



ARBEITSWISSENSCHAFT
INNOVATIONSMANAGEMENT



Telepräsenzte Fernunterstützung in der Industrie

Angelika C. Bullinger-Hoffmann (Hrsg.)

aw&I – Wissenschaft und Praxis

© Verlag aw&I – Wissenschaft und Praxis 2025
Prof. Angelika C. Bullinger-Hoffmann
Technische Universität Chemnitz
Professur Arbeitswissenschaft und Innovationsmanagement

09107 Chemnitz
Tel.: +49 371 531-23210
Mail: awi@tu-chemnitz.de



Die Forschungs- und Entwicklungsarbeit, die zu den Beiträgen dieses Sammelbandes führten, wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ in den Projekten PraeRI (Förderkennzeichen 02L21B000-4) sowie TeleInteraction-XR (Förderkennzeichen 02L21B550-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut.

Die Verantwortung für den Inhalt der Veröffentlichung liegt bei den Autoren.

Reihe aw&I Report, Heft 3

ISSN: 2566-8633

Das Werk, einschließlich seiner Teile, ist urheberrechtlich geschützt. Jede Verwertung ist ohne Zustimmung des Verlages und des Autors unzulässig. Dies gilt insbesondere für die elektronische oder sonstige Vervielfältigung, Übersetzung, Verbreitung und öffentliche Zugänglichmachung.

Bibliografische Information der Deutschen Nationalbibliothek:

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Bildnachweis Titel: Collage aus Werken von CMC/Oculavis sowie der TU Chemnitz



ARBEITSWISSENSCHAFT
INNOVATIONSMANAGEMENT



Telepräsenzte Fernunterstützung in der Industrie

Angelika C. Bullinger-Hoffmann (Hrsg.)

aw&I – Wissenschaft und Praxis

Vorwort

Spätestens in der COVID-19-Pandemie haben wir realisiert, dass viele berufliche Tätigkeiten aus der Ferne – im Homeoffice und per Telearbeit – möglich sind. Nach der Pandemie blieb Telearbeit ein Teil des neuen Normal: Aktuell arbeiten ca. 25 Prozent der Beschäftigten in Deutschland zumindest anteilig im Homeoffice. Zusätzlich ersetzen mittlerweile viele Unternehmen die bisher üblichen Dienstreisen in großem Stil durch Webmeetings.

Das betrifft aber immer noch vorrangig Beschäftigte, die überwiegend geistig und mithin in Büros arbeiten. Für viele Berufe mit einem hohen Anteil manueller Tätigkeiten und von Interaktionsarbeit – sei es in der Pflege, im Handwerk oder am Montageband – ist eine umfassende Telearbeit schlicht nicht vorstellbar. Aber auch in diesen Gebieten existieren neben den wiederkehrenden, eher ausführenden Tätigkeiten auch Aufgaben, bei denen es auf eine Kombination wissensbasierter, kreativ-intellektueller Arbeit mit manuellen Verrichtungen und dialogischer Interaktion ankommt.

In der Industrie sind das vor allem Nicht-Routinen wie Inbetriebnahmen, Wartungen, Instandsetzungen, arbeitsplatznahe Schulungen und Unterweisungen, Audits oder Umbauten und Erweiterungen. Diese erfordern einerseits diagnostische und Problemlösefähigkeiten in Verbindung mit hochspezifischem Fachwissen, andererseits aber eben auch manuelle Verrichtungen – wie das De- und Remontieren, das Ausführen von Mess- und Prüfprozessen oder Einstellarbeiten. Oft genug ist dazu eine Zusammenarbeit zwischen den Betreibern von Produktionsanlagen mit ihren Ausrüstern sowie industrienahen Dienstleistern nötig. Dazu mussten sich bislang vor allem Mitarbeitende – bisweilen auch Geschäftsführende – der Ausrüster und Dienstleister auf Dienstreisen zu den Anlagenbetreibern begeben.

Vor diesem Hintergrund haben sich die beiden Forschungs- und Entwicklungs-Verbundprojekte

- **PraeRI** (Telepräsenzroboter in der Industrie) und
- **TeleInteraction-XR** (Augmented Reality für die Fernwartung)

mit der Frage beschäftigt, wie neue Technologien – eben Telepräsenzroboter und Augmented/Mixed Reality – dabei helfen können, solche Tätigkeiten in Form einer Fernunterstützung zu erledigen.

Ziel – insbesondere der beteiligten Anwendungspartner – war es dabei, durch technologiebasierte Fernunterstützung Dienstreisen zu ersetzen, um so an Reaktionsschnelligkeit, Zeit- und Kosteneffizienz, aber auch an Arbeitgeberattraktivität zu gewinnen. Gerade für Beschäftigtengruppen wie Monteure, Instandhalter und Projektingenieure verbessert sich die Vereinbarkeit von Beruf und Privatem enorm, wenn weniger Dienstreisen mit langen Abwesenheitszeiten von zu Hause anfallen.

In beiden Projekten wurden dazu zunächst Lösungen, die am Markt für andere Anwendungsgebiete bzw. für den privaten Konsum verfügbar sind, im industriellen Umfeld erprobt; Anwendungspotenziale und Entwicklungsbedarfe wurden identifiziert; für den Industrieinsatz nötige Weiterentwicklungen wurden beispielhaft durchgeführt und als Demonstratoren umgesetzt.

Die Ergebnisse weisen die technische Machbarkeit der Fernunterstützung mit Telepräsenzrobotern und Mixed Reality in der Industrie grundsätzlich nach. Sie zeugen auch von einer oft bereits guten Akzeptanz durch die Beschäftigten. Zu den bleibenden Herausforderungen zählen vor allem die Verfügbarkeit stabiler, Breitband-Internetverbindungen, Fragen von Datensicherheit und Datenschutz sowie Abhängigkeiten von den Geschäftsmodellen der Technologiegeber.

Zu all diesen Aspekten berichten die Projektbeteiligten in diesem Sammelband nun selbst.

Als Inhaberin der Professur für Arbeitswissenschaft und Innovationsmanagement, die an beiden Projekten beteiligt war, bedanke ich mich an dieser Stelle sehr herzlich bei den Projektpartnern Actimage GmbH, CMC-Kiesel GmbH, HLS Robotik GmbH, IMSYS Immersive Systeme GmbH & Co. KG, LiGenium GmbH, LS Software Engineering GmbH, oculavis GmbH und ULT AG für die vertrauensvolle und konstruktive Zusammenarbeit. Den Autorinnen und Autoren, die gleichzeitig auch die maßgeblichen Bearbeitenden in den Projekten waren, danke ich auch persönlich für ihr Engagement im jeweiligen Projekt und für ihre wertvollen Beiträge zu dieser Publikation.

Im Namen beider Projektkonsortien bedanke ich mich beim Bundesministerium für Bildung und Forschung, das die Projekte im Programm „Zukunft der Wertschöpfung – Forschung zu Produktion, Dienstleistung und Arbeit“ gefördert hat.

Dem Projektträger Karlsruhe (PTKA), namentlich Frau Dr. Spangenberg und Frau Schulz, danke ich für die ausgezeichnete Betreuung.

Den Leserinnen und Lesern wünsche ich spannende Einblicke und eine Ermutigung, die Arbeitswelt von morgen aktiv mitzugestalten. Denn das wurde in beiden Projekten getan: Neues ausprobieren, nutzen und verbessern!

Chemnitz, März 2025

Angelika C. Bullinger-Hoffmann

Herausgeberin



Angelika C. Bullinger Hoffmann

Prof. Dr. Bullinger-Hoffmann leitet seit April 2012 die Professur Arbeitswissenschaft und Innovationsmanagement der TU Chemnitz. Sie forscht zur menschenzentrierten Gestaltung und Evaluation von Mensch-Maschine-KI-Systemen, mit dem Ziel individueller und gesellschaftlicher Akzeptanz und unternehmerischer Produktivitätssteigerung. Sie hat langjährige Erfahrung in der Einwerbung und Leitung von nationalen und europäischen Projekten, berät Industrieunternehmen und hält regelmäßig Vorträge und Keynotes zur Zukunft von Arbeit und Arbeitenden. Prof. Bullinger-Hoffmann ist Mitglied der Deutschen Akademie für Technikwissenschaften - acatech und Aufsichtsrätin der Paul Hartmann AG. Ergebnisse ihrer Arbeiten sind in Zeitschriften wie Ergonomics, Human Factors, Creativity and Innovation Management, R&D Management, sowie der Zeitschrift für Arbeitswissenschaft publiziert.

Inhaltsverzeichnis

So fern und doch so nah – Telepräsenes Arbeiten in der Industrie Danny Ruffert, Alexander Kögel, Thomas Löffler	1
Die Qual der Wahl – Welcher Telepräsenzroboter ist der Richtige? Francisco Hernandez, Danny Ruffert, Thomas Löffler	11
Ich brauche Unterstützung – Kontaktaufnahme über TPR- Anruffunktion Francisco Hernandez, Danny Ruffert	19
Bitte folge mir – Steuerung eines TPR über Gesten Florian Schnabel	27
Kannst Du mir das zeigen? Entwicklung und Evaluation einer Zeigefunktion für einen TPR Rakshith Venepally, Francisco Hernandez, Danny Ruffert	32
Erweiterung der Softwarefunktionalität von TPR zum Einsatz in der Industrie Sven Eichhorn, Georg Schmidt, Stefan Stüring, Rakshith Venepally	39
Roboter öffnen Türen – Neue Wege der Telepräsenz Nikolas von Lonski	47
Chancen und Grenzen von Telepräsenzrobotern in der Industrie aus Anwendersicht Sven Eichhorn, Florian Schnabel, Tom Heinicker, Georg Rudolph	53
Wer und was ist wo? – Synchronisation von Virtueller und Erweiterter Realität Alexander Kögel	63
Ich habe etwas vorbereitet – 3D-Modell-basierte AR/VR-Kopplung Markus Kiesel, Alexander Kögel	74
Kamera, Licht, Action! - Videostream-basierte AR/VR Kollaboration Daniel Wulf-Misaki	84

Mehr als oberflächlich! – Live Volumetric Streamings für die AR/VR-Kollaboration. Markus Kiesel, Daniel Wulf-Misaki	93
TeleInterACTION! – VR/AR-Kopplung im Praxistest Madeleine Berger, Nicole Holz, Danny Rüffert, Alexander Kögel	104

So fern und doch so nah – Telepräsenztes Arbeiten in der Industrie

Danny Rüffert¹, Alexander Kögel¹, Thomas Löffler¹

¹ Professur Arbeitswissenschaft und Innovationsmanagement, TU Chemnitz

Zusammenfassung

Nicht physisch vor Ort, aber dennoch präsent zu sein, beschreibt den Kontext der Telepräsenz. In diesem Beitrag werden zwei differierende technologische Ansätze der Telepräsenz näher vorgestellt. Im Projekt PraeRI wurde dazu ein Telepräsenzroboter (TPR) verwendet, wogegen das Projekt TeleInteraction-XR auf der Nutzung von virtuellen und erweiterten Realitäten aufbaut. Beide Projekte adressieren das telepräsenztes Arbeiten im industriellen Feld.

1 Motivation

Von 2009 bis 2019 stieg die Zahl der Geschäftsreisen in deutschen Unternehmen um 24 Prozent auf 195,4 Millionen pro Jahr. Diese Zahl sank während der Corona-Pandemie auf 32,7 Millionen, wuchs dann aber wieder auf 116,7 Millionen Reisen im Jahr 2023 an [1]. Gleichzeitig sind die Kosten pro Geschäftsreise innerhalb von 5 Jahren um ca. 100 € auf durchschnittlich 409 € in 2023 gestiegen [2]. Zu diesem Betrag müssen noch die Personalkosten für die Reisezeiten addiert werden, so dass die mit Geschäftsreisen verbundenen Kosten einen erheblichen Faktor bilden. Geschäftsreisen strapazieren zudem die Vereinbarkeit von Beruf und Privatleben und tragen zu CO₂-Emissionen bei. Wie die Corona-Pandemie zeigte, sind Alternativen zu Geschäftsreisen – und damit zu Präsenzterminen – nötig und möglich [3].

Dafür haben sich in erster Linie Videobesprechungen mit PCs, Notebooks und Smartphones etabliert. Sie sind in Büroumgebungen gut geeignet, um „telepräsenztes“ an einem entfernten Ort an Meetings oder bilateralen Besprechungen teilzunehmen.

Videokonferenzen zeigen allerdings Schwächen im industriellen Kontext – etwa, um Besprechungen direkt an Maschinen, technischen Anlagen oder Produkten abzuhalten:

- Das Sichtfeld der PC-, Notebook-, Tablet- und Smartphone-Kameras ist eingeschränkt.
- Besprechungsteilnehmer können sich an den jeweils entfernten Ort nicht frei im Raum bewegen bzw. nicht einmal umsehen.
- Es fehlt es an Interaktionsmöglichkeiten, die über das reine Sehen und Hören hinausgehen, insbesondere an der Möglichkeit, durch Gesten auf etwas zu zeigen.
- In vielen Situationen (z. B. bei Wartungs- oder Inbetriebnahme-Arbeiten) sind „freie Hände“ gefragt, so dass Notebook, Tablet oder Smartphone nicht gehalten werden können.

Vor diesem Hintergrund wurden in zwei vom BMBF geförderten Projekten zwei unterschiedliche Technologien erprobt und für den Einsatz in der Industrie weiterentwickelt. Dabei wurde auch berücksichtigt, dass Industrieservices wie Inbetriebnahme, Wartung und Schadensdetektion war bei und teilweise von den Betreibern ausgeführt werden, aber oft Expertenwissen der Hersteller der Ausrüstungen verlangen. Der Einsatz hochqualifizierter Fachkräfte der Hersteller vor Ort beim Kunden ist aber schon aus logistischen und Kapazitätsgründen nicht immer (sofort) möglich. Die telepräsente Fernunterstützung ist auch daher eine attraktive Alternative.

2 Zwei Lösungsansätze – zwei Projekte

2.1 PraeRI – Präsenzroboter für Industrie

Das Projekt PraeRI hatte zum Ziel, Potenziale, Hemmnisse und Voraussetzungen des Einsatzes von Telepräsenzrobotern in der Industrie zu erforschen und Demonstrator-Lösungen für deren Etablierung zu entwickeln.

Ein Telepräsenzroboter (TPR) ist ein fahrbares, über Fernzugriff steuerbares Videokonferenzsystem. Charakteristisch für die Nutzung von TPR ist, dass sich TPR-Fernanwender, die sich an einem Ort 1 (remote) befinden, mittels Fernsteuerung über das Internet autark in der Umgebung, in der sich der Roboter befindet – dem Ort 2 –, bewegen und dort mit Personen interagieren können [4-6]. Grundsätzlich bestehen TPR-Systeme aus dem vor Ort agierenden Roboter (Sensoren, Aktoren, Bildschirm) und der remote bedienten Steuerungs-Schnittstelle. Abbildung 1 zeigt das Funktionsprinzip der Kommunikation über TPR.

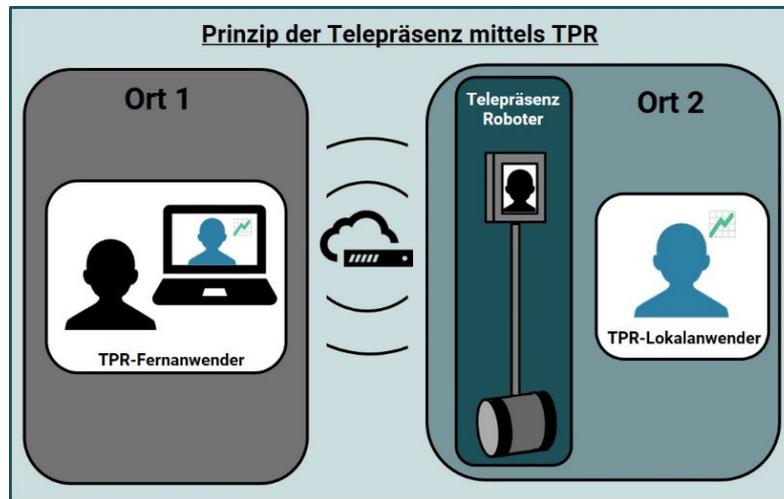


Abbildung 1: Prinzip der Telepräsenz mittels TPR

Die Bewegungsfreiheit des TPR bietet dem Fernanwender die Möglichkeit, die Position im Raum und das Sichtfeld beliebig zu verändern, was bei einem herkömmlichen Videoanruf nicht erreicht werden kann. Die Kommunikation wird durch die scheinbar physische Anwesenheit des Fernanwenders über das auf dem Bildschirm des TPR angezeigte Bewegtbild, über den Ton und die Bewegung des TPR als natürlicher wahrgenommen und steigert somit das Präsenzepfinden [4]. Die Person, die mit dem TPR am Ort 2 interagiert, muss keine zusätzliche Ausrüstung mit sich führen und kann somit freihändig arbeiten.

Die Einsätze von TPR sind in der Vergangenheit mehr oder weniger erfolgreich in den Bereichen Büro, Gesundheitswesen, Pflege sowie Lehre und Forschung getestet worden. Anwendungsfälle waren:

- virtuelle Teilnahmen an Konferenzen und Besprechungen [4, 7],
- virtuelle Beratungen mit medizinischem Personal, um die Kommunikation mit isolierten Patienten bzw. Bewohnern von Altersheimen zu erleichtern [5, 8-10],
- virtuelle Treffen von Lehr- und Lernenden [11, 12].
- virtuelle Teilhabe von Menschen mit eingeschränkter Mobilität [13].

Als spezifische Anwendungsfelder im industriellen Feld eruierten die Projektpartner – auf Basis erster Tests verschiedener TPR-Modelle – folgende Anwendungsfälle:

- Projektbesprechungen an der Anlage,
- Einweisungen, Abnahmen und Unterstützung bei Inbetriebnahmen von Anlagen/Maschinen,
- Anlagenbegehung,
- Qualitätskontrollen,
- Wartungen.

Projektpartner waren Actimage GmbH (Entwicklung von Informations- und Kommunikationstechnologie), HLS Robotik GmbH (Engineering von Produktionsanlagen), LiGenium GmbH (Hersteller von Ladungsträgern, Fördertechnik und Leichtbauprofilen), LS Software Engineering GmbH (Entwicklung von Visualisierungs- und Simulationssoftware) sowie die TU Chemnitz, Professur Arbeitswissenschaft und Innovationsmanagement.



Abbildung 2: Qualitätskontrolle Großladungsträger mit TPR Ohmni Pro

Basierend auf diesen Anwendungsfällen wurden nutzungsspezifische Anforderungen identifiziert, die innerhalb der Projektlaufzeit umgesetzt und evaluiert werden sollten. Tabelle 1 zeigt eine Zusammenfassung wichtiger Anforderungen, um einen handelsüblichen TPR für das Niveau „Industry Ready“ zu befähigen.

Tabelle 1: Weiterentwicklungsbedarfe für TPR

Anforderung	Beschreibung
Zeigefunktion	Die Integration einer Zeigefunktion ermöglicht eine intuitive und effiziente Kommunikation zwischen den Nutzern.
Anruftaste	Ein Anruf kann bisher nur über den Fernanwender initiiert werden. Mit einer Anruftaste kann der Lokalanwender den Fernanwender direkt kontaktieren und somit schneller Unterstützung anfordern.
Gebäudekommunikationsschnittstelle	Die Interaktion mit der Gebäudeautomation (Licht, Türen, Fahrstuhl) erhöht die Unabhängigkeit zu physisch echtem Personal vor Ort.
Gestensteuerung	Bietet die Möglichkeit, dass auch der Lokalanwender den TPR steuern kann. Anwendungsfälle finden sich in Gefahrensituationen oder Arealen mit Geheimhaltung.
Alternative Internetverbindung	Ein TPR sollte unabhängig von lokalen Internetverbindungen sein. Firmeninterne Netzwerke sind aufgrund von Datenschutz und Datengeschwindigkeit oftmals nur begrenzt nutzbar.
Headset Anbindung	Ein Headset verbessert die Kommunikation und den Datenschutz zwischen TPR-Lokal- und Fernanwender
Umgebung abdecken/ausblenden	Produktionslinien oder Ausrüstungen die der Geheimhaltung unterliegen, können bei Bedarf für den Fernanwender ausgeblendet werden.
Digitale Karte	Mittels einer digitalen Karte kann der TPR autonom zum vereinbarten Treffpunkt fahren, ohne dass der Fernanwender am unbekanntem Ort navigieren muss.
Digitale Ausweisfunktion	Um Missbrauch des Video-Feedback-Systems zu verhindern, sollte für den TPR-Fernanwender ein Besucherausweis erstellt werden. So kann der TPR fernbetrieben werden, als ob TPR-Fernanwender vor Ort wäre.
Gesichtserkennung	Benutzeridentifikation vor Ort bei größeren Gesprächsrunden und bei für Fernanwender unbekanntem Lokalanwendern.
Online QM-Protokoll	Eine App für ein QM-Protokoll mit dem Fern- und Lokalanwender gemeinsam Qualitätsmaßnahmen durchführen, am TPR protokollieren und signieren können.

