukunft meinen Job vereinfachen.

Anhang 3: Fragebogen Experiment 1 – Dynamische Projektion

Seite 1 von 3

Evaluation von Assistenzsystemen – FanCowl – Experiment

DYNAMISCHE PROJEKTION – Informationen und Interaktion über das Dynamische Projektionssystem

Abschnitt	Inhalt
A	Datenschutzinformation
B	Angaben zur Person
C	Angaben zum Use Case
D	Fragen zur Usability mittels SUS (System Usability Scale)
E	Fragen zur Nutzerakzeptanz mittels TAM (Technology Acceptance Model)
F	Fragen zu Qualifikation, mentale Belastung und Flexibilität
G	Feedback

A. Datenschutzinformation

Die Befragung läuft anonym. Falls Sie für Rückfragen freiwillig zur Verfügung stehen möchten, ergänzen Sie bitte Ihre Kontaktdaten auf der letzten Seite. Wenn die Signifikanz der Daten nicht gegeben ist und daher eine Rückführung auf eine Person ermöglichen würde, werden die Daten entfernt. Die Daten werden für die Auswertung von Nutzerakzeptanz und Usability und für dieses Projekt verwendet. Interessante Zusammenhänge werden unter Einhaltung der oben angegebenen Punkte wissenschaftlich publiziert.

Unterschrift: _____

B. Angaben zur Person

Allgemeine Daten			
Geschlecht	☐ Männlich(m)	☐ Weiblich(w)	☐ Divers(div)
Altersgruppe	☐ Unter 20	☐ 20 bis 29	☐ 30 bis 39
	☐ 40 bis 49	☐ 50 bis 59	☐ Über 59
Tätigkeit/Position	☐ Führungskraft (z.B. FACC)	☐ MitarbeiterIn (z.B. FACC)	☐ StudentIn TU Wien
	☐ PraktikantIn (z.B. FACC)	☐ BesucherIn	☐ StudentIn Uni/FH:
BrillenträgerIn	☐ ja	☐ nein	☐ keine Angabe

[B.1] Wie häufig nutzen Sie folgende Technologien?

	Täglich	2-9x/Monat	6-12x/Jahr	Einmalig	Nie
Smartphone	☐	☐	☐	☐	☐
Tablet	☐	☐	☐	☐	☐
Augmented Reality AR am Smartphone/Tablet	☐	☐	☐	☐	☐
AR-Datenbrillen	☐	☐	☐	☐	☐
Virtual Reality VR Board	☐	☐	☐	☐	☐
VR Datenbrillen	☐	☐	☐	☐	☐

[B.2] Wie häufig spielen Sie aktuell auf folgenden Geräten?

	Täglich	2-9x/Monat	6-12x/Jahr	Einmalig	Nie
Brett-/Kartenspiele	☐	☐	☐	☐	☐
Gaming am Smartphone/Tablet	☐	☐	☐	☐	☐
Gaming am PC	☐	☐	☐	☐	☐
Gaming mit AR	☐	☐	☐	☐	☐
Gaming in VR	☐	☐	☐	☐	☐

C. Angaben zum Use Case

[C.1] Getesteter Use Case: Fan Cowl – Industrielle Baustellenmontage – Pilotfabrik TU Wien

D. Fragen zur Usability mittels SUS (System Usability Scale)

Aussage	Stimme überhaupt nicht zu 1	Stimme eher nicht zu 2	Weder noch 3	Stimme eher zu 4	Stimme voll zu 5
Ich denke, dass ich das System gerne häufig benutzen würde.	0	0	0	0	0
Ich fand das System unnötig komplex.	0	0	0	0	0
Ich fand das System einfach zu benutzen.	0	0	0	0	0
Ich glaube, ich würde die Hilfe einer technisch versierten Person benötigen, um das System benutzen zu können.	0	0	0	0	0
Ich fand, die verschiedenen Funktionen in diesem System waren gut integriert.	0	0	0	0	0
Ich denke, das System enthielt zu viele Inkonsistenzen/Widersprüche.	0	0	0	0	0
Ich kann mir vorstellen, dass die meisten Menschen den Umgang mit diesem System sehr schnell lernen.	0	0	0	0	0
Ich fand das System sehr umständlich zu nutzen.	0	0	0	0	0
Ich fühlte mich bei der Benutzung des Systems sehr sicher.	0	0	0	0	0
Ich musste eine Menge lernen, bevor ich anfangen konnte das System zu verwenden.	0	0	0	0	0