Im vorliegenden Sammelband wird eine Methode zur TPR-Auswahl vorgestellt [14]. Weitere Beiträge beschreiben die prototypischen Entwicklungen einer Anruftaste [15], einer Gestensteuerung [16], einer Zeigefunktion mittels Laserpointer [17] und einer Ankopplung an die Gebäudesteuerung [18]. Außerdem werden wünschenswerte Erweiterungen der Softwarefunktionalität [19] sowie Chancen und Grenzen des TPR-Einsatzes aus Anwendersicht [20] thematisiert.

2.2 TeleInteraction-XR - Augmented Reality für die Fernwartung

Auch Technologien der Virtual und Augmented Reality (VR/AR) bieten das Potenzial, dass die Expertinnen und Experten telepräsente Unterstützung aus der Ferne leisten und so Reisen zu physischen Präsenzterminen vermeiden.

Ziel des Projekts TeleInteraction-XR war die Entwicklung eines Demonstrators, der im Kontext komplexer Service- und Trainingsaufgaben eine möglichst natürliche kollaborative Mensch-Mensch-Interaktion zwischen einem aus der Ferne unterstützenden Experten oder einer Expertin und einer vor Ort agierenden Person ermöglicht.

Dazu erfolgte eine Weiterentwicklung und Integration von bisher separat genutzten Systemen der Entwicklungspartner CMC-Kiesel (Multi-User VR-Viewer) und Oculavis GmbH (AR-Support) in ein interaktives Gesamtsystem. Verantwortlich für die Anforderungsanalysen, Erprobungen und Evaluationen waren die Anwendungspartnern ULT AG (Herstellung von lufttechnischen Anlagen) und IMSYS Immersive Systeme GmbH & Co. KG (Beratung, Planung und Projektierung von digitalen Arbeitsumgebungen) begleitet von der Professur Arbeitswissenschaft und Innovationsmanagement der TU Chemnitz.

Technisch basiert die Kopplung von VR und AR auf einem identischen Bezugssystem, in dem reale und digitale Informationen in Echtzeit übermittelt werden (siehe Abb. 1). So können Experten von einem beliebigen Ort die ortsferne reale Feldumgebung mit immersiven Visualisierungskonzepten in VR erkunden. Gleichzeitig werden dem Nutzenden vor Ort mittels Augmented Reality in 3D repräsentierte Handlungsanweisungen des Experten oder der Expertin (bspw. in Form digitalisierter Handanleitungen) in Echtzeit und räumlich korrekt visualisiert.

Beide Personen sind räumlich getrennt, aber jeweils virtuell präsent. Der vorhandene Sprachkanal über die bestehende Multi-User-Anwendung ergänzt die visuellen Rückmeldungen und ermöglicht eine multimodale Interaktion. Dadurch kann der Experte bzw. die Expertin trotz physischer Distanz als „Third-person“ in natürlicher Weise „hinter“ der Person im Feld über deren „Schulter schauen“ oder diese als „first-person“ begleiten.

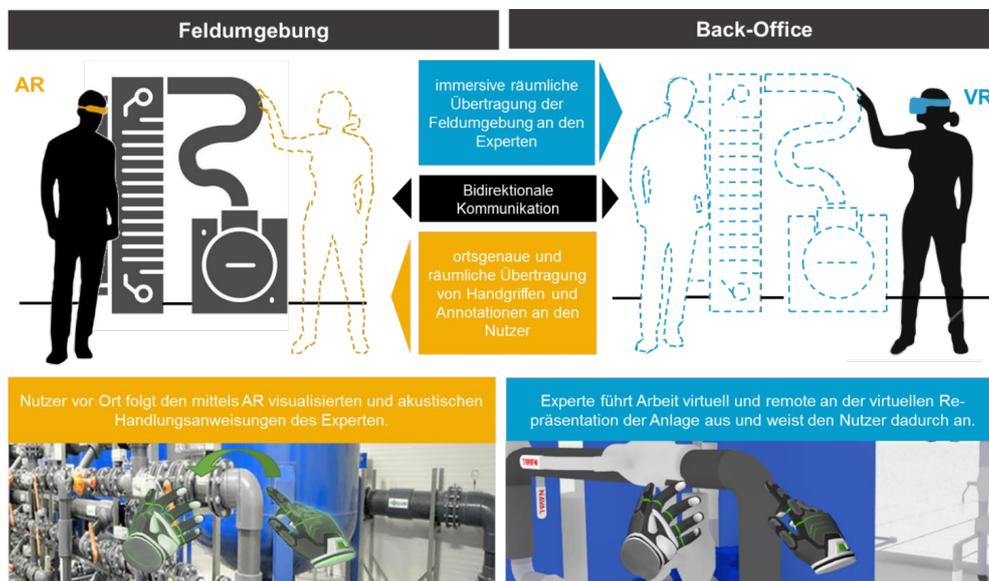


Abbildung 3: Fernunterstützung durch VR/AR-Kollaboration

Zur Umsetzung der VR/AR-Kollaboration wurden folgende drei aufeinander aufbauende Entwicklungs- bzw. Integrationsschritte durchgeführt, die im weiteren Verlauf des Sammelbands genauer beschrieben werden:

1. Vorab erstellte 3D-Modelle [21],
2. Live Videostream [22] und
3. Live Volumetric Streaming [23].

Grundlagen der VR/AR-Kollaboration sind in [24] und praktische Erfahrungen in [25] dargelegt.

3 Zusammenfassung

Sowohl der Einsatz von TPR als auch die VR/AR-Kollaboration zeigen ein hohes Potenzial, Industrieservices durch telepräsenste Expertinnen und Experten aus der Ferne zu unterstützen. Dadurch können der Ressourceneinsatz für vor Ort-Besuche reduziert, die Reaktionsschnelligkeit erhöht und die Attraktivität der Arbeit erhöht werden. Praktische Limitationen und Herausforderungen werden besonders in den Beiträgen [20, 25] erörtert.

Die Forschungs- und Entwicklungsarbeit wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und

der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ in den Projekten PraeRI (Förderkennzeichen 02L21B000-4) sowie TeleInteraction-XR (Förderkennzeichen 02L21B550-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt der Veröffentlichung liegt bei den Autoren.

4 Literaturverzeichnis

- [1] Verband Deutsches Reisemanagement. (6. August, 2024). Anzahl der Geschäftsreisen von deutschen Unternehmen in den Jahren von 2004 bis 2023 (in Millionen) [Graph]. In Statista. Zugriff am 13. März 2025, von <https://de.statista.com/statistik/daten/studie/72112/umfrage/anzahl-der-geschaeftsreisen-seit-2004/>
- [2] Verband Deutsches Reisemanagement. (6. August, 2024). Durchschnittliche Kosten pro Geschäftsreise in deutschen Unternehmen von 2005 bis 2023 [Graph]. In Statista. Zugriff am 13. März 2025, von <https://de.statista.com/statistik/daten/studie/72171/umfrage/tourismus-kosten-pro-geschaeftsreise-seit-2005/>
- [3] Hoffmann, A. & Wintermann, O. (2020). ZUKUNFTSSTUDIE MÜNCHNER KREIS. Sonderstudie zur Corona-Pandemie. Bertelsmann Stiftung. Online verfügbar unter <https://www.bertelsmann-stiftung.de/de/publikationen/publikation/did/sonderstudie-corona-all>, zuletzt geprüft am 13.03.2025
- [4] Björnftot, P., Bergqvist, J., & Kaptelinin, V. (2018). Non-technical users' first encounters with a robotic telepresence technology: An empirical study of office workers. In: Paladyn, Journal of Behavioral Robotics 9 (1), S. 307–322. DOI: 10.1515/pjbr-2018-0022.
- [5] Kehl, C. (2018). Robotik und assistive Neurotechnologien in der Pflege - gesellschaftliche Herausforderungen. Vertiefung des Projekts »Mensch-Maschine-Entgrenzung«. Büro für Technikfolgen-Abschätzung beim Deutschen Bundestag (TAB). Online verfügbar unter <https://www.tab-beim-bundestag.de/de/pdf/publikationen/berichte/TAB-Arbeitsbericht-ab177.pdf>, zuletzt geprüft am 13.03.2025.
- [6] Kristoffersson, A., Coradeschi, S., & Loutfi, A. (2013). A Review of Mobile Robotic Telepresence. In: Advances in Human-Computer Interaction 2013 (5), S. 1–17. DOI: 10.1155/2013/902316.
- [7] Lee, M. K., & Takayama, L. (2011). Now, I have a body. In: Desney Tan, Geraldine Fitzpatrick, Carl Gutwin, Bo Begole und Wendy A. Kellogg (Hg.): Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11. the 2011 annual conference. Vancouver, BC, Canada, 07.05.2011 - 12.05.2011. New York, New York, USA: ACM Press, S. 33.
- [8] Lee, WH., Park, J., & Park., C. H. (2018). Acceptability of Tele-assistive Robotic Nurse for Human-Robot Collaboration in Medical Environment. In: Kanda, Šabanović et al. (Hg.) 2018 – Companion of the 2018 ACM/IEEE, S. 171–172.
- [9] Valérie, J. (2019). Einsatz von Telepräsenzrobotern für ärztliche Konsultationen. Online verfügbar unter <https://irf.fhnw.ch/bitstream/handle/11654/30275/Jungo%20Val%C3%A9rie.pdf?sequence=1>, zuletzt geprüft am 13.03.2025.

- [10] Tischler, N. W. (2017). „Assistive Social Robots“ für Personen mit Demenz in der institutionellen Pflege. Online verfügbar unter https://online.medunigraz.at/mug_online/wbAbs.showMaskAbsBe-treuer?pOrgNr=1&pPersNr=67508&pSort=4%3B2%3B6&pPageNr=1, zuletzt geprüft am 22.07.2020.
- [11] Puarungroj, W., Boonsirisumpun, N. (2020). Multiple Device Controlled Design for Implementing Telepresence Robot in Schools. In: Simon K. S. Cheung, Richard Li, Kongkiti Phusavat, Naraphorn Paoprasert und Lam-For Kwok (Hg.): Blended Learning. Education in a Smart Learning Environment. Cham: Springer International Publishing (12218), S. 405–415.
- [12] Gallon, L. (2014). Immersion dans un TP en téléprésence.
- [13] Herring, S. (2013). Telepresence robots for academics. DOI: 10.1002/meet.14505001156.
- [14] Hernandez, F., Rüffert, D., Löffler, T. (2025). Die Qual der Wahl – Welcher Telepräsenzroboter ist der Richtige? In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [15] Hernandez, F., Rüffert, D. (2025). Ich brauche Unterstützung – Kontaktaufnahme über TPR-Anruhfunktion. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [16] Schnabel, F. (2025). Bitte folge mir – Steuerung eines TPR über Gesten. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [17] Venepally, R., Hernandez, F., Rüffert, D. (2025). Kannst Du mir das zeigen? Entwicklung und Evaluation einer Zeigefunktion für einen TPR. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [18] von Lonski, N. (2025). Roboter öffnen Türen – Neue Wege der Telepräsenz. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis. DOI
- [19] Eichhorn, E., Schmidt, G., Stüring, S., Venepally, R. (2025). Erweiterung der Softwarefunktionalität von TPR zum Einsatz in der Industrie. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [20] Eichhorn, E., Schnabel, F., Heinicker, T., Rudolph, G. (2025). Chancen und Grenzen von Telepräsenzrobotern in der Industrie aus Anwendersicht. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [21] Kiesel, M., Kögel, A. (2025). Ich habe etwas vorbereitet – 3D-Modell-basierte AR-/VR-Kopplung. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [22] Wulf-Misaki, D. (2025). Kamera, Licht, Action! - Videostream-basierte AR/VR Kollaboration. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [23] Kiesel, M., Wulf-Misaki, D. (2025). Mehr als oberflächlich! – Live Volumetric Streamings für die AR/VR-Kollaboration. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.

- [24] Kögel, A. (2025). Wer und was ist wo? - Synchronisation von Virtueller und Erweiterter Realität. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenste Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [24] Berger, M., Holz, N., Rüffert, D., Kögel, A. (2025). TeleInterACTION! – VR/AR-Kopplung im Praxistest. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenste Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.

Autoren



Rüffert, Danny

Danny Rüffert studierte Sports Engineering an der Technischen Universität Chemnitz. Seit 2014 arbeitet er als wissenschaftlicher Mitarbeiter an der Professur Arbeitswissenschaft und Innovationsmanagement im Cluster Industrial Engineering. Seine Aufgaben umfassen die Themenfelder Alterssimulation und Ergonomie sowie die Forschung an Mensch-Maschine-Schnittstellen.



Kögel, Alexander

Alexander Kögel studierte Medientechnik (B.Eng.) und Medienmanagement (M.Eng.) an der HTWK Leipzig. Seit 2017 ist er an der TU Chemnitz als wissenschaftlicher Mitarbeiter tätig. Schwerpunkte sind VR-Anwendungen, das Präsenz-Empfinden in virtuellen Umgebungen und der Erschließung der Potenziale der Digitalisierung, Vernetzung und Industrie 4.0.



Löffler, Thomas

Dr.-Ing. Thomas Löffler studierte Maschinenbau an der Technischen Universität Chemnitz. Er arbeitete anschließend als wissenschaftlicher Mitarbeiter an der Professur Fabrikplanung und Fabrikbetrieb, als Werkstrukturplaner bei bse engineering Leipzig GmbH sowie als Geschäftsführer des Instituts IREGIA e. V. Seit 2011 ist er an der Professur Arbeitswissenschaft und Innovationsmanagement als Clusterleiter Industrial Engineering tätig.

Die Qual der Wahl – Welcher Telepräsenzroboter ist der Richtige?

Francisco Hernandez¹, Danny Ruffert¹, Thomas Löffler¹

¹TU Chemnitz – Professur Arbeitswissenschaft und Innovationsmanagement

Zusammenfassung

Der Beitrag beschreibt eine strukturierte Vorgehensweise zur Bewertung der Eignung von Telepräsenzrobotern (TPR) für den industriellen Einsatz. Mithilfe einer Nutzwertanalyse und eines strukturierten Bewertungsprozesses wurden sechs vorausgewählte TPR-Modelle bezüglich 28 Kriterien in fünf übergeordneten Kategorien untersucht. Für die Anforderungen im Projekt PraeRI zeigten die Modelle Double 3 und Ohmni Pro die höchsten Nutzwerte. Der Bewertungsrahmen kann als praxisnahe Entscheidungshilfe für die Auswahl von TPR weitergenutzt werden.

1 Einleitung

Telepräsenzroboter (TPR) sind in der Industrie weniger verbreitet als in Büros, Schulen oder im Gesundheitswesen [1-5]. Daher ist es von entscheidender Bedeutung, zu bestimmen, welche TPR für industrielle Anforderungen geeignet sind. Aktuell existieren jedoch weder für die spezifischen Einsatzbedingungen in der Industrie noch allgemein für Leistungsparameter von TPR standardisierte Test- und Beurteilungsverfahren. Vorstellbar wären etwa standardisierte Nutzungszyklen, um die Betriebsdauer mit einer Batterieladung zu ermitteln.

Das Konsortium des Projekts PraeRI umfasste nicht den nötigen Kreis von Partnern, um allgemeinverbindliche Standards für die Bewertung von TPR zu entwickeln. Jedoch wurden aus Perspektive der am Projekt beteiligten Anwendungs-, Entwicklungs- und Wissenschaftspartnern eine Vorgehensweise und ein Bewertungsrahmen geschaffen. Beides diente zunächst dazu, geeignete TPR für die Anwendungsfälle im Projekt PraeRI [6] auszuwählen, kann aber orientierend auch darüber hinaus nachgenutzt werden.

2 Vorgehensweise

Als methodischer Ansatz der Bewertung diente die Nutzwertanalyse. Bei der Auswahl der Bewertungskriterien wurde initial auf Vorarbeiten von Hernandez et al. [7] zu industriellen Anforderungen an TPR zurück gegriffen. Folgende Vorgehensschritte wurden durchlaufen:

2.1 Bildung eines Bewertungsteams

Im Rahmen des Projektes wurde ein zwölfköpfiges Bewertungsteams gebildet. Es umfasste potenzielle Anwendende, Entscheider und Wissenschaftler.

2.2 Analyse von Basisanforderungen und Marktanalyse

Ausgehend von Basisanforderungen wie der Verfügbarkeit im deutschen Markt und dem Vorhandensein einer Fahrfunktion – in Schulen kommen TPR zum Einsatz, die getragen werden müssen – wurden sechs TPR-Modelle zur Bewertung ausgewählt (siehe Abbildung 1).



Abbildung 1: Untersuchte Roboter: (a) GoBe Robot von Blue Ocean Robotics – Denmark [8], (b) UBBO-Expert von Axyn Robotics – France [9], (c) Ava von Ava Robotics – USA [10], (d) Temi von Temi Robots – USA [11], (e) Double3 von Double Robotics – USA [12], (f) Ohmni Pro von Ohmnilabs – USA [13].

2.3 Eruiere von Datenquellen und Testaktivitäten

Um im nächsten Schritt realistische Bewertungskriterien festlegen zu können, wurden zunächst die verfügbaren Daten und Erprobungsmöglichkeiten der TPR eruiert. Prinzipiell standen Datenblätter und technische Beschreibungen der Hersteller zur Verfügung. In der Regel konnten auch die Benutzeroberflächen online beurteilt werden. Einige Modelle standen für Remote-Tests zur Verfügung. Drei Modelle Double3, Ohmni und Temi waren bereits beim Partner TU Chemnitz im Einsatz.

2.4 Festlegung der Bewertungskriterien

Insgesamt wurden 28 einzelne Bewertungskriterien identifiziert, die zu fünf Kategorien zusammengefasst wurden. Diese bilden den Bewertungsrahmen und die Grundlage für die Nutzwertanalyse.

1. **Kosten:** Im Projekt wurden vereinfachend allein die Anschaffungskosten verglichen. Je nach Geschäftsmodell der Hersteller können bestimmte Kosten für Hard- und besonders für Software-Optionen als Anschaffungskosten oder Betriebskosten anfallen (z. B. Freischaltung Entwicklermodus, Lizenzen für Flottenmanagement-Software und Inhouse-Hosting). Diese sollten dann in einer Investitionsrechnung (z. B. Kapitalwertmethode) einfließen.
2. **Administrativ-koordinativer Aufwand:** Bestimmte Aufwände zur Sicherstellung der Gesetzeskonformität, bei der Beschaffung, Wartung und von Weiterentwicklungsmöglichkeiten lassen sich nicht rational-zuverlässig in Kosten ausdrücken, spielen aber eine Rolle. Im Projekt beurteilte das Bewertungsteam folgende Kriterien: Einhaltung der DSGVO, Lieferzeit, Niederlassung in Deutschland, Support des Herstellers, Zugänglichkeit Entwicklermodus.
3. **Technische Basisleistungsfähigkeit:** Innerhalb dieser Kategorie wurden folgende technischen Parameter verglichen: Akkulaufzeit bei aktiver Videoübertragung und im passiven Zustand, Ladezeit, Gewicht des TPR, Stabilität/Kippfestigkeit, Transportierbarkeit.
4. **Benutzerfreundlichkeit:** In dieser Kategorie wurden anhand der Erprobung und Bemusterung folgende Kriterien beurteilt: Übersichtlichkeit der Benutzeroberfläche, Intuitivität der Steuerung der Hauptfunktionen (Fahren, Video-Call), Intuitivität der Bedienung von weiteren Funktionen (z. B. Bildschirmteilen), Reaktionszeit/Verzögerung, Konnektivität

(Webbrowser, Verbindungsmöglichkeit zu Mobiltelefon, Einfachheit der Einladung von Video-Call-Teilnehmenden).

5. **Zusatzfunktionalitäten:** In dieser Kategorie erfolgte eine Bewertung von Kriterien/Merkmalen, die über die Basisfunktionalität hinaus gehen oder die die Einsatzmöglichkeiten und den Komfort erhöhen. Die Bewertung umfasste: Konnektivität (Wi-Fi, Bluetooth), Bildschirmgröße, Höhenverstellbarkeit (motorisch-ferngesteuert, manuell), Kamera (zusätzliche Frontkamera, Sichtfeld, zusätzliche Navigationskamera), Verbreitung des Betriebssystems, Umfang und Qualität der Sensorik, Sonstige Extras (z. B. 2D Mapping, 3D Mapping, Follow-me-Funktion).

2.5 Durchführung der TPR-Erprobung und Datenakquise

Mitglieder des Bewertungsteams erprobten die sechs TPR und dokumentierten Nutzungserfahrungen. In den verfügbaren Datenblättern nicht enthaltene Angaben wurden teilweise nacherhoben.

2.6 Bewertungs-Workshop

Nach den Erprobungen, der Datenakquise und der Zuordnung von Daten und Informationen zu den Bewertungskriterien fand ein Bewertungs-Workshop mit dem gesamten Bewertungsteam statt. Im Workshop wurde eine Nutzwertanalyse für die ausgewählten TPR-Modelle erstellt. Dazu wurden vom Bewertungsteam

- (1) die Bewertungskategorien und die einzelnen Bewertungskriterien untereinander gewichtet und
- (2) jedes Bewertungskriterium für jedes TPR-Modell bewertet.

Die Bewertung erfolgte mittels vierstufiger Punkteskala nach VDI 2225 Blatt 3 [14] (0 unbefriedigend, 1 gerade noch akzeptabel, 2 ausreichend, 3 gut, 4 sehr gut). Anschließend wurde für jedes Bewertungskriterium der Teilnutzwert berechnet ($\text{Gewichtung} \times \text{Bewertung}$). Die Summe aller Teilnutzwerte eines TPR gibt seinen Nutzwert an.

3 Ergebnisse

Bewertungskategorien wurden wie folgt gewichtet:

- Kosten: 20 Prozent
- Administrativ-koordinativer Aufwand: 15 Prozent
- Technische Basisleistungsfähigkeit: 15 Prozent

- Benutzerfreundlichkeit: 20 Prozent
- Zusatzfunktionalitäten: 30 Prozent

Tabelle 1 zeigt für jedes TPR-Modell die kumulierten Teilnutzwerte auf Ebene der Bewertungskategorien, den Nutzwert und den Zielerreichungsgrad in Prozent. Letzterer entspricht dem Verhältnis von Nutzwert zum theoretisch maximal erreichbaren Nutzwert 4.

Tabelle 1: Ergebnisse der Nutzwertanalyse

Kategorie	GoBe Robot	UBBO Expert	AVA	Tem 3	Double 3	Ohmni Pro
Kosten	0,4	0,6	0,0	0,8	0,6	0,6
Administrativ-koordinativer Aufwand	0,5	0,5	0,5	0,5	0,4	0,4
Technische Basisleistungsfähigkeit	0,5	0,4	0,4	0,5	0,5	0,5
Benutzerfreundlichkeit	0,5	0,4	0,7	0,3	0,7	0,7
Zusatzfunktionalitäten	0,8	0,8	1,1	0,6	1,0	0,9
Nutzwert	2,7	2,7	2,7	2,6	3,2	3,1
Zielerreichungsgrad	68,0%	66,4%	66,6%	65,6%	80,5%	78,6%

Bezüglich der **Kosten** wurde – wie oben geschildert – vereinfachend nur ein Bewertungskriterium – die Anschaffungskosten resp. der Preis – betrachtet. Die Preisspanne der sechs TPR reichte von ca. 2.000 Euro bis ca. 50.000 Euro und entsprechend deutlich sind die Unterschiede im Teilnutzwert.

In der Kategorie **Administrativ-koordinativer Aufwand** zeigten sich zwischen den TPR-Modellen keine gravierenden Differenzen.

Bezüglich der **Technischen Basisleistungsfähigkeit** variiert der Teilnutzwert leicht. Dies ist vor allem auf unterschiedlich lange Akkuladezeiten (3 bis 10 Stunden) und unterschiedlich hohe Gewichte (12 bis 60 kg) in Verbindung mit entsprechender Transportierbarkeit zurückzuführen.

Große Unterschiede existieren bei der **Benutzerfreundlichkeit**. Maßgeblich waren hier das Kriterium, wie einfach Teilnehmende Zugang zu einem Video-Call erhalten können (Direktlink vs. aufwändige Registrierung), sowie die Kriterien Intuitivität der Bedienung und Reaktionszeit/Verzögerung.

In der Kategorie **Zusatzfunktionalitäten** schwanken die kumulierten Teilnutzwerte ebenfalls stark. Bildschirmgröße und Bildqualität sowie die Höhenverstellbarkeit haben hier einen großen Einfluss.

Insgesamt erhielten die Modelle Double 3 und OhmniPro die besten Bewertungsergebnisse. Beide Modelle erreichen jeweils um die 80 Prozent des

maximal möglichen Nutzwertes. Dieser nicht sehr hohe Wert – nach Wiendahl et al. [15] sind Lösungen im Bereich von Fabrikarüstungen mit einem Zielerreichungsgrad unter 80 Prozent erfahrungsgemäß nicht wettbewerbsfähig – deutet auch darauf hin, dass die marktverfügbaren TPR für die industrielle Praxis noch nicht ausgereift sind. Der vom Projekt PraeRI adressierte Bedarf, TPR für die industrielle Nutzung weiter zu entwickeln, wird damit bestätigt.

4 Fazit

Für die im Projekt PraeRI vorgesehenen Anwendungsszenarien erwiesen sich die TPR-Modelle Double 3 und Ohmni Pro als am besten geeignet. Die Nutzwerte liegen jedoch nur im akzeptablen Bereich, so dass ein Bedarf für Verbesserungen und Weiterentwicklung besteht.

Der entwickelte Bewertungsrahmen und die Vorgehensweise können für andere Anwendungsszenarien adaptiert und nachgenutzt werden.

Die Forschungs- und Entwicklungsarbeit wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ im Projekt PraeRI (Förderkennzeichen 02L21B000-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt der Veröffentlichung liegt bei den Autoren.

5 Literaturverzeichnis

- [1] Björnfot, P., Bergqvist, J., & Kaptelinin, V. (2018). Non-technical users' first encounters with a robotic telepresence technology: An empirical study of office workers. In: Paladyn, Journal of Behavioral Robotics 9 (1), S. 307–322. DOI: 10.1515/pjbr-2018-0022.
- [2] Lee, M. K., & Takayama, L. (2011). Now, I have a body. In: Desney Tan, Geraldine Fitzpatrick, Carl Gutwin, Bo Begole und Wendy A. Kellogg (Hg.): Proceedings of the 2011 annual conference on Human factors in computing systems - CHI '11. the 2011 annual conference. Vancouver, BC, Canada, 07.05.2011 - 12.05.2011. New York, New York, USA: ACM Press, S. 33.
- [3] Lee, WH., Park, J., & Park., C. H. (2018). Acceptability of Tele-assistive Robotic Nurse for Human-Robot Collaboration in Medical Environment. In: Kanda, Šabanović et al. (Hg.) 2018 – Companion of the 2018 ACM/IEEE, S. 171–172. doi: 10.1145/3173386.3177084.
- [4] Hernandez, F., Birke, J., Bullinger, A.C. (2023). The Tribrid-Meeting-Setup – Improving Hybrid Meetings Using a Telepresence Robot. In: Streitz, N.A., Konomi, S. (eds) Distributed,

- Ambient and Pervasive Interactions. HCII 2023. Lecture Notes in Computer Science, vol 14037. Springer, Cham. https://doi.org/10.1007/978-3-031-34609-5_26
- [5] Hernandez, F., Rueffert, D., Hoffmann, H., Bullinger, A.C. (2024). Vom Pixel zur Präsenz - Ferninteraktion mit Telepräsenzrobotern. *Industry 4.0 Science*, Bd. 2024, Nr. 5, Sep. 2024, doi: 10.30844/I4SD.24.5.18.
- [6] Rüffert, D., Kögel, A., Löffler, T. (2025). So fern und doch so nah – Telepräsen-tes Arbei-ten in der Industrie. In: Bullinger-Hoffmann, A.C. (2025). *Telepräsen- te Fernunterstüt- zung in der Industrie*. aw&I – Wissenschaft und Praxis. DOI
- [7] Hernandez, F., Waechter, M., Bullinger, A.C. (2021). A First Approach for Implementing a Telepresence Robot in an Industrial Environment. In: Nunes, I.L. (eds) *Advances in Human Factors and System Interactions*. AHFE 2021. Lecture Notes in Networks and Systems, vol 265. Springer, Cham. https://doi.org/10.1007/978-3-030-79816-1_18
- [8] BLUE OCEAN ROBOTICS. GoBe Robot. <https://gobe.blue-ocean-robotics.com> [letzter Ab- ruf im Januar 2025]
- [9] Axyn Robotics. UBBO-Expert. <https://axyn.fr/ubbo-expert> [letzter Abruf im Januar 2025]
- [10] AVA ROBOTICS. AVA Robot. <https://www.avarobotics.com/ava-robot> [letzter Abruf im Ja- nuar 2025]
- [11] temi USA inc. Temi3. <https://www.robotemi.com/robots/> [letzter Abruf im Januar 2025]
- [12] Double Robotics. Double3. <https://www.doublerobotics.com> [letzter Abruf im Januar 2025]
- [13] OhmniLabs. OhmniPro. <https://ohmnilabs.com> [letzter Abruf im Januar 2025]
- [14] VDI 2225 Blatt 3, Design engineering methodics - Engineering design at optimum cost - Valuation of costs. [Online]. Verfügbar unter: <https://www.beuth.de/en/technical-rule/vdi-2225-blatt-3/10940959>
- [15] Wiendahl, H.-H., Reichardt, J., Nyhuis, P. (2024). *Handbuch Fabrikplanung: Konzept, Ge- staltung und Umsetzung wandlungsfähiger Produktionsstätten, 3., Vollständig überarbei- tete Auflage*. in Hanser eLibrary. München: Hanser Verlag, 2024. doi: 10.3139/9783446473607.

Autoren



Hernandez, Francisco

Francisco Hernandez studierte Fahrzeugtechnik in Bogotá (Kolumbien) bevor er sein Masterstudium in Automobilproduktion und -technik an der TU Chemnitz absolvierte. Seit Mai 2019 arbeitet er an der Professur unter dem Cluster Industrial Engineering mit dem Schwerpunkt Industrie 4.0 im Bereich kollaborierende Roboter, Augmented Reality und Tele-Präsenz-Roboter.



Ruffert, Danny

Danny Ruffert studierte Sports Engineering an der Technischen Universität Chemnitz. Seit 2014 arbeitet er als wissenschaftlicher Mitarbeiter an der Professur Arbeitswissenschaft und Innovationsmanagement im Cluster Industrial Engineering. Seine Aufgaben umfassen die Themenfelder Alterssimulation und Ergonomie sowie die Forschung an Mensch-Maschine-Schnittstellen.



Löffler, Thomas

Dr.-Ing. Thomas Löffler studierte Maschinenbau an der Technischen Universität Chemnitz. Er arbeitete anschließend als wissenschaftlicher Mitarbeiter an der Professur Fabrikplanung und Fabrikbetrieb, als Werkstrukturplaner bei bse engineering Leipzig GmbH sowie als Geschäftsführer des Instituts IREGIA e. V. Seit 2011 ist er an der Professur Arbeitswissenschaft und Innovationsmanagement als Clusterleiter Industrial Engineering tätig.

Ich brauche Unterstützung – Kontaktaufnahme über TPR- Anruffunktion

Francisco Hernandez¹, Danny Rüffert¹

¹TU Chemnitz – Professur Arbeitswissenschaft und Innovationsmanagement

Zusammenfassung

Dieser Beitrag beschreibt die Entwicklung einer Anruf-Funktion für den Telepräsenzroboter (TPR) Double 3 von Double Robotics, die eine bidirektionale Kommunikation zwischen Personen vor Ort und dem TPR-Fahrenden ermöglicht. Während bisher nur der TPR-Fahrende die Kommunikation direkt über den TPR initiieren konnte, können nun auch Personen vor Ort den Kontakt starten, was Verzögerungen reduziert und die Effizienz der Zusammenarbeit steigert. Die technische Umsetzung umfasst eine grafische Benutzeroberfläche mit einem Call-Button sowie eine serverseitige Implementierung. Die modulare Architektur erleichtert zukünftige Anpassungen an andere TPR-Modelle. Die Funktionsfähigkeit wurde in Nutzertests (n = 10) geprüft. Die Arbeit bietet Potenziale für zukünftige Entwicklungen, darunter die Integration von Instant-Messaging-Diensten.

1 Einleitung

Telepräsenzroboter (TPR) haben sich als Werkzeuge etabliert, die eine Brücke zwischen physischer und digitaler Präsenz schlagen können [1]. Bislang war es allerdings nur TPR-Fahrenden möglich, die Kommunikation direkt über den TPR zu initiieren, was bei dringendem Unterstützungsbedarf vor Ort zu Verzögerungen führen kann. Die Person vor Ort konnte nur über zusätzliche Kommunikationswege (E-Mail, Telefon) Telepräsenz-Termine vereinbaren. Dies ist insbesondere in Situationen, in denen schnelle Entscheidungen und Anweisungen erforderlich sind, problematisch [2].

Der im Projekt PraeRI entwickelte Lösungsansatz ist die Implementierung einer Anruf-Funktion in der Benutzeroberfläche eines TPRs, die es Personen vor Ort ermöglicht den TPR-Fahrenden direkt vom TPR aus zu kontaktieren. Diese Funktion würde nicht nur die Benutzerfreundlichkeit des Systems verbessern, sondern trägt auch zur Optimierung von Kommunikationsprozessen bei, indem sie die Effizienz der Zusammenarbeit steigert. In Abbildung 1 ist das grundlegende Funktionsschema abgebildet.

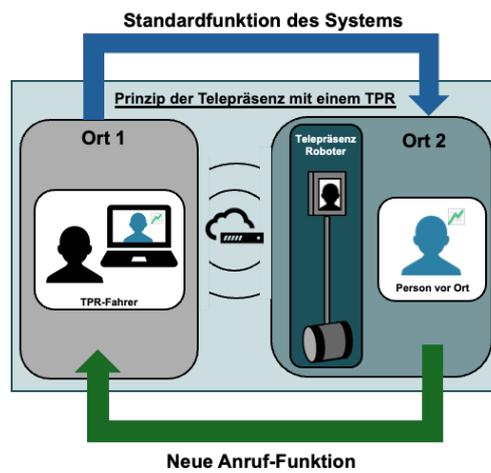


Abbildung 1: Neue Möglichkeit zur Kontaktaufnahme zwischen der Person vor Ort und dem TPR-Fahrer (in Anlehnung an [2])

2 Konzeption einer Anruf-Funktion für den TPR Double 3

Der Double 3 von Double Robotics zeichnet sich durch seine Mobilität, sein autonomes Navigationssystem und seine einfache Steuerung aus. Mit Funktionen wie der „Click-to-Drive“-Navigation und einer Pan-Tilt-Zoom-Kamera bietet der TPR viele Vorteile. Fehlend ist jedoch auch beim Double 3 die Möglichkeit, Anrufe direkt vom TPR zu initiieren.

Für die Entwicklung eines Konzepts einer Anruf-Funktion erfolgte eine Literaturrecherche, bei der bestehende Lösungen zur Mensch-Roboter-Kommunikation analysiert wurden. Ziel war es, die technischen und nutzerzentrierten Anforderungen für eine Anruf-Funktion, insbesondere für den Nutzenden vor Ort, besser zu verstehen. Zu den wichtigsten

Anforderungen gehören eine einfache Bedienbarkeit, die Möglichkeit zur asynchronen Kommunikation sowie ein klares und intuitives Nutzungserlebnis [3], [4], [5].

Kern der Anruf-Funktion ist ein klar sichtbarer und intuitiv nutzbarer Call-Button auf der Nutzeroberfläche vom TPR. Nach dem Drücken des Call-Buttons werden Nutzende durch eine Schritt-für-Schritt-Navigation geführt, bei der relevante Informationen wie der Name des TPR, der Standort und die Dringlichkeitsstufe abgefragt werden. Diese Informationen werden an den Server übermittelt, der eine E-Mail-Benachrichtigung an den TPR-Fahrenden sendet. Diese E-Mail enthält alle notwendigen Details zur Anrufanfrage sowie zwei interaktive Links, über die der TPR-Fahrende den Anruf annehmen oder ablehnen kann [6].

Um eine kontinuierliche Rückmeldung an den Nutzenden zu gewährleisten, wurde ein Echtzeit-Abfrageverfahren vorgesehen. Dieses überprüft den Status der Anfrage in regelmäßigen Intervallen und informiert Nutzende, ob der Anruf angenommen oder abgelehnt wurde. Abbildung 2 zeigt die Netzwerkarchitektur der entwickelten Anruf-Funktion.

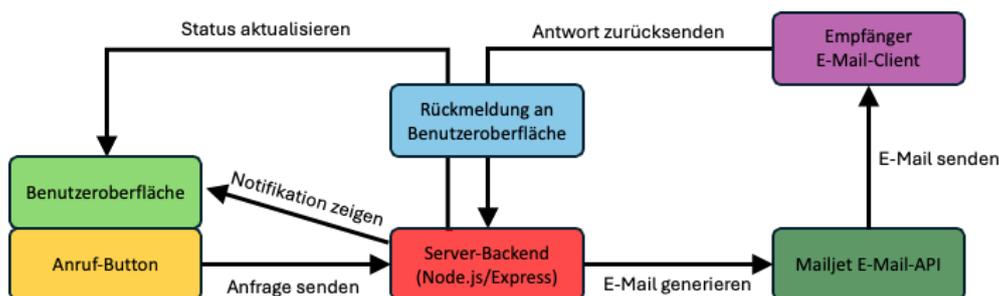


Abbildung 2: Netzwerkarchitektur für die Anruf-Funktion Interface

3 Implementierung eines Demonstrators

Das Konzept wurde als funktionsfähiger Demonstrator umgesetzt. Zur Strukturierung und Gestaltung bzw. der Formatierung der Benutzeroberfläche wurden HTML als Markup Language und CSS als Stylesheet Language verwendet [7]. JavaScript ermöglicht Interaktivität und Benutzeraktionen wie das Drücken eines Call-Buttons oder das Auswählen eines Empfängers. Parallel dazu wurde die serverseitige E-Mail-Funktionalität mit Node.js und der Mailjet-API entwickelt. Node.js besitzt eine asynchrone Architektur und

unterstützt die Verarbeitung von Anfragen in Echtzeit. Mailjet-API ermöglicht eine einfache und zuverlässige E-Mail-Kommunikation und erleichtert die Integration mit serverseitigen Anwendungen [6],[8]. Damit wurde der Mechanismus geschaffen, der automatisierte E-Mail-Anfragen an den TPR-Fahrenden sendet.