E. Fragen zur Nutzerakzeptanz

Das Assistenzsystem... (Informationen und Interaktion über das Dynamische Projektionssystem)	---	--	-	+	++	+++	Das Assistenzsystem... (Informationen und Interaktion über das Dynamische Projektionssystem)
bietet mir schlechtere Unterstützung über die auszuführende Tätigkeit	0	0	0	0	0	0	bietet mir bessere Unterstützung über die auszuführende Tätigkeit
verschlechtert meine Arbeitsleistung	0	0	0	0	0	0	verbessert meine Arbeitsleistung.
geht nicht auf meine Bedürfnisse ein.	0	0	0	0	0	0	geht auf meine Bedürfnisse ein.
kostet mir mehr Zeit im Prozessablauf.	0	0	0	0	0	0	spart mir Zeit im Prozessablauf.
bremst mich beim Erledigen meiner Aufgaben.	0	0	0	0	0	0	ermöglicht mir meine Aufgaben schneller zu erledigen.
schränkt mich ein, wodurch ich weniger Arbeit leisten kann als sonst.	0	0	0	0	0	0	ermöglicht mir mehr Arbeit zu leisten, als es sonst möglich wäre.
erhöht die Zeit, die ich für unproduktive Aktivitäten aufwende.	0	0	0	0	0	0	reduziert die Zeit, die ich für unproduktive Aktivitäten aufwende.
verschlechtert meine Effektivität bei der Arbeit.	0	0	0	0	0	0	verbessert meine Effektivität bei der Arbeit.
verschlechtert die Qualität meiner Arbeit.	0	0	0	0	0	0	verbessert die Qualität meiner Arbeit.
verschlechtert meine Produktivität am Arbeitsplatz.	0	0	0	0	0	0	erhöht meine Produktivität am Arbeitsplatz.
erschwert mir meine Arbeit.	0	0	0	0	0	0	erleichtert mir meine Arbeit.
finde ich für meinen Job nicht nützlich.	0	0	0	0	0	0	finde ich für meinen Job nützlich.
könnte mir in Zukunft meinen Job erschweren.	0	0	0	0	0	0	könnte mir in Zukunft meinen Job vereinfachen.

F. Fragen zu Qualifikation, mentale Belastung und Flexibilität

Das Assistenzsystem...	---	--	-	-/+	+	++	+++	Das Assistenzsystem...
erhöht meine erforderte Qualifikation. (dadurch benötige ich eine höhere Qualifikation)	0	0	0	0	0	0	0	Reduziert meine erforderte Qualifikation. (dadurch benötige ich eine geringere Qualifikation)
erhöht meine mentale Belastung.	0	0	0	0	0	0	0	reduziert meine mentale Belastung.
reduziert die Flexibilität des Prozesses.	0	0	0	0	0	0	0	erhöht die Flexibilität des Prozesses.

G. Feedback

[G.1] Haben Sie Feedback zum Fragebogen und zum Experiment / Evaluierung?

[G.2] Möchten Sie uns noch etwas mitteilen?

[G.3] Falls Sie für Rückfragen zur Verfügung stehen, geben Sie bitte Ihren Namen und Ihre E-Mail-Adresse an:

Anhang 4: Fragebogen Experiment 2 – In-view-/In-situ-Projektion

Fragebogens zur Bewertung des Assistenzsystems

Datenschutzinformation

Die Befragung läuft anonym. Falls Sie für Rückfragen freiwillig zur Verfügung stehen möchten, ergänzen Sie bitte Ihre Kontaktdaten. Die Daten werden für die Auswertung und für die Weiterentwicklung des bestehenden Projekts verwendet. Interessante Zusammenhänge werden unter Einhaltung der oben angegebenen Punkte wissenschaftlich publiziert. Um Ihre Daten ANONYM über die einzelnen Befragungsrunden verarbeiten zu können, wird von Ihnen ein CODE generiert, der eine anonyme Zuordnung der Fragebögen ermöglicht.