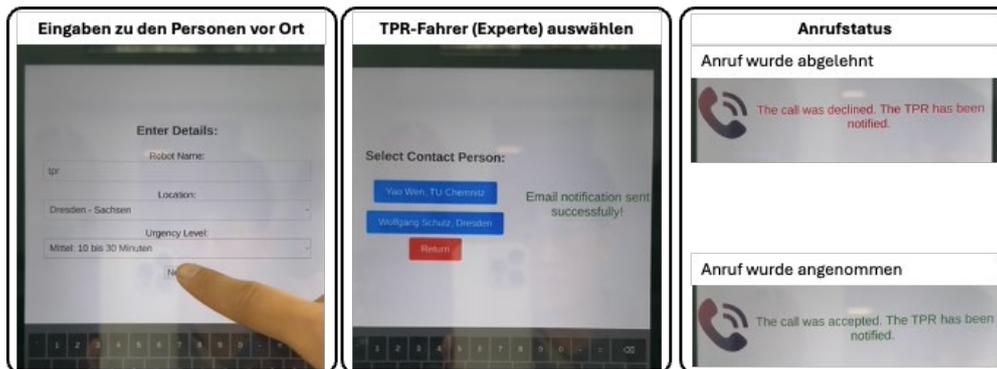


Abbildung 3: Eingabe- und Anrufstatus auf dem Bildschirm des TPR

Im Gegensatz zu einer direkten Anrufverbindung ermöglicht die E-Mail-basierte Lösung dem TPR-Fahrenden, flexibel auf Anfragen zu reagieren, auch wenn er nicht sofort verfügbar ist (siehe Abbildung 4). Parallel dazu erhalten Nutzende auf dem TPR-Bildschirm Rückmeldungen in Echtzeit, ob die E-Mail erfolgreich gesendet wurde und ob die Anfrage vom TPR-Fahrenden akzeptiert oder abgelehnt wurde. Dies wird durch ein kontinuierliches Statusabfragesystem realisiert.



Abbildung 4: E-Mail-Anfrage von der Person vor Ort an den TPR-Fahrer

4 Erprobung des Demonstrators

Der Demonstrator der Anruf-Funktion wurde abschließend in einem Nutzertest (n=10) erprobt. Bewertet wurde über den INTUI-Fragebogen [9]. Ziel dieser Erprobung war es, die Benutzerfreundlichkeit der Anruf-Funktion zu erheben und die intuitiven Aspekte der Benutzeroberfläche zu untersuchen. Die Dimensionen "Mühelosigkeit", "Bauchgefühl", "Vertrautheit" und "Magische Erfahrung" standen hierbei im Fokus. Das Feedback der Testpersonen wurde genutzt, um Verbesserungspotenziale zu identifizieren und die Lösung weiter zu optimieren.

Die Ergebnisse der Nutzertests verdeutlichen die hohe Benutzerfreundlichkeit der Lösung. Die Tester gaben besonders positive Bewertungen in den Dimensionen "Mühelosigkeit", "Vertrautheit" und "Bauchgefühl" ab, was darauf hindeutet, dass die Bedienung als einfach, vertraut und angenehm empfunden wurde. Die Dimension "Magische Erfahrung" erhielt ebenfalls gute Bewertungen, da die Nutzenden das innovative und reibungslose Funktionieren der Anruf-Funktion positiv bewerteten. Die Ergebnisse der Tests wurden genutzt, um kleinere Optimierungen an der Benutzeroberfläche vorzunehmen, darunter die Klarheit der Navigationsschritte und die Darstellung der Rückmeldungen.

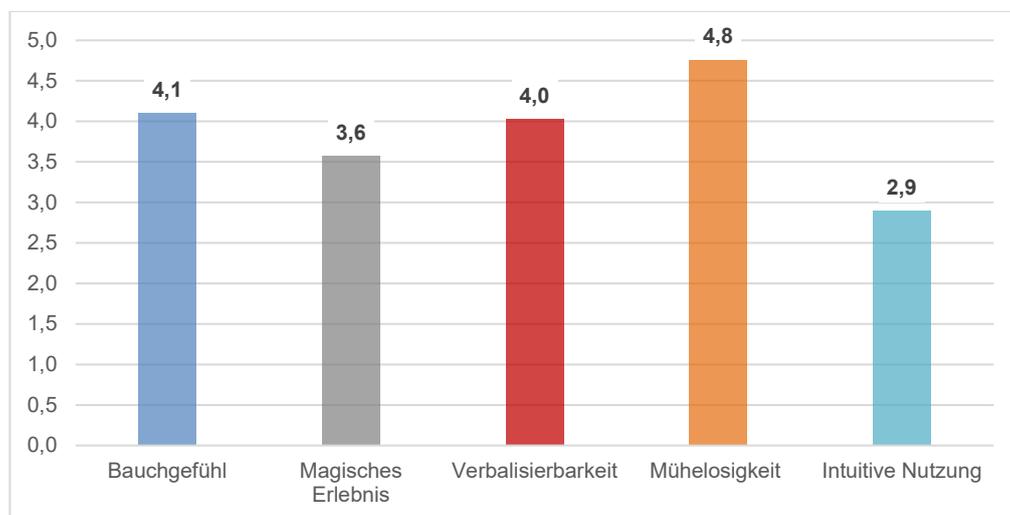


Abbildung 5: Ergebnisse der Nutzertest

5 Diskussion und Ausblick

Insgesamt deuten die Ergebnisse auf eine deutliche Verbesserung der Kommunikationsfähigkeit zwischen TPR-Fahrenden und den Nutzenden vor Ort hin. Letztere haben nun die Möglichkeit, ohne zusätzliche Endgeräte direkt über den TPR den TPR-Fahrenden zu kontaktieren.

Durch die E-Mail-basierte Lösung anstatt einer Echtzeit-Anruffunktion ist eine asynchrone Kommunikation möglich. Das heißt, die Kommunikation kann in den Fällen, in denen der TPR-Fahrende gerade nicht erreichbar ist, zeitversetzt stattfinden. Dabei kann auch die Dringlichkeit berücksichtigt werden. Im Hinblick auf die Skalierbarkeit bietet die modulare Architektur der Lösung klare Vorteile. Durch die Trennung der Benutzeroberfläche, der Serverlogik und der E-Mail-Kommunikation kann die Lösung leicht an neue Anforderungen angepasst oder auf andere TPR-Modelle übertragen werden.

Dennoch besteht Potenzial zur Verbesserung. Die größte Herausforderung ist die vollständige Integration der Anruf-Funktion in die proprietäre Benutzeroberfläche des Double 3. Aktuell läuft die Funktion als separate, zusätzliche Webanwendung. Eine direkte Integration in die bestehende Benutzeroberfläche könnte die Bedienfreundlichkeit weiter erhöhen und den Anlernaufwand für Nutzende verringern. Ein weiteres Verbesserungspotenzial besteht in der Diversifizierung der Kommunikationskanäle. Neben der E-Mail-Funktion könnte eine zukünftige Erweiterung mit einem Instant-Messaging-Dienst ergänzt werden. Die Einbindung vertrauter Messaging-Dienste ermöglicht es Personen vor Ort, Anfragen in Echtzeit zu stellen und visuelle Informationen wie Bilder oder Videos zu übermitteln, wodurch die Effizienz der Problemlösung erhöht werden kann.

Zukünftige Arbeiten könnten auch die Integration von Künstlicher Intelligenz (KI) erwägen. Eine KI könnte zum Beispiel die Dringlichkeit von Anfragen automatisch bewerten und die Anrufe entsprechend priorisieren.

6 Fazit

Die als Demonstrator entwickelte Anruf-Funktion ermöglicht es Personen vor Ort, ohne zusätzliche Endgeräte direkt über den TPR Kontakt mit dem TPR-Fahrenden aufzunehmen. Diese Funktion kann erhebliche Vorteile bieten in industriellen Umgebungen, in denen schnelle Reaktionen erforderlich sind und nicht alle Akteure in jeder Situation Zugang zu allen Kommunikationsmöglichkeiten haben.

Die Forschungs- und Entwicklungsarbeit wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ im Projekt PraeRI (Förderkennzeichen 02L21B000-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt der Veröffentlichung liegt bei den Autoren.

7 Literaturverzeichnis

- [1] Hernandez, F., Waechter, M., Bullinger, A.C. (2021). A First Approach for Implementing a Telepresence Robot in an Industrial Environment. In: Nunes, I.L. (eds) Advances in Human Factors and System Interactions. AHFE 2021. Lecture Notes in Networks and Systems, vol 265. Springer, Cham. https://doi.org/10.1007/978-3-030-79816-1_18
- [2] Hernandez, F., Rueffert, D., Hoffmann, H., Bullinger, A.C. (2024). Vom Pixel zur Präsenz - Ferninteraktion mit Telepräsenzrobotern. Industry 4.0 Science, Bd. 2024, Nr. 5, Sep. 2024, doi: 10.30844/I4SD.24.5.18.
- [3] Naumann, A.B., Pohlmeier, A.E., Husslein, S., Kindsmüller, M.C., Mohs, C., Israel, J.H. (2008). Design for intuitive use: beyond usability. In CHI '08 Extended Abstracts on Human Factors in Computing Systems, Florence Italy: ACM, Apr. 2008, S. 2375–2378. doi: 10.1145/1358628.1358688.
- [4] Bader, F., Schön, E. M., Thomaschewski, J. (2017). Heuristics Considering UX and Quality Criteria for Heuristics. In: IJIMAI, Bd. 4, Nr. 6, S. 48, doi: 10.9781/ijimai.2017.05.001.
- [5] Jacobsen J., Meyer, L. (o. J.). Praxisbuch Usability und UX: Was jeder wissen sollte, der Websites und Apps entwickelt. <https://www.medimops.de/undefined/jens-jacobsen-praxisbuch-usability-und-ux-was-jeder-wissen-sollte-der-websites-und-apps-entwickelt-bewaehrte-methoden-praxisnah-erklaert-gebundene-ausgabe-M03836244233.html> (letzter Zugriff: 11. Dezember 2024)
- [6] Mailjet. E-Mail-Zustelldienst für Marketing- und Entwicklerteams. <https://www.mailjet.com/de/> (letzter Zugriff: 11. Dezember 2024)
- [7] Duckett, J., Ruppert, G., Moore, J., Duckett, J. (2025). JavaScript & jQuery: interaktive Websites entwickeln, 1. Aufl. Weinheim: Wiley-VCH
- [8] Prediger R., Winzinger, R. (2015). Node.js: Professionell hochperformante Software entwickeln. München: Hanser, 2015. doi: 10.3139/9783446437586.
- [9] Ullrich D., Diefenbach, S. (2011). Erlebnis intuitive Interaktion — ein phänomenologischer Ansatz. i-com, Bd. 10, Nr. 3, S. 63–68, Nov. 2011, doi: 10.1524/icom.2011.0036.

Autoren



Hernandez, Francisco

Francisco Hernandez studierte Fahrzeugtechnik in Bogotá (Kolumbien) bevor er sein Masterstudium in Automobilproduktion und -technik an der TU Chemnitz absolvierte. Seit Mai 2019 arbeitet er an der Professur unter dem Cluster Industrial Engineering mit dem Schwerpunkt Industrie 4.0 im Bereich kollaborierende Roboter, Augmented Reality und Tele-Präsenz-Roboter.



Ruffert, Danny

Danny Ruffert studierte Sports Engineering an der Technischen Universität Chemnitz. Seit 2014 arbeitet er als wissenschaftlicher Mitarbeiter an der Professur Arbeitswissenschaft und Innovationsmanagement im Cluster Industrial Engineering. Seine Aufgaben umfassen die Themenfelder Alterssimulation und Ergonomie sowie die Forschung an Mensch-Maschine-Schnittstellen.

Bitte folge mir – Steuerung eines TPR über Gesten

Florian Schnabel¹

¹HLS Automation and Robotic GmbH

Zusammenfassung

Telepräsenzroboter zeichnen sich dadurch aus, dass sie von einem Nutzer aus der Ferne über das Internet gesteuert werden. Da jedoch auch ein Bedarf besteht, Telepräsenzroboter an ihrem Einsatzort von lokalen Nutzern steuern zu können, wurde für ausgewählte Fahrfunktionen eine Gestensteuerung entwickelt und deren Funktionsfähigkeit nachgewiesen.

1 Einleitung

Telepräsenzroboter (TPR) zeichnen sich dadurch aus, dass sie, anders als herkömmliche Videokonferenzsysteme, von dem Nutzenden aus der Ferne (TPR-Fahrender) autonom gesteuert werden können: Die Person, die den TPR steuert kann den TPR zur von ihr gewünschten Positionen im Raum fahren, die Orientierung im Raum ändern und – je nach Modell – die Positionshöhe und Ausrichtung der Kamera sowie den Zoomfaktor anpassen.

Die Person, die vor Ort mit dem TPR interagiert, hat diese Interaktionsmöglichkeiten bei den marktverfügbaren TPR nicht. Sie kann den TPR bestenfalls heben oder tragen, um ihn beispielsweise an eine Position zu bringen, die der TPR-Fahrende nicht kennt, die ihm nur schwer zu beschreiben oder zu zeigen ist oder zu der er nur mit Aufwand navigieren kann (z. B. durch schmale Passagen).

In solchen Fällen wäre es vorteilhaft, wenn die Person vor Ort dem TPR auch direkt Instruktionen wie „Stopp“, „Start“ oder „Folge mir“ erteilen kann. Damit dafür keine zusätzlichen Endgeräte nötig werden und die mit dem TPR-Einsatz angestrebte intuitive Interaktionsweise erhalten bleibt, bietet sich hierfür eine Gestensteuerung an. Eine solche Gestensteuerung wurde im Projekt PraeRI für den TPR Double 3 prototypisch entwickelt.

2 Umsetzung und Demonstrator

Konzept

Die Steuerung des TPR soll durch die Detektion spezifischer Handgesten erfolgen. Dazu soll das Kamerasignal des Roboters analysiert werden, um die Position und Bewegungen der Hand des Nutzens vor Ort zu interpretieren. Die daraus abgeleiteten Gesten sollen in Fahr- und Bedienbefehle umgewandelt werden, die über die Steuerschnittstellen des Roboters ausgeführt werden.

Technische Rahmenbedingungen

Der TPR Double 3 bietet keine Möglichkeit, Software direkt auf dem Gerät zu installieren. Es existieren weder Hard- noch Software-Schnittstellen, die einen Zugriff auf interne Steuerungsmechanismen oder Signale ermöglichen. Einzig die Nutzung des vom Hersteller bereitgestellten Webinterfaces des TPR ist möglich. Über dieses Interface steuert der TPR-Fahrende den Roboter normalerweise aus der Ferne manuell.

Technische Lösung

Unter den gegebenen technischen Rahmenbedingungen wurde folgende Lösung verfolgt: Über das Webinterface wird eine Verbindung mit dem TPR hergestellt. Darüber wird das Kamerabild des TPR – wie im normalen Betrieb auch – als Videostream übertragen. Von dem Videostream auf dem Bildschirm des TPR-Fahrenden werden Screenshots erfasst und bezüglich des Vorkommens von bestimmten Handgesten analysiert. Werden relevante Handgesten erkannt, werden diese in Steuerbefehle übersetzt. Dazu werden virtuelle Tastendrücke auf dem PC des Nutzens generiert, die das Webinterface – analog zu Eingaben des TPR-Fahrenden – als Steuerbefehle für den Roboter interpretiert. Die Umsetzung der technischen Lösung erfolgte in mehreren Schritten.

(1) Vorbereitung der Entwicklungsumgebung

Eine Python-Umgebung wurde in einer virtuellen Maschine (VMWARE) eingerichtet, um die Softwareentwicklung und -tests unabhängig vom Hauptsystem durchzuführen.

(2) Erfassung des Videostreams

Ein aktiver Anruf mit dem TPR wurde initiiert, und Screenshots des Bildschirms des TPR-Fahrenden wurden in festgelegten Zeitintervallen aufgenommen. Diese Screenshots bildeten die Grundlage für die Analyse der Handgesten.

(3) Erkennung der Handgesten

Für die Erkennung der Handgesten wurde eine Mediapipe-Bibliothek eingesetzt. Erste Tests zeigten, dass eine Hand mit den Orientierungspunkten prinzipiell zuverlässig erkannt wird. Es traten jedoch mehrere spezifische Herausforderungen auf, die iterative Anpassungen erforderten. Dazu zählten:

Fingerzählung: Die Zählung der in die Kamera gezeigten Finger bereitete Probleme, da der anfänglich eingesetzte Algorithmus, der die Fingerspitzenhöhe relativ zu anderen Fingergelenken bewertete, fehlerhafte Ergebnisse lieferte. Der Algorithmus arbeitete allein auf der Annahme, dass die Spitzen der Finger, die in der betreffenden Geste nicht gezeigt werden, tiefer liegen als andere Gelenkpunkte der Finger. Die Analyse der relativen Höhe und der Distanz der Gelenkpunkte war jedoch anfällig für Verzerrungen bei Handrotationen. Daher wurde in einer Überarbeitung des Algorithmus die Positionsbewertung um die relative Lage des Daumens zu den anderen Fingern ergänzt. Der finale, nun stabilere Algorithmus kombiniert die relativen Höhen und Seitenverhältnisse der Gelenkpunkte, um zu bestimmen, ob der Daumen links oder rechts von den anderen Fingern liegt.

Erkennung mehrerer Hände: Fehlinterpretationen der Gesten entstanden auch, wenn gleichzeitig mehrerer Hände erkannt wurden. Als einfache Lösung für dieses Problem wird die Handgestenerkennung deaktiviert, sobald mehr als eine Hand sichtbar ist. Ein Hinweis auf dem Bildschirm informiert den Nutzenden und fordert dazu auf, nur eine Hand zu zeigen.

(4) Systemintegration

Abschließend erfolgte die Integration der Gestenerkennung in den Steuerungsablauf des TPR. Dazu wurden die Gesten (Anzahl erkannter Finger) mit virtuellen Tastatureingaben verknüpft, die wiederum – in gleicher Weise wie reale Tastatureingaben des TPR-Fahrers – bei Übertragung durch das Webinterface Fahrbefehle im TPR auslösen.

3 Erprobung

Funktionsfähigkeit der Lösung wurde für drei Gesten (Abbildung 1) erprobt:

- Geste zwei Finger = Taste <Pfeil oben> = Fahrbefehl „Vorwärts“
- Geste fünf Finger = Taste <Beenden> = Fahrbefehl „Stopp“
- Geste vier Finger = Taste <Pfeil unten> = Fahrbefehl „Rückwärts“

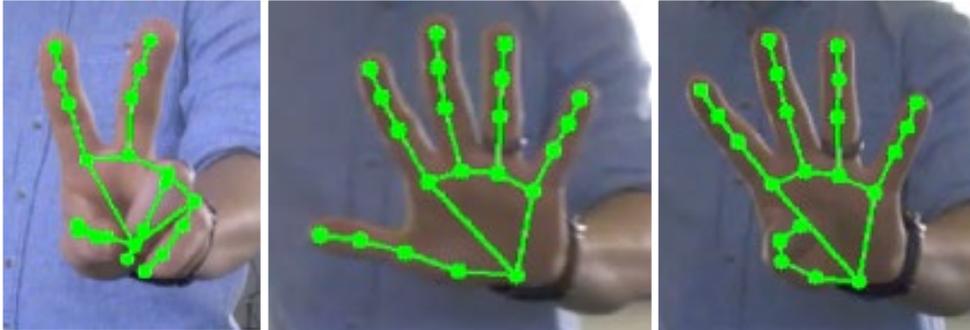


Abbildung 1: Gesten für „Vorwärts“ (links), „Stopp“ (Mitte) und „Rückwärts“ (rechts)

In den Tests zeigte sich, dass die Steuerung mit den ausgewählten Gesten unter üblichen Lichtbedingungen in Innenräumen zuverlässig funktioniert und der Roboter sich auf diese Weise fahren und stoppen lässt. Auch das Erkennen einer zweiten, die Gestenerkennung potenziell störenden Hand und das dann notwendige Unterbrechen der Gestenerkennung sowie das Instruieren der Person vor Ort, die zweite Hand aus dem Kamerabild zu entfernen, funktioniert zuverlässig.

4 Zusammenfassung und Ausblick

Es konnte gezeigt werden, dass der TPR Double 3 bezüglich der Funktion Vorwärts- und Rückwärtsfahren sowie Stoppen mittels Gestensteuerung zuverlässig gesteuert werden kann. Die technische Lösung wurde – da keine anderen Schnittstellen zur Verfügung standen – das Webinterface des TPR.