- 1.1. Bitte geben Sie die vierte Ziffer Ihres Geburtsjahres ein: ____ (z.B. 199**2**)
- 1.2. Bitte geben Sie den ersten Buchstaben Ihres Geburtsortes ein: ____ (z.B. **B**aden)
- 1.3. Bitte geben Sie den ersten Buchstaben des Vornamens Ihrer Mutter ein: ____ (z.B. **S**usanne)
- 1.4. Bitte geben Sie den ersten Buchstaben des Vornamens Ihres Vaters ein: ____ (z.B. **G**ünter)

1. Bitte kreuzen Sie an welchen Prototyp sie getestet haben:
- ☐ Spatial Augmented Reality – **Static IN-VIEW-Arbeitsanweisung**
 - ☐ Spatial Augmented Reality – **IN-SITU-Arbeitsanweisung**

Angaben zur Person und Technologieverwendung:

Geschlecht	<i>O Männlich(m)</i>	<i>O Weiblich(w)</i>	<i>O Divers(div)</i>
Altersgruppe	<i>O Unter 20</i>	<i>O 20 bis 29</i>	<i>O 30 bis 39</i>
	<i>O 40 bis 49</i>	<i>O 50 bis 59</i>	<i>O Über 59</i>
Tätigkeit/Position	<i>O Führungskraft</i>	<i>O Arbeiter/Angestellte</i>	<i>O StudentIn</i> <i>O Andere:</i>
Technischer Background (Beruflich/Ausbildung/)	<i>O JA</i>	<i>O NEIN</i>	

Wie häufig nutzen Sie folgende Technologien?

	<i>Täglich</i>	<i>2-9x/Monat</i>	<i>6-12x/Jahr</i>	<i>Einmalig</i>	<i>Nie</i>
Smartphone	<i>O</i>	<i>O</i>	<i>O</i>	<i>O</i>	<i>O</i>
Tablet	<i>O</i>	<i>O</i>	<i>O</i>	<i>O</i>	<i>O</i>
Augmented Reality (AR) am Smartphone/Tablet	<i>O</i>	<i>O</i>	<i>O</i>	<i>O</i>	<i>O</i>
AR Brillen	<i>O</i>	<i>O</i>	<i>O</i>	<i>O</i>	<i>O</i>
Virtual Reality (VR)	<i>O</i>	<i>O</i>	<i>O</i>	<i>O</i>	<i>O</i>

2. Bitte bewerten Sie Ihre Erfahrung (User-Experience) mit dem Assistenzsystem im Hinblick auf die folgenden Aussagen. (Structured Usability Score - SUS)

	<i>Ich stimme auf keinen Fall zu</i>	<i>Ich stimme nicht zu</i>	<i>neutral</i>	<i>Ich stimme zu</i>	<i>Ich stimme voll zu</i>
Ich denke, dass ich dieses System häufig nutzen möchte.					
Ich fand das System unnötig komplex.					
Ich fand das System einfach zu bedienen.					
Ich denke, ich würde die Unterstützung einer technischen Fachkraft benötigen, um dieses System benutzen zu können.					
Ich fand, dass die verschiedenen Funktionen in diesem System gut integriert sind.					
Ich fand, dass es in diesem System zu viele Inkonsistenzen gab.					
Ich könnte mir vorstellen, dass die meisten Menschen den Umgang mit diesem System sehr schnell lernen würden					
Ich fand das System sehr umständlich in der Benutzung.					
Ich fühlte mich sehr sicher im Gebrauch mit dem System.					
Ich musste eine Menge Dinge lernen, bevor ich mit dieses System arbeiten konnte.					

3. Bitte beurteilen Sie Ihre Ansichten über das Assistenzsystem im Hinblick auf die folgenden Aussagen. (Technologie-/Nutzer- Akzeptanz mit „Technology Compatibility“)

	Ich stimme auf keinen Fall zu	Ich stimme nicht zu	neutral	Ich stimme zu	Ich stimme voll zu
Die Nutzung des Assistenzsystems ist eine neue Erfahrung für mich					
Die Nutzung des Assistenzsystems ist nicht mit dem vergleichbar, was ich bisher gemacht habe.					
Die Nutzung des Assistenzsystems unterscheidet sich von anderen Erfahrungen, die ich gemacht habe.					
Die Nutzung des Assistenzsystems ist eine neue Arbeitserfahrung für mich.					
Die Anwendung des Assistenzsystems in der Praxis würde meinen eigenen Werten widersprechen.					
Die Nutzung des Assistenzsystems passt nicht zu meiner Sicht der Welt.					
Die Verwendung des Assistenzsystems steht im Widerspruch zu dem, wofür Computer meiner Meinung nach eingesetzt werden sollten.					
Die Verwendung des Assistenzsystems ist für eine Person mit meinen Wertvorstellungen bezüglich der Rolle von Computern nicht geeignet.					
Die Nutzung des Assistenzsystems widerspricht meinen Wertvorstellungen darüber, wie gearbeitet werden sollte					
Die Verwendung eines solchen Assistenzsystems würde zu meiner bevorzugten Arbeitsroutine passen.					
Das Assistenzsystem würde es mir ermöglichen, in der von mir bevorzugten Weise zu arbeiten.					
Die Nutzung des Assistenzsystems würde gut zu meiner Arbeitsweise passen.					
Die Nutzung des Assistenzsystems würde meiner bevorzugten Arbeitsmethode entsprechen.					