Künftige Entwicklungen könnten die Funktionalität auch direkt an Bord von TPR realisieren, um die Robustheit zu erhöhen, Latenzen zu vermeiden und die über das Internet zu übertragende Datenmenge zu reduzieren. Außerdem ist es nötig, weitere Gesten zu berücksichtigen, um einen größeren Funktionsumfang (z. B. Richtungsänderung, „Folge-mir-Funktion“) umsetzen zu können. Schließlich besteht auch der Bedarf, die Robustheit der Gestenerkennung bei schwierigen Lichtverhältnissen zu erhöhen.

Außerdem bietet die Kombination mit einer Sprachsteuerung vielversprechende Perspektiven – etwa für die Nutzung von TPR durch Menschen mit verschiedenen Einschränkungen.

Die Forschungs- und Entwicklungsarbeit wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ im Projekt PraeRI (Förderkennzeichen 02L21B000-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt der Veröffentlichung liegt beim Autor.

Autor**Schnabel, Florian**

Florian Schnabel arbeitet seit 2023 für die HLS Robotic Automation GmbH. Er ist für die Abwicklung von Automatisierungsprojekten als Projektleiter verantwortlich. Im Rahmen des Forschungsprojektes PraeRI hat er seine Diplomarbeit zum Thema Telepräsenzroboter geschrieben.

Kannst Du mir das zeigen? Entwicklung und Evaluation einer Zeigefunktion für einen Telepräsenzroboter

Rakshith Venepally¹, Francisco Hernandez², Danny Rüffert²

¹LS Software Engineering GmbH, ²TU Chemnitz – Professur Arbeitswissenschaft und Innovationsmanagement

Zusammenfassung

Bei der natürlichen Kommunikation von Personen spielt neben der Sprache auch die Gestik eine große Rolle. Die Möglichkeit, auf etwas Relevantes zu zeigen – hat der Fernanwender bei marktverfügbaren Telepräsenzrobotern (TPR) nicht. In dieser Arbeit wird die Entwicklung und Evaluation einer für TPR neuartigen Zeigefunktion mittels schwenkbaren Laserpointers vorgestellt. Die Evaluation zeigt, dass der entwickelte Ansatz vielversprechend ist, jedoch Optimierungspotenzial in Bezug auf Sichtbarkeit und Steuerungspräzision bietet. Zukünftige Verbesserungen, wie eine Point-to-Click-Funktion und optimierte Videostreams, werden diskutiert.

1 Einleitung

Telepräsenzroboter (TPR) haben das Vermögen, die natürliche Kommunikation über die Ferne zwischen einem Lokal- und Fernanwender zu verbessern. Im Gegensatz zu herkömmlichen Methoden wie Videotelefonie oder beim Einsatz von Videokonferenzsystemen, können die Nutzer die Kamera und damit ihr Sichtfeld selbstbestimmt positionieren. Gleichzeitig zeigen aktuelle TPR immer noch Schwachstellen gegenüber natürlicher Kommunikation. Natürliche Kommunikation zeichnet sich neben der Sprache auch durch eine unbewusste bzw. bewusste Gestenkommunikation aus. Gerade das bewusste Zeigen auf Objekte ist derzeit nicht über TPR möglich, aber eine wichtige Funktion, die im industriellen Umfeld benötigt wird [1].

Der vorliegende Beitrag beschreibt die Entwicklung einer Zeigefunktion, die auf dem TPR Ohmni Pro (Anbieter: OhmniLabs) integriert wurde. Der Schwerpunkt liegt auf der Hardware- und Softwareentwicklung, der Integration sowie der Evaluation des Systems.

2 Entwicklung der Zeigefunktion

In die Entwicklung der Zeigefunktion erfolgte in einem mehrstufigen Vorgehen aus Anforderungsanalyse, Komponentenauswahl und Umsetzung.

Die Anforderungen wurden qualitativ in Fokusgruppen und quantitativ mittels Fragebogen (n=25) erhoben. Der Probandenpool bestand aus verschiedenen Berufs- und Altersgruppen. Alle Teilnehmer waren mit virtuellen Arbeitsmethoden vertraut. Spezifische Anforderungen betrafen die Integration ins bestehende TPR-System, Modularität, Zugänglichkeit und Kosten.

Bei der Lösungssuche mit Fokusgruppen zeigte sich, dass ein bewegbarer Laserpointer am besten geeignet ist, um eine Zeigefunktion umzusetzen. Im weiteren Verlauf der Entwicklung wurde von den Probanden dann bewertet, an welcher Stelle des TPRs eine Zeigefunktion angebracht werden und welchen Bewegungsumfang diese haben sollte. Nach Meinung der Probanden befindet sich der geeignete Montageort für den Laserpointer direkt über dem Lautsprecher des Ohmni Pro TPR und der Pointer sollte sich um jeweils 180° neigen und schwenken lassen.

Folgende Komponenten wurden verwendet:

- **Microcontroller-Board (Arduino Nano v3):** eine Mikrocontrollerplatine, die wegen ihrer Anpassungsfähigkeit und Benutzerfreundlichkeit für eine Vielzahl elektronischer Anwendungen bekannt ist [2].
- **Servomotor (SG90-Servomotor):** sind mit allen Mikrocontrollern kompatibel und bieten ausreichend Drehmoment, um eine präzise kontrollierte Bewegung des Laserpointers in zwei Achsen (vertikal Neigen und horizontal Schwenken) zu ermöglichen [3, 4].
- **Laserpointer (KY-008):** besteht aus einer kleinen Platine mit unterstützenden Komponenten und der darauf installierten Laserdiode. Die Diode erzeugt einen für das Zeigen geeigneten konzentrierten, rot leuchtenden Lichtstrahl. Das Modul lässt sich problemlos mit Mikrocontrollern wie den Arduino verbinden [5].
- **Ohmni FlexAdapter:** gehört zum Ohmni Erweiterungsset und gewährleistet die korrekte Stromversorgung (Umwandlung von 5 V auf 3,3 V) von externen Geräten, hier zwischen Arduino und Ohmni Pro [6].
- **3-D gedruckte Plattform:** selbstentwickeltes Bauteil zum Schutz der elektronischen Komponenten und als Halterung am Ohmni Pro.

Abbildung 1 zeigt das Funktionsschema und Abbildung 2 die auf der Plattform fertig montierten Komponenten.

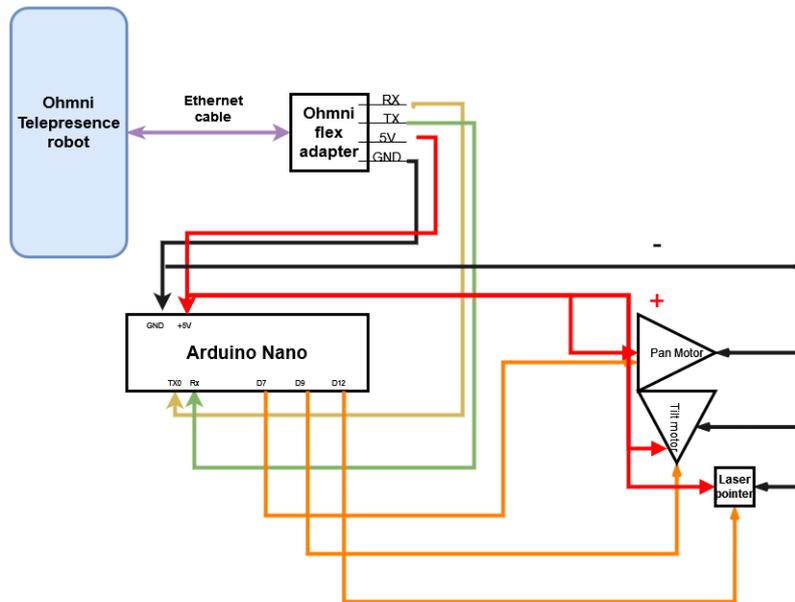


Abbildung 1: Funktionsschema Zeigefunktion

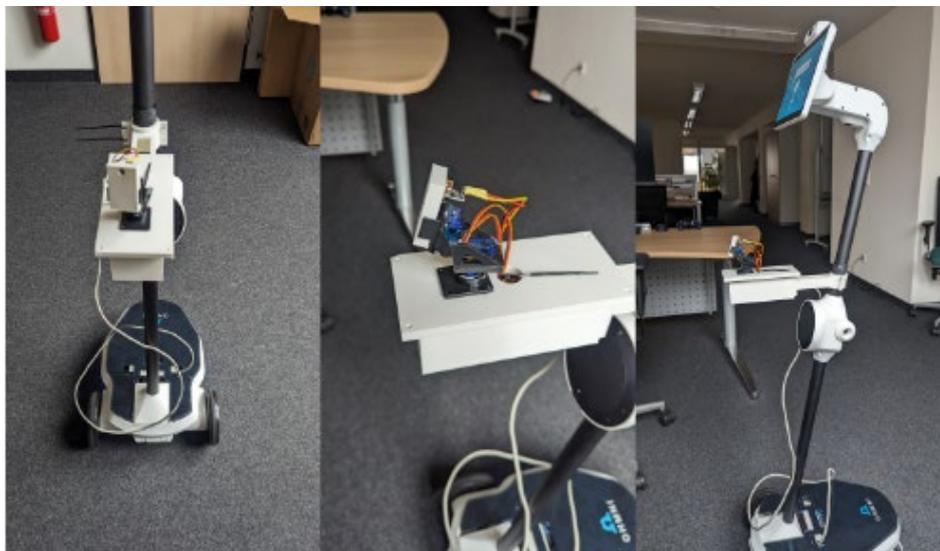


Abbildung 2: Montierte Hardwarekomponenten

Die Ansteuerung des Laserpointers erfolgt über einen in HTML programmierten virtuellen Joystick in Kombination mit einem JavaScript-Plugin. Dies ermöglicht die Übertragung von Positionswerten des virtuellen Joysticks an die Servomotoren. Eine Reset-Funktion erlaubt es, die Position des Laserpointers schnell zurückzusetzen. Abbildung 3 zeigt das User-Interface.

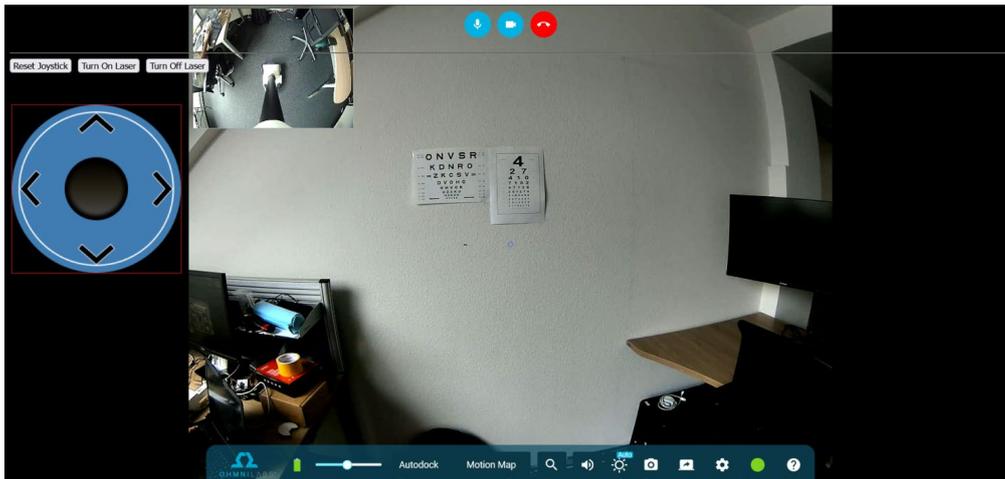


Abbildung 3: User-Interface Zeigefunktion/Joystick

3 Evaluation

Die Zeigefunktion wurde von zehn Probanden, die mit den Grundfunktionen des TPR bereits vertraut waren, getestet. Dazu war der Laserpointer an zwei Orten A und B auf je vier Zielmarken unterschiedlicher Größe und Geometrie zu richten. Die beiden Orte A und B waren unterschiedlich gut belichtet bzw. beleuchtet (Abbildung 4). Die Reihenfolge (Standort A-B oder B-A) wurde jeweils zufällig gewählt. Der Abstand des TPR zu den Zielmarken war für alle Probanden gleich.

Die Präzision der Ausrichtung der Laserpointer wurde anhand der gemessenen Abweichung von der Zielmitte bestimmt. Zusätzlich wurde die Zeit bis zum Abschluss der Zielmarkierung aufgenommen. Die subjektiv empfundene Gebrauchstauglichkeit wurde mit der System Usability Scale (SUS) erhoben [7, 8, 9]. Dieser Fragebogen besteht aus 10 Items und verwendet ein 5-stufige Likert-Skala. Die Bewertungen aller Items werden zu einem Gesamt-Score aggregiert. Die maximal erreichbare Punktzahl beträgt 100 (beste vorstellbare Lösung). Ab 85 Punkte wird von einer ausgezeichneten, ab ca. 70 Punkten von einer akzeptablen Lösung gesprochen.

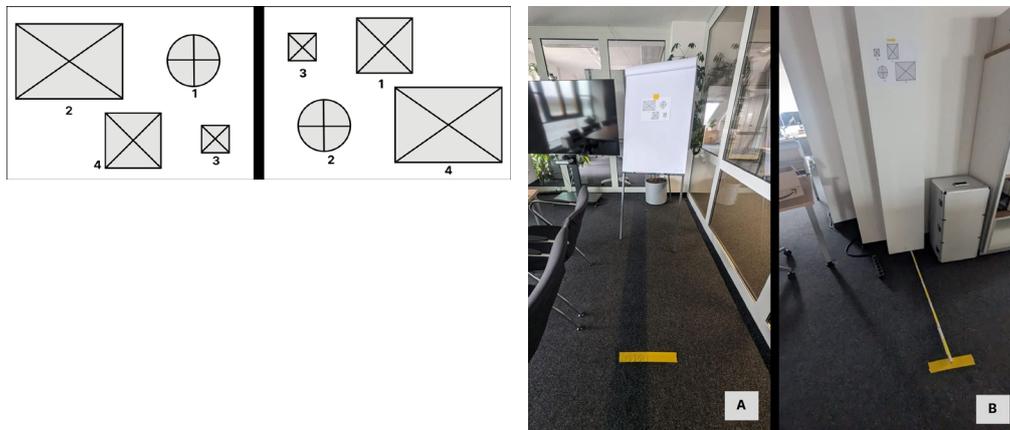


Abbildung 4: Zielmarken und Teststandort A und B

Die Probanden waren alle in der Lage, den Prototypen des Laserpointers zu bedienen. Test erbrachten im Einzelnen folgende Ergebnisse:

Präzision: Die Probanden hatten große Probleme, die Zielmarken genau zu treffen. Die mittlere Abweichung vom Zielmittelpunkt – berechnet für jeden Probanden und jeden Testdurchlauf (Standort A und B) – betrug bis zu knapp 2 Zentimeter. Die Abweichungen waren unabhängig von den Lichtverhältnissen. Auch im jeweils zweiten Testdurchlauf (Standort A oder B) reduzierten sich die Abweichungen vom Zielmittelpunkt kaum. Allerdings verringerte sich – mit offenbar wachsender Vertrautheit mit dem virtuellen Joystick – die Zeit bis zum Abschluss der Zielmarkierung.

Gebrauchstauglichkeit: Die subjektiv wahrgenommene Gebrauchstauglichkeit wurde auf der SUS mit 64,5 Punkte bewertet. Die Lösung ist daher noch nicht akzeptabel.

Lichtverhältnisse: Eine dunklere Umgebung verbessert die Wahrnehmung des Laserpointers, wirkt sich jedoch nicht auf die Präzision aus.

Qualitatives Probanden-Feedback: Die Probanden gaben an, dass die Benutzeroberfläche durchaus intuitiv bedienbar sei, jedoch sensumotorisch Schwierigkeiten bei der Ausrichtung des Laserstrahls mit dem virtuellen Joystick bestehen. Besonders bei sehr hellen Hintergründen – wie im Test bei dem Whiteboard am Standort A – ist das Licht des Laserpointers zudem schlecht zu erkennen.

4 Zusammenfassung und Fazit

Der entwickelte Laserpointer bietet die Möglichkeit, dass Fernanwender eines TPR die Aufmerksamkeit von Lokalanwendern durch eine Zeigefunktion lenken können. Die technische Machbarkeit wurde für den TPR Ohmni nachgewiesen. Bei der Übertragung auf andere TPR sind deren Hard- und Software-schnittstellen sowie verfügbaren Entwicklerumgebungen zu berücksichtigen.

Die im Prototyp aus pragmatischen Gründen gewählte Ansteuerung über einen virtuellen Joystick bedarf jedoch der Weiterentwicklung oder alternativer Lösungen wie einer Point and Click-funktion in der Nutzeroberfläche des Fernanwenders. Weitere Verbesserungspotenziale betreffen die Sichtbarkeit des Pointers bei heller Umgebung und die Latenz des Videostreams.

Die Forschungs- und Entwicklungsarbeit wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ im Projekt PraeRI (Förderkennzeichen 02L21B000-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt der Veröffentlichung liegt bei den Autoren.

5 Literaturverzeichnis

- [1] Hernandez, F., Waechter, M., Bullinger, A. C. (2021) A first approach for implementing a telepresence robot in an industrial environment. In Advances in Human Factors and System Interactions: Proceedings of the AHFE 2021 Virtual Conference on Human Factors and Systems Interaction, July 25-29, 2021, USA, pages 141–146. Springer
- [2] Arduino CC. Arduino nano documentation. <https://docs.arduino.cc/hardware/nano/>
- [3] Department of Electrical and electronic Engineering - Imperial College London. Sg90 datasheet by department of electrical and electronic engineering - imperial college London. http://www.ee.ic.ac.uk/pcheung/teaching/DE1_EE/stores/sg90_datasheet.pdf.
- [4] Adafruit. Servo controlled pan and tilt device. <https://www.adafruit.com/product/1967>.
- [5] Arduino. Arduino modules. <https://arduinomodules.info/ky-008-laser-transmitter-module/>.
- [6] Ohmnilabs-Ohmni flexAdapter. Ohmni telepresence robot developer documentation. <https://docs.ohmnilabs.com/developertext/#ohmni-flexadapter>
- [7] Brooke, J. et al. (1996) SUS -a quick and dirty usability scale. Usability evaluation in industry, 189(194):4–7, 1996
- [8] Brooke. J (2013) Sus: a retrospective. Journal of usability studies, 8(2): 29–40, 2013
- [9] 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

Autoren



Venepally, Rakshith

Rakshith Venepally studierte Maschinenbau in Indien und schloss 2024 sein Masterstudium in Digital Engineering an der Otto-von-Guericke-Universität Magdeburg ab. Seit 2022 arbeitet er bei LS Software & Engineering GmbH mit den Schwerpunkten Visualisierung, Digital Twins und einer Masterarbeit über die Erweiterung eines Telepräsenzroboters.



Hernandez, Francisco

Francisco Hernandez studierte Fahrzeugtechnik in Bogotá (Kolumbien) bevor er sein Masterstudium in Automobilproduktion und -technik an der TU Chemnitz absolvierte. Seit Mai 2019 arbeitet er an der Professur unter dem Cluster Industrial Engineering mit dem Schwerpunkt Industrie 4.0 im Bereich kollaborierende Roboter, Augmented Reality und Tele-Präsenz-Roboter.



Rüffert, Danny

Danny Rüffert studierte Sports Engineering an der Technischen Universität Chemnitz. Seit 2014 arbeitet er als wissenschaftlicher Mitarbeiter an der Professur Arbeitswissenschaft und Innovationsmanagement im Cluster Industrial Engineering. Seine Aufgaben umfassen die Themenfelder Alterssimulation und Ergonomie sowie die Forschung an Mensch-Maschine-Schnittstellen.

Roboter öffnen Türen – Neue Wege der Telepräsenz

Nikolas von Lonski¹

¹Actimage GmbH

Zusammenfassung

Die Einbindung eines Telepräsenzroboters (TPR) in bestehende KNX-Systeme eröffnet neue Möglichkeiten zur physischen Interaktion mit Gebäuden. Actimage GmbH entwickelte hierfür einen KNX-USB-Treiber, der auf Linux basiert und über RF-Technologie kommuniziert. Die Implementierung erfolgte vollständig in Rust, wodurch Software- und Netzwerkisolierung erheblich verbessert und kostspielige mechanische Komponenten vermieden werden.