4. Bitte bewerten Sie Ihre Erfahrungen mit der Lösung der Aufgabe mit dem Assistenzsystem im Hinblick auf die folgenden Aussagen (NASA-RTLX)

	hoch		mittel		gering
Wie geistig anspruchsvoll war die Aufgabe?					
Wie körperlich anstrengend war die Aufgabe?					
Wie eilig oder gehetzt war das Arbeitstempo?					
Wie erfolgreich waren Sie dabei, das zu tun, was von Ihnen verlangt wurde?					
Wie intensiv mussten Sie arbeiten, um ihr Leistungsniveau zu erreichen?					
Wie unsicher, entmutigend, irritiert, gestresst und genervt waren Sie?					

5. Welche Probleme sind Ihnen bei der Arbeit mit dem Assistenzsystem aufgefallen?

6. Was möchten Sie uns positives bei der Arbeit mit dem Assistenzsystem mitteilen?

7. Wie beurteilen Sie die Gesamt-Ergonomie des Assistenzsystems?
o sehr gut o gut o mittel o schlecht o sehr schlecht

8. Wie viele Minuten könnten Sie sich vorstellen, mit dem Assistenzsystem zu arbeiten...

	unter 5 Minuten	5-15 min	15-30 min	30-60 min	über 60 min	Ganze Schicht (8h)
... ohne Pause?						
... pro Tag mit Pausen?						

Gestensteuerung und Interaktion

9. Wie haben Sie die Gestenkontrolle empfunden?

10. Wie würden Sie die Ergonomie der Gestensteuerung insgesamt beurteilen?

☐ sehr gut ☐ gut ☐ mittel ☐ schlecht ☐ sehr schlecht

11. Würden Sie die Form der Gestensteuerung als Interaktion mit dem System bevorzugen?

☐ JA Warum: _____ ☐ Nein Warum: _____

12. Wie haben Sie die Informationsführung (also die textliche oder grafische Anzeige der Matten-Positionen) empfunden?

Anhang 5: Fragebogen Experiment 3 – TABLET-Anweisung/DYPRO-In-situ-Projektion

Fragebogens zur Bewertung des Assistenzsystems

Datenschutzinformation

Die Befragung läuft anonym. Falls Sie für Rückfragen freiwillig zur Verfügung stehen möchten, ergänzen Sie bitte Ihre Kontaktdaten. Die Daten werden für die Auswertung und für die Weiterentwicklung des bestehenden Projekts verwendet. Interessante Zusammenhänge werden unter Einhaltung der oben angegebenen Punkte wissenschaftlich publiziert. Um Ihre Daten ANONYM über die einzelnen Befragungsrunden verarbeiten zu können, wird von Ihnen ein CODE generiert, der eine anonyme Zuordnung der Fragebögen ermöglicht.

- 1.1. Bitte geben Sie die vierte Ziffer Ihres Geburtsjahres ein: _____ (z.B. 1992)
- 1.2. Bitte geben Sie den ersten Buchstaben Ihres Geburtsortes ein: _____ (z.B. Baden)
- 1.3. Bitte geben Sie den ersten Buchstaben des Vornamens Ihrer Mutter ein: _____ (z.B. Susanne)
- 1.4. Bitte geben Sie den ersten Buchstaben des Vornamens Ihres Vaters ein: _____ (z.B. Günter)

1. Bitte kreuzen Sie an welchen Prototyp sie getestet haben:

- ☐ LEGO-Experiment - TABLET-Anweisung
- ☐ LEGO-Experiment - Projektionssystem

Angaben zur Person und Technologieverwendung:

Geschlecht	<i>O Männlich(m)</i>	<i>O Weiblich(w)</i>	<i>O Divers(div)</i>
Altersgruppe	<i>O</i> Unter 20 <i>O</i> 40 bis 49	<i>O</i> 20 bis 29 <i>O</i> 50 bis 59	<i>O</i> 30 bis 39 <i>O</i> Über 59
Tätigkeit/Position	<i>O</i> Führungskraft	<i>O</i> Arbeiter/Angestellte	<i>O</i> StudentIn <i>O</i> Andere:
Technischer Background (Beruflich/Ausbildung/)	<i>O</i> JA	<i>O</i> NEIN	

Wie häufig nutzen Sie folgende Technologien?