1 Einleitung

Telepräsenzroboter (TPR) ermöglichen Nutzenden, aus der Ferne physisch präsent zu sein und sich autonom in Gebäuden zu bewegen. Bisher waren direkte physische Aktionen, wie das Öffnen von Türen oder das Einschalten von Lichtanlagen, nur durch mechanische Roboterarme oder komplexe Umbauten möglich. Actimage adressierte dieses Problem, indem ein TPR in das weit verbreitete KNX-Gebäudeautomationssystem integriert wurde.

2 Technisches Konzept

Ziel der Integration war eine einfache und kostengünstige Lösung, die sowohl Sicherheit als auch Bedienkomfort sicherstellt. Es sollten keine tiefgreifenden Eingriffe in bestehende KNX-Installationen erforderlich sein.

Der zentrale Baustein der Lösung ist ein speziell entwickelter KNX-USB-Treiber, der auf Linux läuft. Die Kommunikation mit dem KNX-System erfolgt kabellos über Funk im Radiofrequenzbereich (RF), wodurch eine physische

Trennung zum Firmen IP-Netzwerk sichergestellt wird. Abbildung 1 zeigt das Konzept im Überblick.

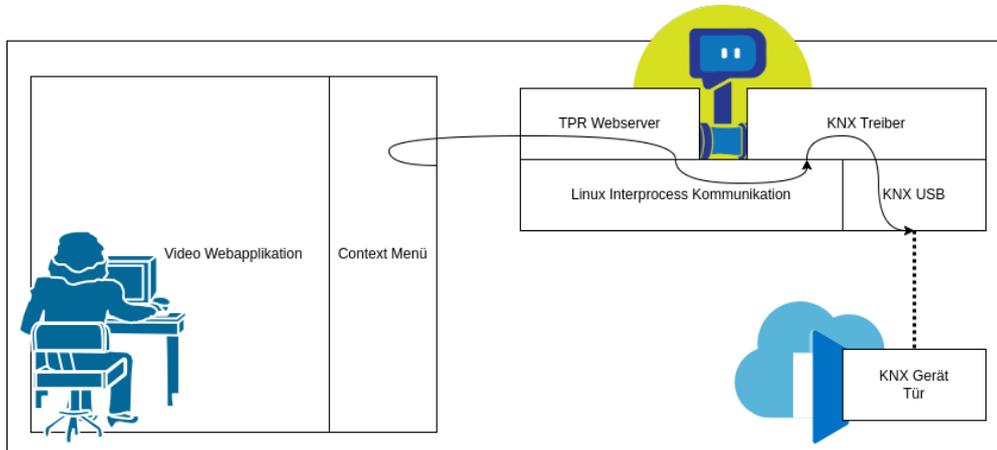


Abbildung 1: Konzept der Verbindung zwischen TPR und KNX

3 Technische Umsetzung

Der Treiber wurde in der Programmiersprache Rust implementiert, da diese eine hohe Speichersicherheit und Fehlervermeidung bietet. Durch Rust konnten typische Schwachstellen vermieden werden, was insbesondere für den industriellen Einsatz von hoher Bedeutung ist.

Die RF-basierte Kommunikation ermöglicht eine vollständige physische Trennung zwischen TPR und IP-Netzwerk. Dies sorgt für erhöhte Sicherheit, da selbst im Falle einer Kompromittierung des Roboters kein Zugriff auf das sensible Gebäudenetzwerk möglich ist.

Vergleich mit alternativen Technologien:

ZigBee: ZigBee wird überwiegend stationär eingesetzt. Bei mobilen Anwendungen können in ZigBee-Netzen beim Übergang zwischen Knotenpunkten Unterbrechungen oder Verzögerungen entstehen.

Infrarot (IR): Infrarot-basierte Lösungen bieten eine geringe Reichweite und sind stark sichtlinienabhängig. Diese Abhängigkeit erschwert den Einsatz in

komplexen, verwinkelten Gebäuden, wodurch IR für einen mobilen Roboter ungeeignet ist.

Kabelgebundene Lösungen: Obwohl kabelgebundene Lösungen stabile Verbindungen bieten, reduzieren sie drastisch die Mobilität des Roboters und erfordern komplexe Umbauten im Gebäude. Dies widerspricht den Zielen der einfachen und flexiblen Integration.

Nahfeldscanner: Technologien wie RFID und NFC bieten nur geringe Reichweite (wenige Zentimeter) und sind daher für großräumige, mobile Anwendungen ungeeignet. Eine unterbrechungsfreie Kommunikation ist nicht möglich.

4 Einsatzmöglichkeiten und Vorteile

Die entwickelte Lösung ermöglicht einfache physische Interaktionen des Telepräsenzroboters mit der Gebäudeinfrastruktur, wie beispielsweise:

- Öffnen von Türen und Toren
- Aktivieren von Lichtanlagen
- Steuerung weiterer KNX-fähiger Geräte

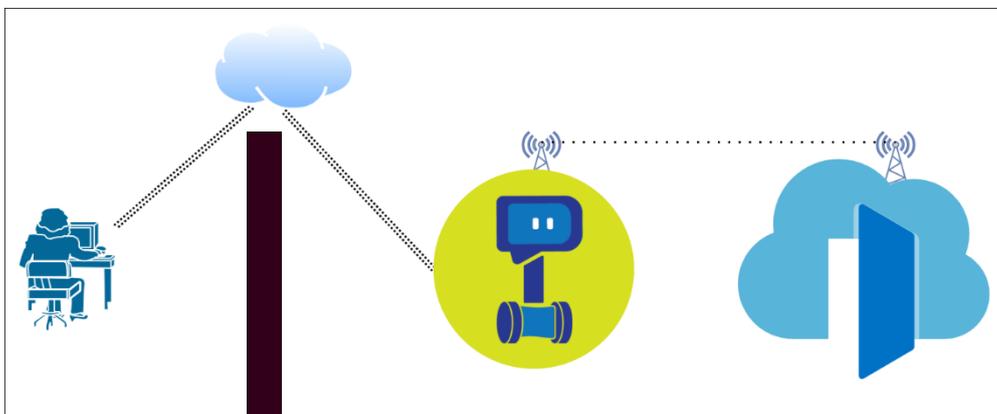


Abbildung 2: Schematische Darstellung des Einsatzfalls Türöffnung

Diese Funktionalitäten wären ohne die Integration nur durch teure mechanische Arme realisierbar oder gar nicht möglich. Somit reduziert diese

Lösung die Kosten erheblich und steigert gleichzeitig die Einsatzflexibilität des TPR.

Ein weiteres, komplexes Anwendungsszenario ist die Bedienung eines Fahrstuhls, um einem Telepräsenzroboter Zugang zu einem sicherheitsrelevanten Kontrollraum zu ermöglichen. Mithilfe der KNX-Integration kann der Roboter den Fahrstuhl eigenständig rufen und bedienen. Nach Aktivierung des KNX-gesteuerten Fahrstuhls fährt der Roboter autonom zur gewünschten Etage und kann anschließend durch KNX-gesteuerte Türen direkt in den Kontrollraum gelangen. Dieses Szenario unterstreicht den besonderen Mehrwert der entwickelten Lösung, da keine zusätzliche mechanische Hardware erforderlich ist und komplexe, sicherheitskritische Räume nun flexibel für TPR erschlossen werden können.

5 Praxistests und Validierung

In praktischen Tests wurde nachgewiesen, dass der Roboter zuverlässig Funktionen in der Gebäudeinfrastruktur aktivieren konnte. Die RF-Kommunikation erwies sich als stabil und robust gegenüber typischen Störungen in Industrieumgebungen.

Abbildung 3 und Abbildung 4 zeigen Szenen aus den Tests.

Ein Video von der Erprobung ist verfügbar unter <https://youtu.be/GkM66GymBuA?si=nDSwbvW AoMh0W86>.



Abbildung 3: Test des Einsatzfalls Aufzug



Abbildung 4: Test des Einsatzfalls Türöffnung

6 Zusammenfassung und Ausblick

Die erfolgreiche Integration eines TPR in das KNX-System mittels Rust und RF-Technologie bietet eine sichere, flexible und kosteneffiziente Lösung für die Gebäudeautomation. Zukünftige Entwicklungen könnten weitere Funktionen hinzufügen und die Kommunikation weiter optimieren, um noch anspruchsvollere Interaktionen zu ermöglichen.

Die Forschungs- und Entwicklungsarbeit wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ im Projekt PraeRI (Förderkennzeichen 02L21B000-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt der Veröffentlichung liegt beim Autor.

Autoren



von Lonski, Nikolas

Nikolas von Lonski ist als Software-Engineer bei der Actimage GmbH tätig und als leitender Entwickler verantwortlich für die Entwicklung innovativer Lösungen im Bereich embedded IoT. Im Forschungsprojekt PraeRI beschäftigte er sich intensiv mit der Integration von Telepräsenzrobotern in KNX-Systeme.

Chancen und Grenzen von Telepräsenzrobotern in der Industrie aus Anwendersicht

Sven Eichhorn¹, Florian Schnabel², Tom Heinicker², Georg Rudolph²

¹ligenium GmbH, Chemnitz

²HLS Robotic Automation GmbH, Meerane

Zusammenfassung

Aufbauend auf einer Darstellung der Chancen und Grenzen des Einsatzes aktueller Telepräsenzroboter (TPR) im industriellen Umfeld werden Lösungsvorschläge unterbreitet, mit deren Hilfe die Einsatzbereiche von TPR in der Industrie erweitert werden können.

1 Einleitung

TPR sind technische Systeme, die über das Internet durch TPR-Fahrende aus der Ferne und nahezu in Echtzeit gesteuert werden [1, 2, 3]. Dadurch wird eine aktive Teilnahme am „Hier und Jetzt“ möglich, ohne direkt vor Ort zu sein. Der Begriff „Roboter“ verleitet zur Annahme, dass eine mannigfaltige Interaktion mit der Umgebung möglich ist. De facto beschränken sich verfügbare TRP allerdings auf Funktionen eines „fahrbaren Tablets“ – bei einigen Modellen mit einer „Ablage für leichtes Transportgut“. Anwendungen liegen z. B. in der Pflege, im Krankenhaus oder in Bürobereichen.

Der Frage, welche Chancen und Grenzen sich für TPR im industriellen Umfeld ergeben, wird nachfolgend aus Sicht eines Systemintegrators für Industrieroboter (HLS Robotic Automation GmbH) und eines Produzenten für Logistiklösungen bzw. Maschinenelemente in Holzbauweise (ligenium GmbH) sowie deren Kundschaft aus Automobil- und Baubranche nachgegangen. Grundlage sind Tests mit den TPR Ohmni von Ohmnilabs und Double 3 von Double Robotics, die im Zeitraum 2022 bis 2024 im Projekt PraeRI erfolgten.

2 Chancen und Grenzen

2.1 Chancen

Grundlegend geht es bei den Chancen darum, Expertinnen und Experten oder Personen mit Entscheidungsbefugnis nicht mehr auf langwierige und kostspielige Geschäftsreisen zu schicken, sondern sie per TPR „just in time“ an entfernten Orten einsetzen zu können. Dadurch können sie in kürzerer Zeit verschiedenste Aufgaben wahrnehmen. Die Reisezeit verkürzt sich auf die Verbindungsdauer des TPR-Fahrenden zum Roboter. Der Mehrwert des TPR zur etablierten Videokonferenz besteht dabei in dessen Interaktionsfähigkeit. Der TPR-Fahrende kann sich autonom, d. h. theoretisch ohne Hilfestellung vor Ort, bewegen – soweit der TPR bereits am richtigen Ort ist.

Darauf aufbauend ergeben sich Vorteile wie

- Minimierung von Reisezeiten und -kosten,
- Erhöhung der Produktivität durch beschleunigten Projektablauf,
- parallele Betreuung von Projekten,
- schnelles, bedarfsweises Zuschalten von Fachkräften (z. B. bei Reparatur, Wartung- und Montagearbeiten),
- Durchführung von Schulungen und Monitoring-Aufgaben aus der Ferne,
- bessere Vereinbarkeit von Beruf und Privatem und
- Reduktion von reisebedingten Klimagas-Emissionen.

Als Vision des Einsatzes von TPR in der Industrie kann angesehen werden, dass der TPR eine speziell benötigte fachkundige und helfende Person vor Ort per Anruf ersetzen kann. Aus dieser hohen Erwartung ergibt sich eine lange Wunschliste an zusätzlichen Fähigkeiten und Funktionen von TPR-Lösungen für den industriellen Einsatz, da marktverfügbare TPR schnell an Grenzen stoßen.

2.2 Grenzen

Folgende Punkte begrenzen den Einsatz der marktverfügbaren TPR in der Industrie sehr deutlich.

1 Mangelnde Robustheit des TPR

1.1 Mangelhafte Fahreigenschaften bei Barrieren

1.2 Mangelhafter Schutz vor Verschmutzung, Vandalismus, Diebstahl

Die getesteten TPR sind für „Büroumgebungen“ konzipiert. Eine robuste Nutzung im Umfeld einer Fertigung oder Montageumgebung ist hinsichtlich der Fahrbarkeit bei Bodenunebenheiten, Schwellen, Kabeln und Schläuchen eingeschränkt. Aus der Erfahrung im Umgang mit Transport- und Transporthilfsmitteln sowie mit anderen Betriebsmitteln in der Industrie hielten die Praktiker, die die TPR testeten, auch den Schutz vor Verschmutzungen und anderer Einwirkungen als zu gering.

2 Fehlende oder eingeschränkte nutzbare Schnittstellen für Datenübertragung

Durch fehlende oder nicht offen zugängliche Schnittstellen ist eine zeitgemäße Kommunikation mit anderen Geräten nicht oder nur eingeschränkt möglich. Das Videobild lässt sich nicht ohne Weiteres auf andere Monitore oder Beamer duplizieren. Falls in einer lauten Umgebung der Ton auf Kopfhörer übertragen werden soll, ist nicht mit allen TPR-Modellen eine einfache Kopplung per Bluetooth möglich.

3 Fehlende Funktionen

3.1 keine oder beschränkte Transportfunktion

3.2 fehlende Aktorik

3.3 fehlende Möglichkeit, die Sichtbarkeit des TPR in Gefahrensituationen zu erhöhen

3.4 fehlende, sichere Verwahrmöglichkeit bei Langzeiteinsatz

Die Personen, die in den Tests mit den TPR interagierten, äußerten wiederholt Wünsche nach zusätzlichen Funktionen. Mitarbeitende vor Ort konnten sich vorstellen, dass der TPR auch Arbeitsmittel transportiert und ggf. als „hilfende“ Hand unterstützt. TPR-Fahrende wünschten sich eine verbesserte Interaktion, indem sie vor Ort etwas zeigen oder sogar bedienen können. Wenn sich die Test-Nutzenden vorstellen sollten, den TPR längere Zeit bei einem Kunden oder einer Kundin einzusetzen, kam auch die Frage auf, wie der TPR, wenn er nicht genutzt wird, z. B. in einer Werkhalle sicher verwahrt werden kann.

4 Mangelnde Bildqualität

4.1 mangelhafte Bildqualität/Auflösung der Kamera für Detailansichten

4.2 schlechte Bildqualität am Display bei schwierigen Lichtverhältnissen

Während der Tests gab es immer wieder Herausforderungen mit der Bildqualität – vor allem wenn die Umgebung nicht nur für das sichere Fahren des TPR, sondern auch für andere Zwecke observiert werden musste, z. B. um in Prüfprozessen kleine Details wie Schrift oder Schrauben zu erkennen. Weitere Herausforderungen stellen sich bei ungünstigen Lichtverhältnissen, etwa durch Schattenwürfe, Blendung und ungleichmäßige Belichtung/Beleuchtung.

5 Schlechte Verbindungsqualität

Grundsätzlich benötigt der TPR zum Fahren und für den Video-Call eine permanent stabile Internetverbindung mit angemessener Bandbreite. Die Erprobungen bei den Anwendungspartnern und bei deren Kundinnen und Kunden zeigten, dass diese Konnektivität nicht an allen Standorten gegeben ist. Das lag teils an der lokalen Verfügbarkeit des Internets (im ländlichen Raum), aber auch an der jeweiligen innerbetrieblichen WLAN-Ausleuchtung. Die Probleme verschärfen sich, wenn die Bandbreite der Internetverbindung durch Zusatzfunktionen beansprucht wird – etwa durch eine zusätzliche Satelliten-Kamera. Diese wurde genutzt, um für Inspektionszwecke Details an Maschinen und Ausrüstungen aufzunehmen, die von der im TPR integrierten Kamera nicht erfasst werden konnten, da sie sich z. B. in Bodennähe oder in verdeckten Positionen befanden. Bei schlechter Internetverbindung und Betrieb der Zusatzkamera verringert sich die Bildqualität des Video-Calls drastisch.

6 Mangelhafte Daten- und Informationssicherheit

6.1 Sicherheitsprobleme aufgrund Kamerafunktion

6.1.1 Geheimhaltung (z. B. Betriebsgeheimnisse)

6.1.2 Persönlichkeitsrechte (z. B. Recht am eigenen Bild)

6.2 IT-Sicherheitsprobleme

Das Thema Daten- und Informationssicherheit bzw. informelle Selbstbestimmung ist beim Einsatz von TPR auf Grund der über das Internet übertragenen Videodaten und die standardmäßig über die Server der TPR-Hersteller laufenden Verbindungen herausfordernd, weil damit gesetzliche Anforderungen und Normen nicht oder nur unsicher eingehalten werden. In Gesprächen mit Kundinnen und Kunden aus der Automobil- und Automobilzuliefererindustrie wurde deutlich, dass dabei die Branchennorm Trusted Information Security Assessment Exchange (TISAX) gegenüber gesetzlichen Vorschriften und der internationale Norm ISO/IEC 27001 zu Informationssicherheits-Management-Systemen noch einmal verschärfte Anforderungen stellt. Außerdem besteht eine hohe Sensibilität von Betriebsräten, die Persönlichkeitsrechte von Beschäftigten zu schützen.

3 Lösungsansätze

Die übergeordneten Zielstellungen für die Ausführung eines TPR zum vorteilhaften Einsatz in der Industrieanwendung sind

- eine robustere Bauweise,
- passende Schnittstellen und Funktionen, mit denen Mehrwerte geschaffen werden können, sowie
- eine niedrige Integrationsschwelle in vorhandene sicherheitsgerechte IT-Infrastrukturen sowie der datenschutzgerechte Betrieb.

Für die einzelnen Herausforderungen existieren folgende Lösungsmöglichkeiten:

1 Mangelnde Robustheit des TPR

1.1 Mangelhafte Fahreigenschaften bei Barrieren

1.2 Mangelhafter Schutz vor Verschmutzung, Vandalismus, Diebstahl

Das Fahrwerk und Chassis sollten – vergleichbar zu fahrerlosen Transportsystemen – robuster, aber trotzdem leicht, ausgeführt werden, um einen langlebigeren, stoßunempfindlicheren und verschmutzungsresistenten Betrieb zu ermöglichen. Ein besonderer Fokus liegt auf dem An- oder Überfahren von Kanten bzw. Schwellen mit größeren und stabileren Rädern, dem Schutz bei Kontakt mit anderen Fahrzeugen der Intralogistik (z. B. Stapler) und einer gewissen „Geländegängigkeit“ (z. B. bei Baustellenzustand, unebener Boden). Eine geschlossene, in sich stabile, aber federnde Struktur sollte die Elektronik vor äußeren dynamischen Einflüssen und Medien schützen und einen gewissen Abnutzungsvorrat gegenüber Vandalismus bieten.

2 Fehlende Schnittstellen für Datenübertragung

In Verbindung und als Voraussetzung für Punkt 3 werden zusätzliche oder, bei einigen TPR-Ausführungen, auch nur einfacher zugängliche Schnittstellen für die Datenübertragung benötigt. Prioritär erscheinen folgende Schnittstellen für folgende Funktionen:

- einfach zu aktivierendes Bluetooth zur Verbindung mit Kopfhörern und anderen Endgeräten,
- einfach und frei wählbare Art der Internetverbindung (verschiedene WLAN und Mobilfunknetze),
- HDMI, Display-Port oder frei auswählbares WLAN, um einen weiteren Videoausgang (z. B. Monitor, Beamer) nutzen zu können.