	<i>Täglich</i>	<i>2-9x/Monat</i>	<i>6-12x/Jahr</i>	<i>Einmalig</i>	<i>Nie</i>
Smartphone	☐	☐	☐	☐	☐
Tablet	☐	☐	☐	☐	☐
Augmented Reality (AR) am Smartphone/Tablet	☐	☐	☐	☐	☐
AR Brillen	☐	☐	☐	☐	☐
Virtual Reality (VR)	☐	☐	☐	☐	☐

2. Bitte bewerten Sie Ihre Erfahrung (User-Experience) mit dem Assistenzsystem im Hinblick auf die folgenden Aussagen. (System Usability Score - SUS)

	Ich stimme auf keinen Fall zu	Ich stimme nicht zu	neutral	Ich stimme zu	Ich stimme voll zu
Ich denke, dass ich dieses System häufig nutzen möchte.					
Ich fand das System unnötig komplex.					
Ich fand das System einfach zu bedienen.					
Ich denke, ich würde die Unterstützung einer technischen Fachkraft benötigen, um dieses System benutzen zu können.					
Ich fand, dass die verschiedenen Funktionen in diesem System gut integriert sind.					
Ich fand, dass es in diesem System zu viele Inkonsistenzen gab.					
Ich könnte mir vorstellen, dass die meisten Menschen den Umgang mit diesem System sehr schnell lernen würden					
Ich fand das System sehr umständlich in der Benutzung.					
Ich fühlte mich sehr sicher im Gebrauch mit dem System.					
Ich musste eine Menge Dinge lernen, bevor ich mit dieses System arbeiten konnte.					

3. Bitte bewerten Sie Ihre Erfahrungen mit der Lösung der Aufgabe mit dem Assistenzsystem im Hinblick auf die folgenden Aussagen (NASA-RTLX)

	hoch		mittel		gering
Wie geistig anspruchsvoll war die Aufgabe?					
Wie körperlich anstrengend war die Aufgabe?					
Wie eilig oder gehetzt war das Arbeitstempo?					
Wie erfolgreich waren Sie dabei, das zu tun, was von Ihnen verlangt wurde?					
Wie intensiv mussten Sie arbeiten, um ihr Leistungsniveau zu erreichen?					
Wie unsicher, entmutigend, irritiert, gestresst und genervt waren Sie?					

4. Welche Probleme sind Ihnen bei der Arbeit mit dem Assistenzsystem (Tablet oder Projektionssystem + Gestensteuerung) aufgefallen?

5. Was möchten Sie uns positives bei der Arbeit mit dem Assistenzsystem mitteilen?

6. Wie beurteilen Sie die Gesamt-Ergonomie des Assistenzsystems?
☐ sehr gut ☐ gut ☐ mittel ☐ schlecht ☐ sehr schlecht

7. Wie viele Minuten könnten Sie sich vorstellen, mit dem Assistenzsystem zu arbeiten...

	unter 5 Minuten	5-15 min	15-30 min	30-60 min	über 60 min	Ganze Schicht (8h)
... ohne Pause?						
... pro Tag mit Pausen?						

(Bei Verwendung des Projektionssystems auszufüllen)

Gestensteuerung und Interaktion

8. Wie haben Sie die Gestensteuerung empfunden?

9. Wie würden Sie die Ergonomie der Gestensteuerung insgesamt beurteilen?
☐ sehr gut ☐ gut ☐ mittel ☐ schlecht ☐ sehr schlecht

10. Würden Sie die Form der Gestensteuerung als Interaktion mit dem System bevorzugen?
☐ JA Warum: _____ Nein Warum: _____

11. Wie haben Sie die Informationsführung (Führung zur Entnahmeposition / Führung zur Montageposition) empfunden?