3 Fehlende Funktionen

3.1 keine oder beschränkte Transportfunktion

Gerade bei Montage- bzw. Wartungseinsätzen bietet sich eine Erweiterung von TPR zu mobilen Serviceplattformen an, die auch nennenswerte Lasten (z. B. Werkzeuge, Prüfmittel, Material, ca. 100 kg) transportieren können. Die Serviceplattform wird somit zu einem mitfahrenden, ggf. sogar „handreichenden“ (s. u.) Helfer oder wahlweise zu einem mit Expertise unterstützenden „TPR-Kollegen“. Zusätzlich wäre für die Serviceplattform auch die Integration von Software-Assistenten wie dem in [5] vorgestellten Protokoll-Assistent und Assembly-Tutor hilfreich.

3.2 fehlende Aktorik

Um die derzeit fehlende Aktorik des TPR zu adressieren, können zwei Lösungsansätze verfolgt werden:

Der TPR kann mit zusätzlicher *eigener* Aktorik ausgestattet werden. Hierfür ist der in [4] beschriebene Laserpointer für eine optische Zeigefunktion prioritär. Er erlaubt eine natürlichere Kommunikation, indem der TPR-Fahrende der Person vor Ort etwas zeigen kann. Darauf aufbauend sind Lösungen zur haptischen Interaktion (z. B. mit Aufzug-Ruftaste, Türklinken, Not-Aus-Schalter) und Manipulatoren und/oder Greifer wünschenswert, die einer „helfenden“ Hand nahekommen. Solche von einem Menschen gesteuerten Roboterarme werden bereits in der Reha-Technik eingesetzt.

Der TPR kann in die Lage versetzt werden, *fremde* Aktoren anderer Systeme – insbesondere der Gebäudetechnik – anzusteuern. Dies wurde prototypisch für die Ansteuerung von Automatik-Türen über einen KNX-Bus umgesetzt [6].

3.3 fehlende Möglichkeit, die Sichtbarkeit des TPR in Gefahrensituationen zu erhöhen

TPR könnten über Warnleuchten und ggf. akustisch ihre Position, das Aktivitätslevel und den Zustand des TPR anzeigen. Dies erscheint besonders beim Einsatz in Hallen mit Staplerverkehr von Bedeutung.

3.4 fehlende, sichere Verwahrmöglichkeit bei Langzeiteinsatz

Der TPR kann in den Zeiten, in denen er nicht aktiv genutzt wird, durch eine Verriegelung in der Ladestation bzw. durch eine mobile, aber stabile Umhausung („Garage“, Shelter) geschützt werden. Die Ladestation oder die Umhausung sollte durch ein Bügelschloss oder eine Kette mit einer festen baulichen Einrichtung (z. B. Geländer, Stütze, Regal) verbunden werden.

4 Mangelnde Bildqualität

4.1 mangelhafte Bildqualität/Auflösung der Kamera für Detailansichten

Durch eine Kombination aus einer (Haupt-)Kamera mit besserer Auflösung und einer zusätzlichen, externen, zuschaltbaren Satelliten-Kamera, die von der vor Ort befindlichen Person von Hand positioniert werden kann, können die Anforderungen an die Bildqualität adäquat erfüllt werden. Die Satellitenkamera wurde im Projekt experimentell erprobt und die Machbarkeit nachgewiesen. Jedoch ist die softwareseitige Integration zur Steuerung der Kamera durch den TPR-Fahrenden zu verbessern.

4.2 schlechte Bildqualität am Display bei schwierigen Lichtverhältnissen

Eine durch den TPR-Fahrenden zuschaltbare Beleuchtung durch LED-Ringe oder verteilt auf mehrere, bewegliche Strahler (vgl. Laserpointer [4]) könnte dazu beitragen, die Ausleuchtung des von der Kamera erfassten Bereichs unabhängig von den Umgebungsbedingungen zu verbessern.

5 Schlechte Verbindungsqualität

Der TPR sollte die notwendige Konnektivität vorzugsweise „mitbringen“, um im Zweifel von der IT-Infrastruktur des Anwendungsortes unabhängig zu sein. Dazu sollte ein eigener mobiler WLAN-Hotspot, idealerweise mit mehreren SIM-Steckplätzen für Karten verschiedener Mobilfunkanbieter, zum Einsatz kommen.

6 Mangelhafte Daten- und Informationssicherheit

6.1 Sicherheitsprobleme aufgrund Kamerafunktion

6.1.1 Geheimhaltung (z. B. Betriebsgeheimnisse)

6.1.2 Persönlichkeitsrechte (z. B. Recht am eigenen Bild)

6.2 IT-Sicherheitsprobleme aufgrund Datenübertragung per Internet und Datenspeicherung auf Servern des TPR-Herstellers bzw. in der Cloud

Für die Verbesserung der Informationssicherheit und die formale Absicherung der Konformität sind mehrere Lösungsansätze zu verfolgen.

Die TPR-Hersteller selbst müssen die DSGVO-Konformität sicherstellen. Das erfordert u. a. mindestens ein Hosting in Europa. Darüber hinaus sollten die TPR-Hersteller das Self-hosting von TPR durch die Käufer/Betreiber erleichtern.

In und zwischen Organisationen, in denen der TPR zum Einsatz kommt, sind mindestens folgende Vereinbarungen zu empfehlen:

- Einverständniserklärung aller Personen, die im Kamerabild erscheinen könnten, dass ihr Bild für den Zweck des TPR-Einsatzes übertragen werden darf,
- Einverständniserklärung der Organisation, in der der TPR-Einsatz erfolgt (ggf. mit Spezifizierung örtlicher Einsatzbereiche) und in Verbindung mit
- Geheimhaltungs-Vereinbarung (ggf. mit Spezifizierung zur Art der Datenübertragung und Datenspeicherung).

Sofern der TPR in der Konstellation zum Einsatz kommt, dass Mitarbeitende eines Ausrüsters/Dienstleisters den TPR bei einem Kunden oder einer Kundin fährt, wäre folgende Option interessant: Der Kunde bzw. die Kundin – insbesondere, wenn es sich um ein größeres Unternehmen handelt, dass viele Ausrüster/Dienstleister beauftragt – beschafft und hostet einen bzw. mehrere TPR selbst und betreibt diese sicherheitskonform in der eigenen IT-Infrastruktur. Der TPR wird dann externen Dienstleistenden gegen Entgelt oder im Rahmen von vertraglichen Vereinbarungen bedarfsweise zur Verfügung gestellt.

4 Zusammenfassung und Ausblick

TPR bieten in der Industrie gerade in Kooperationen zwischen Ausrüstern, Dienstleistenden und Anlagenbetreibern ein hohes Potenzial, durch eine telepräsenste Fernunterstützung an Reaktionsschnelligkeit zu gewinnen und Geschäftsreisen einzusparen. Akteure aus der Praxis, die im Projekt PraeRI marktverfügbare TPR getestet haben, bestätigen das Potenziale.

Gleichzeitig wurde in den Tests sichtbar, dass die verfügbaren TPR über die Anwendung als „Besprechungsroboter“ hinaus im industriellen Umfeld nur beschränkt eingesetzt werden können. Konkrete Einsatzgrenzen wurden identifiziert und Lösungsansätze vorgeschlagen. Erste Softwareanpassungen und Funktionserweiterungen sind im Rahmen des Projektes bereits erfolgt und in [4, 5, 6] beschrieben.

TPR-Hersteller sollten die Vorschläge und prototypischen Lösungen aufgreifen und in marktfähige Produkte umsetzen.

Die Forschungs- und Entwicklungsarbeit wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ in den Projekten PraeRI (Förderkennzeichen 02L21B000-4) sowie TeleInteraction-XR (Förderkennzeichen 02L21B550-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt der Veröffentlichung liegt bei den Autoren.

5 Literaturverzeichnis

- [1] Birke et al. (2023) Entwicklung eines Erprobungskonzeptes für den Einsatz von Telepräsenzrobotern GfA, Sankt Augustin (Hrsg.): Frühjahrskongress 2023, Hannover B.6.2 Nachhaltig Arbeiten und Lernen - Analyse und Gestaltung lernförderlicher und nachhaltiger Arbeitssysteme und Arbeits- und Lernprozesse
- [2] Hernandez, F., Birke, J., Bullinger, A.C. (2023). The Tribid-Meeting-Setup – Improving Hybrid Meetings Using a Telepresence Robot. In: Streitz, N.A., Konomi, S. (eds) Distributed, Ambient and Pervasive Interactions. HCII 2023. Lecture Notes in Computer Science, vol 14037. Springer, Cham. https://doi.org/10.1007/978-3-031-34609-5_2
- [3] Hernandez, F., Ruffert, D., Hoffmann, H., Bullinger, A.C. (2024). Vom Pixel zur Präsenz - Ferninteraktion mit Telepräsenzrobotern. Industry 4.0 Science, Bd. 2024, Nr. 5, Sep. 2024, doi: 10.30844/I4SD.24.5.18.
- [4] Venepally, R., Hernandez, F., Ruffert, D. (2025). Kannst Du mir das zeigen? Entwicklung und Evaluation einer Zeigefunktion für einen TPR. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [5] Eichhorn, E., Schmidt, G., Stüring, S., Venepally, R. (2025). Erweiterung der Softwarefunktionalität von TPR zum Einsatz in der Industrie. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [6] von Lonski, N. (2025). Roboter öffnen Türen – Neue Wege der Telepräsenz. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.

Autoren



Eichhorn, Sven

Dr.-Ing. Sven Eichhorn beschäftigt sich seit 2005 mit dem vorteilhaften Einsatz von Holz im Maschinenbau. Er hat 2018 die ligenium GmbH mitgeründet und ist seit 2024 Leiter des Bereiches Forschung und Technologie bei ligenium.



Georg Rudolph

Georg Rudolph arbeitet zwischen 2020 und 2024 als Leiter der Robotic Automation GmbH. In dieser Zeit war er als Projektkoordinator maßgeblich am Forschungsprojekt PraeRI beteiligt. Nun ist er als Geschäftsführer für die Autania Services GmbH im Bereich Fertigung elektrischer Schaltanlagen tätig.



Tom Heinicker

Tom Heinicker arbeitet zwischen 2020 und 2024 als Projektleiter von Automatisierungsprojekten für die HLS Robotic Automation GmbH. Im Forschungsprojekt PraeRI leitet er die Erprobungen im industriellen Umfeld.



Florian Schnabel

Florian Schnabel arbeitet seit 2023 für die HLS Robotic Automation GmbH. Er ist für die Abwicklung von Automatisierungsprojekten als Projektleiter verantwortlich. Im Rahmen des Forschungsprojektes PraeRI hat er seine Diplomarbeit zum Thema Telepräsenzroboter geschrieben.

Erweiterung der Softwarefunktionalität von TPR zum Einsatz in der Industrie

Sven Eichhorn¹, Georg Schmidt², Stefan Stüring², Rakshith Venepally²

¹ligenium GmbH, Chemnitz ²LSE Software & Engineering GmbH, Magdeburg

Zusammenfassung

Der Beitrag skizziert, wie durch Erweiterung der Softwarefunktionalität des Telepräsenzroboters (TPR) Ohmni eine Verbesserung der Leistungsfähigkeit von TPR in der industriellen Praxis erreicht werden kann. Unter anderem wurden eine Objekterkennung, eine Protokollfunktion und Tutor-Funktion integriert.

1 Einleitung

Die Grundfunktionalitäten von TPR sind auf die ferngesteuerte Bewegung und eine Interaktion per Display des Fahrenden mit der Umgebung am Zielort reduziert [1, 2, 3, 4]. Um TPR besser im industriellen Umfeld zu nutzen, wurden Bedarfe und potenzielle Lösungsansätze für zusätzliche Funktionen gesammelt [1]. Dabei spielte die Perspektive eines Produzenten für Logistiklösungen bzw. Maschinenelemente und Maschinen in Holzbauweise – des Projektpartners ligenium GmbH – und dessen Kundschaft aus der Automobil- und Baubranche eine zentrale Rolle.

Aus der Vielfalt der Bedarfe werden nachfolgend ausgewählte Funktionen betrachtet, für die allein auf Basis von Softwareerweiterungen die Einsetzbarkeit von TPR in der Industrie verbessert werden kann. Es wurden Funktionen fokussiert, für die mit den Möglichkeiten des Projektteams prototypische Lösungen realisiert werden konnten.

Dabei handelt es sich, neben dem Softwaretreiber für eine Zeigefunktion, die in einem separaten Beitrag [5] beschrieben ist, um folgende Softwareerweiterungen:

1. automatisierte Objekterkennung,
2. digitaler Assistent für Qualitätsprozesse und
3. digitaler Assistent / Tutor für Montageaufgaben.

Die Bedarfsanalysen basieren auf 2022 bis 2024 durchgeführten Tests mit dem TPR Ohmni von Ohmnilabs. Für diesen TPR wurden auch die Demonstratoren entwickelt.

2 Objekterkennung

Mit der automatisierten Objekterkennung waren zwei Ziele verbunden:

- Sie sollte eine automatisierte Navigation ermöglichen, die Hindernisse im Fahrweg des TPR intelligent meidet, und die es erlaubt, dass der Roboter mit seiner Umgebung intelligent interagiert.
- Sie sollte mit einer Anonymisierung von Objekten und Menschen verknüpft werden können, um einen besseren Schutz von Persönlichkeitsrechten und Betriebsgeheimnissen zu gewährleisten.

Für die technische Umsetzung der automatischen Objekterkennung wurde das Konzept verfolgt, den Video-Call des TPR mit Hilfe der sogenannten YOLO-Technologie zu analysieren. YOLO ist ein erstmals von Redmon et al. [6] vorgestelltes Framework für Objekterkennungssysteme, das auf neuronalen Faltungsnetzen basiert und das die Wahrscheinlichkeit ermittelt, mit der sich ein Objekt zunächst in Vierecken eines regelmäßig über ein Bild gelegten Rasters und davon abgeleitet in Begrenzungsrechtecken (Bounding Boxes) befindet.

Ein YOLO-Netzwerk konnte erfolgreich auf dem TPR implementiert werden (vgl. Abbildung 1). Die funktionierende YOLO-Objekterkennung eröffnet die Möglichkeit, darauf aufbauend weitere der oben genannten Funktionalitäten zu realisieren. Im Rahmen des Projektes wurden dafür mögliche technische Konzepte eruiert, deren Umsetzung erfordert jedoch noch weitere Forschungsarbeit:

- Interaktion mit Hindernissen: Die YOLO-Objekterkennung könnte in Verbindung mit einem Lidar-Sensor eingesetzt werden, um es dem TPR zu ermöglichen, automatisch Hindernisse zu erkennen und zu umfahren. Ein Einsatzfall wäre das Erkennen von gesperrten Verkehrswegen: Wenn die Objekterkennung einen Verkehrsleitkegel, ein Absperrband oder ähnliches detektiert, meidet der TPR diesen Bereich und findet automatisiert eine Alternativroute. Eine Voraussetzung dafür wäre, dass die Software die Objekte und die Verhaltensalternativen durch eine entsprechende Programmierung oder ein Training kennt.
- Schutz von Persönlichkeitsrechten und Betriebsgeheimnissen: Mit einer ähnlichen Logik könnte eine Software die von dem YOLO-Netzwerk erkannten Objekte (und Personen) bezüglich ihrer Schutzwürdigkeit interpretieren und zum Beispiel automatisiert Unschärfefilter in den Video-Stream des TPR einfügen.

In beiden Fällen ist es zu empfehlen, die von dem YOLO-Netzwerk berechnete Wahrscheinlichkeit, dass sich ein Objekt in der Bounding Box befindet, bei der weiteren Interpretation und Ableitung von Handlungen spezifisch zu berücksichtigen: In Anwendungsfällen, in denen der Schutz von Persönlichkeitsrechten im Vordergrund steht, sollten auch mit geringer Wahrscheinlichkeit als Personen erkannte Bildausschnitte vorsorglich mit einem Unschärfefilter unkenntlich gemacht werden. Dagegen würde eine weiträumige und zeitaufwändige Umfahrung eines Hallenbereichs eher

erst dann initiiert werden, wenn das YOLO-Netzwerk mit hoher Wahrscheinlichkeit ein Absperrungsobjekt erkannt hat.



Abbildung 1: Anwendung der YOLO-Objekterkennung auf dem Ohmni TPR

3 Digitaler Protokoll-Assistent

Marktverfügbare TPR unterstützen zwar die Kommunikation zwischen remote arbeitenden Expertinnen und Experten und Vor-Ort-Mitarbeitenden in Bild und Ton. Sobald diese Kommunikation oder sich daraus ergebende Handlungen jedoch dokumentiert bzw. protokolliert werden müssen, sind dafür zusätzliche Endgeräte bzw. Papier und Stift nötig. Protokolle sind in der Industrie zum Beispiel bei Abnahmen und anderen Prüfprozessen erforderlich. Solche Prüfprozesse stellen geeignete Anwendungsfälle für die Fernunterstützung durch TPR dar.

Daher wurde eine Software-Erweiterung zur Unterstützung der Protokollierung von Qualitätsprozessen entwickelt und getestet. Zu Grunde gelegt wurden beispielhaft Qualitätsprozesse der Igenium GmbH.

Die Umsetzung erfolgte als Android-App. Dies hat den Vorteil, dass die Anwendung – wie gewünscht – direkt auf dem TPR aber auch auf Mobiltelefonen laufen kann. Letztere stehen den Vor-Ort-Mitarbeitenden in den meisten Fällen zur Verfügung, während TPR eher nur punktuell im Einsatz sein werden.

Die App kann parallel zum Video-Call auf dem TPR eingesetzt werden und ermöglicht die papierlose Dokumentation der durchgeführten Arbeiten in einem speziell für die Arbeitsaufgabe zugeschnittenen Protokoll. Das Protokoll ist sowohl durch den Remote-Experten oder die Remote-Expertin als auch durch die Vor-Ort-Mitarbeitenden

editierbar. Per Cloud werden die Daten synchron gehalten. Das Protokoll ist individuell anpassbar und kann für Prüfungen bei verschiedenen Montage-, Wartungs- bzw. Reparaturprozessen bearbeitet werden (vgl. Abbildung 3).

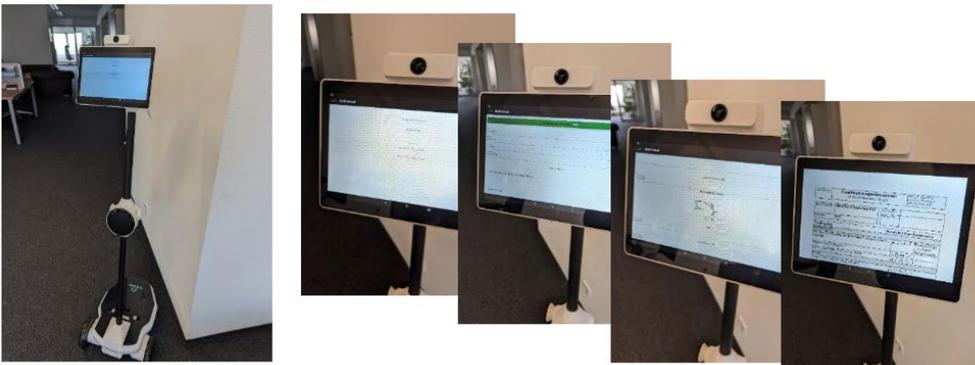
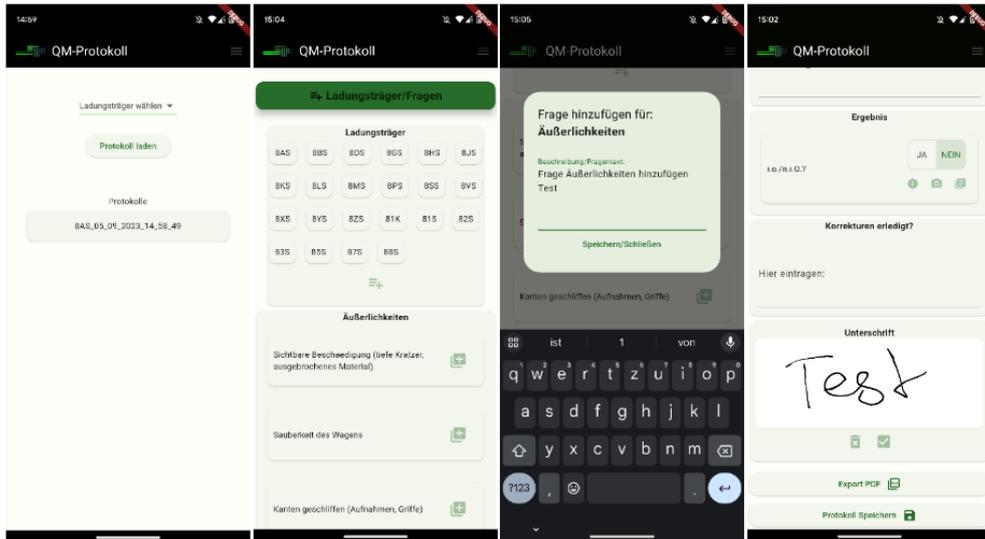


Abbildung 2: QM-App für Android auf dem Mobiltelefon und dem Ohmi TPR

Der Protokoll-Assistent bzw. die App sind für folgende Szenarien nutzbar:

Szenario A: Die Remote-Fachkraft unterstützt in Vorbereitung und /oder während des QM-Prozesses durch die Erstellung und Vorgabe des Protokolls, welches von der vor Ort arbeitenden Person genutzt und abgearbeitet wird. Das Protokoll wird auf dem TPR oder dem Mobiltelefon bereitgestellt und schrittweise ausgefüllt.

Szenario B: Die Remote-Fachkraft und die vor Ort arbeitende Person interagieren und kommunizieren mit TPR und ggf. dem Mobiltelefon live, um schnelle Anpassungen am Protokoll gemeinsam durchzuführen.

Szenario C: Die vor Ort arbeitende Person generiert das Protokoll nach vorhandener, abgeänderter Sachlage selbst und aktualisiert damit das Wissen der Remote-Fachkraft.

Szenario A und B können durch nachfolgend beschriebene Assembly Tutor Web-App bereichert werden.

4 Assembly Tutor

Ähnlich wie beim oben beschriebenen Protokoll-Assistent treten bei der Fernunterstützung von Montage-, Wartungs- und Reparaturaufgaben auch Situationen auf, in denen der Video-Call des TPR allein nicht hinreichend geeignet erscheint. Vielmehr wären schrittweise Anleitungen mit Bildunterstützung, Checklisten und ähnlichem wünschenswert. Dazu wurde prototypisch eine Assembly Tutor Web-App entwickelt.

Das Anwendungsszenario ist, dass bisher ungeschultes Vor-Ort-Personal durch den Remote-Experten oder eine Remote-Expertin befähigt wird, unbekannte Montageaufgaben inklusive nahtloser Dokumentation der durchgeführten Arbeiten durchzuführen. Für die Protokollierung ist oben dargestellte Protokoll-App nutzbar. Für den Assembly Tutor wird die Arbeitsfolge in Vorbereitung der Fernunterstützung auf Basis vorhandener CAD-Daten und der Fachexpertise erstellt und vor Ort „nur“ abgerufen – als Ergänzung und zur besseren Erklärung im Video-Call oder auch unabhängig vom Video-Call für ein Selbststudium. Die webbasierte Tutor-App kann auf jedem Gerät mit Browser ausgeführt werden.

Abbildung 4 zeigt aus CAD-Daten generierte Szenen, die für den Assembly Tutor App verfügbar gemacht wurden.



Abbildung 3: CAD-basierte Szenen für die Assembly Tutor Web-App am Beispiel des Ladungsträgers 8GS der ligenium GmbH

Abbildung 5 zeigt einen Auszug aus einem Workflow mit verschiedenen Funktionalitäten der Assembly Tutor Web-App.



Abbildung 4: Assembly Tutor Web-App – Auszug aus einem Workflow am Beispiel des Ladungsträgers 8GS der ligenium GmbH und ausgewählte Funktionalitäten

5 Zusammenfassung und Ausblick

Die beschriebenen Erweiterungen der Softwarefunktionalität des TPR ermöglichen in Summe eine Verbesserung der Leistungsfähigkeit und der Nutzungserfahrung in der industriellen Praxis. Der TPR wird mehr zum digitalen Helfer.

Aus der automatischen Objekterkennung resultiert eine verbesserte Navigation und ein verbesserter Datenschutz. Der Protokoll-Assistent für Qualitätsbelange erweitert – mit und ohne die Lernumgebung (webbasierte Tutor-App) – die Einsatzmöglichkeiten und steigert die Reaktions-Geschwindigkeit der Mitarbeitenden vor Ort.

Zukünftig ist es wichtig, die Hardware des TPR in Symbiose mit der vorgeschlagenen Software gezielt weiter zu entwickeln, um den digitalen Helfer weiter für seine angedachten Aufgaben zu qualifizieren.

Die Forschungs- und Entwicklungsarbeit wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ im Projekt PraeRI (Förderkennzeichen 02L21B000-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt der Veröffentlichung liegt bei den Autoren.

6 Literaturverzeichnis

- [1] Eichhorn, E., Schnabel, F., Heinicker, T., Rudolph, G. (2025). Chancen und Grenzen von Telepräsenzrobotern in der Industrie aus Anwendersicht. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [2] Birke, J. et al. (2023) Entwicklung eines Erprobungskonzeptes für den Einsatz von Telepräsenzrobotern GfA, Sankt Augustin (Hrsg.): Frühjahrskongress 2023, Hannover B.6.2 Nachhaltig Arbeiten und Lernen - Analyse und Gestaltung lernförderlicher und nachhaltiger Arbeitssysteme und Arbeits- und Lernprozesse
- [3] Hernandez, F., Birke, J., Bullinger, A.C. (2023). The Tribrid-Meeting-Setup – Improving Hybrid Meetings Using a Telepresence Robot. In: Streitz, N.A., Konomi, S. (eds) Distributed, Ambient and Pervasive Interactions. HCII 2023. Lecture Notes in Computer Science, vol 14037. Springer, Cham. https://doi.org/10.1007/978-3-031-34609-5_2
- [4] Hernandez, R., Rueffert, D., Hoffmann, H., Bullinger, A.C. (2024): Vom Pixel zur Präsenz Ferninteraktion mit Telepräsenzrobotern Industry 4.0 SciencenAusgaben40. Jahrgang, 2024, Ausgabe 5, Seite 18-25 <https://doi.org/10.30844/I4SD.24.5.186>
- [5] Venepally, R., Hernandez, F., Ruffert, D. (2025). Kannst Du mir das zeigen? Entwicklung und Evaluation einer Zeigefunktion für einen TPR. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenzte Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [6] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. <https://arxiv.org/abs/1506.02640>

Autoren



Eichhorn, Sven

Dr.-Ing. Sven Eichhorn beschäftigt sich seit 2005 mit dem vorteilhaften Einsatz von Holz im Maschinenbau. Er hat 2018 die ligenium GmbH mitgeründet und ist seit 2024 Leiter des Bereiches Forschung und Technologie bei ligenium.

Schmidt, Georg

Georg Schmidt hat seine Masterarbeit in Zusammenarbeit mit der LS Software & Engineering GmbH erarbeitet und war seit 2023 im Projekt mit der APP-Entwicklung betraut.

Stüring, Stefan

Stefan Stüring arbeitet seit mehr als 20 Jahren im Bereich Simulation und Visualisierung für Anwendungen in der Produktion und hat 2020 die LS Software & Engineering GmbH als SpinOff aus der LIVINGSOLIDS GmbH gegründet. Das Unternehmen beschäftigt sich mit der Entwicklung innovativer Visualisierungs- und Simulationslösungen für produzierende Unternehmen.



Venepally, Rakshith

Rakshith Venepally studierte Maschinenbau in Indien und schloss 2024 sein Masterstudium in Digital Engineering an der Otto-von-Guericke-Universität Magdeburg ab. Seit 2022 arbeitet er bei LS Software & Engineering GmbH mit den Schwerpunkten Visualisierung, Digital Twins und einer Masterarbeit über die Erweiterung eines Telepräsenzroboters.

Wer und was ist wo? – Synchronisation von Virtueller und Erweiterter Realität

Alexander Kögel¹

¹TU Chemnitz – Professur Arbeitswissenschaft und Innovationsmanagement

Zusammenfassung

Virtual- und Augmented Reality (VR / AR) werden in der Industrie bislang meist als alleinstehende Systeme genutzt. Damit AR-Anwendende im Feld sich von Experten bzw. Expertinnen, die sich an einem entfernten Ort in einer VR befinden, unterstützen lassen können, müssen beide Welten kontinuierlich synchronisiert werden. Dafür sind verschiedene Konzepte nutzbar.

1 Einleitung

Die VR- und AR-Technologien erfuhren in den letzten Jahren bedeutende Fortschritte: Immer schnellere und dabei zunehmend mobile und energieeffiziente Rechnerhardware erlaubte deutliche Steigerungen hinsichtlich der Auflösung, Anzahl der darstellbaren Polygone und der erzielbaren Bildwiederholraten. Auch Eingabe- und Interaktionsmethoden wurden stetig verbessert – etwa durch das Inside-Out Tracking ohne externe Sensoren, spezielle Marker und das automatische Erfassen von Hand- und Fingerpositionen. Dies erlaubt es, Gesten und Bewegungen des menschlichen Körpers – bis hin zur Mimik der Nutzenden – ohne störende Zusatzhardware wie Datenhandschuhe oder in den Händen getragene Controller zu erfassen. VR- und AR-Systeme können heute mobil und autark, ohne Anbindung an ortsfeste Rechnersysteme betrieben werden. Gleichzeitig stieg die Verbreitung solcher Systeme durch immer erschwinglichere, auch für den Consumer erhältliche, Produkte.

Im industriellen Kontext eignet sich VR insbesondere für Aufgaben im Engineering; AR ist für Anwendungen im Feld (z. B. Bei der Wartung oder Inbetriebnahme von Maschinen und Anlagen) prädestiniert.

2 VR/AR-Kopplung zur Fernunterstützung

Für die Fernunterstützung von Personen, die im Feld z. B. eine Wartung an einer Maschine ausführen, durch Expertinnen oder Experten, die an einem entfernten Standort über Wissen und digitale Daten zu der Maschine sowie eine VR-Umgebung verfügen, bietet es sich an, VR- und AR-Systeme miteinander zu koppeln. Das ermöglicht eine gemeinsame Interaktion in einer hybriden Umgebung. Dabei können Avatare eine entscheidende Rolle spielen, da sie es den Nutzenden erlauben, sich gegenseitig so wahrzunehmen und miteinander zu interagieren, als wären sie physisch im selben.

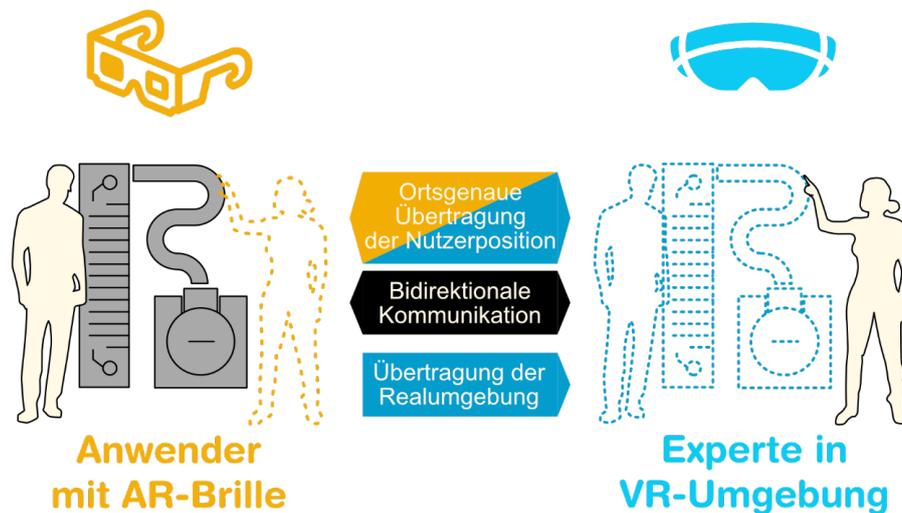


Abbildung 1: Konzept der VR-AR-Kopplung im Projekt TeleInteraction-XR

Die Person im Feld sieht dabei die reale Umgebung, in welcher das Abbild der Person aus der VR als virtueller Avatar dargestellt ist. Die Person, die das VR-System nutzt, sieht sowohl die Person im Feld als auch deren Umgebung (bzw. den für die Unterstützung notwendigen Teil dieser Umgebung) als virtuelle Abbilder. Die Positionierung der Personen untereinander, als auch relativ zur Realumgebung, wird dabei erfasst und bidirektional übertragen. So können sich beide Personen zielgerichtet bestimmten Orten (Points of Interest, POI) zuwenden, etwa einem Bedienpanel einer Anlage, und dort mittels Gesten auf ein Objekt zeigen. Das jeweilige Gegenüber kann lage- und ausrichtungsrichtig das Ziel der Geste wahrnehmen.

3 Eigenschaften von VR, AR und Avataren

Systeme der **Virtuelle Realität (VR)** umgeben den Nutzer vollständig mit einer digitalen Umgebung und schotten ihn von der realen physischen Welt ab. Sie zeichnen sich durch folgende Eigenschaften aus:

Virtualität: VR erzeugt eine umfassend computergenerierte Umgebung, in die Nutzende vollständig eintauchen kann (Immersion). Diese virtuelle Welt kann reale oder fiktive Szenarien darstellen und bietet eine hohe Gestaltungsfreiheit. Die Virtualität erlaubt es, beliebige Szenarien zu kreieren, sei es die Nachbildung realer Orte oder fantastischer Welten, was eine hohe Erlebnisvielfalt ermöglicht.

Interaktion: Nutzende können mit der virtuellen Umgebung und den darin befindlichen Objekten interagieren. Diese Interaktionen sind essenziell für das Gefühl der Präsenz in der virtuellen Welt. Die virtuelle Umgebung reagiert dabei dynamisch auf die Handlungen der Nutzenden, wodurch eine wechselseitige Beeinflussung zwischen Nutzenden und Umgebung entsteht.

Echtzeit: Für ein immersives VR-Erlebnis ist es notwendig, dass alle Interaktionen und Bewegungen in Echtzeit verarbeitet und dargestellt werden. Verzögerungen (Latenzen) können das Gefühl der Präsenz erheblich beeinträchtigen. Moderne VR-Systeme streben Latenzzeiten unter 20 Millisekunden an, um eine flüssige und realistische Darstellung zu gewährleisten.

Stereoskopie: VR nutzt stereoskopische Techniken, um dem Nutzer ein dreidimensionales Seherlebnis zu bieten. Durch die Darstellung leicht unterschiedlicher Bilder für jedes Auge wird ein Tiefeneindruck erzeugt, der zur Immersion beiträgt: Die Stereoskopie verstärkt das Gefühl, tatsächlich in der virtuellen Umgebung anwesend zu sein, und verbessert die räumliche Wahrnehmung.

Maßstabstreue: Eine maßstabsgetreue Darstellung ist notwendig, um die Proportionen der virtuellen Umgebung und der darin enthaltenen Objekte korrekt wiederzugeben. Dies hilft dem Nutzenden, Entfernungen und Größenverhältnisse realistisch einzuschätzen. Maßstabstreue ist besonders wichtig für Anwendungen wie Architekturvisualisierungen und Simulationen, bei denen eine präzise Darstellung essenziell ist.

Augmented Reality (AR) ergänzt die reale Welt durch digitale Informationen und Objekte, ohne die Nutzenden von ihrer Umgebung abzuschotten.

Charakteristische Eigenschaften von AR sind:

Überlagerung der realen Welt: AR-Systeme überlagern die reale Welt mit digitalen Informationen. Diese Überlagerungen können sowohl visuelle als auch

auditive Informationen umfassen und dienen dazu, die Wahrnehmung der Realität zu erweitern. AR kann kontextabhängige Informationen liefern, die Nutzenden in Echtzeit zur Verfügung gestellt werden.

Interaktive Ergänzungen: AR ermöglicht interaktive Ergänzungen zur realen Umgebung. Nutzende können mit den eingeblendeten Informationen und Objekten interagieren, was neue Möglichkeiten für Anwendungen wie Navigation, Bildung und Unterhaltung eröffnet. Beispielsweise können Nutzende mit virtuellen Anleitungen interagieren, die über reale Objekte gelegt werden.

Mobilität: AR-Anwendungen sind in der Regel für mobile Geräte wie Smartphones und Tablets vorgesehen. Dies ermöglicht eine flexible und ortsunabhängige Nutzung. Tragbare AR-Brillen und Headsets bieten zudem eine freihändige und immersive Nutzungserfahrung.

Echtzeit-Rendering: Wie bei VR ist auch bei AR eine Echtzeit-Verarbeitung notwendig, um die digitalen Überlagerungen nahtlos in die reale Umgebung zu integrieren. Verzögerungen würden die Erfahrung der Nutzenden negativ beeinflussen. Echtzeit-Rendering ist entscheidend für Anwendungen, bei denen präzise und sofortige Reaktionen erforderlich sind, wie in der industriellen Wartung oder der Chirurgie.

AR ist besonders nützlich für Aufgaben, bei denen aus der realen Welt heraus der Zugriff auf zusätzliche Informationen und Anweisungen in Echtzeit erforderlich ist. AR ist ideal für Anwendungen im Feld, da es ihnen ermöglicht, relevante Daten und Anweisungen direkt in ihrem Sichtfeld zu erhalten, ohne ihre Arbeit unterbrechen zu müssen. Beispiele hierfür sind Wartungstechniker, die Reparaturanweisungen direkt auf Maschinen projiziert bekommen, oder medizinisches Personal, das während einer Operation zusätzliche Informationen erhält. Reize aus der Umgebung, welche beispielsweise zum Erkennen von Gefahren notwendig sind, werden nicht systembedingt ausgeblendet. Schlecht designte audiovisuelle Inhalte, welche über das AR-System präsentiert werden, können aber dennoch die Sicht blockieren, akustische Signale übertönen und generell die Aufmerksamkeit von Nutzenden von seiner Umgebung ablenken, was bei der Gestaltung der Inhalte beachtet werden muss.

Avatare in der Virtual Reality (VR) und Augmented Reality (AR) sind digitale Darstellung von Nutzenden, die als deren Stellvertretende in der virtuellen oder erweiterten Umgebung fungieren. Diese Darstellung kann in Form eines realistischen 3D-Modells, einer abstrahierten Darstellung oder, im einfachsten Fall, durch ein beliebiges grafisches Symbol erfolgen.

Im Rahmen des Konzepts der Mensch-Computer-Interaktion ist ein Avatar eine wahrnehmbare digitale Darstellung, deren Verhaltensweisen die Verhaltensweisen eines bestimmten Menschen widerspiegeln.[1]

Die Entwicklung von Avataren hat ihren Ursprung in den frühen Computerspielen und Online-Communities der 1980er Jahre. „Der Begriff und das Konzept tauchten tatsächlich bereits 1984 in Online-Multiuser-Dungeons oder MUDs (Rollenspielumgebungen) auf, und das Konzept, wenn auch nicht der Begriff, erscheint in fiktiven Werken, die bis Mitte der 1970er Jahre zurückreichen“ [1]. Die zunehmende Rechenleistung ermöglicht immer realistischere und immersivere Avatar-Darstellungen. Zur Vermittlung von spezifischen Inhalten und für eine hinreichende Präsenz-Erfahrung – gerade bei der Anwendung für Instruktionen und Schulungen – haben sich jedoch auch einfachere, nur auf den Oberkörper, Kopf und die Hände beschränkte Darstellungen als sinnvoll erweisen [2].

4 Datensynchronisierung

4.1 Identifizierung relevanter Daten

Ausgehend von dem oben beschriebenen Konzept der VR-AR-Kopplung im Projekt TeleInteraction-XR (vgl. Abbildung 1), wurden drei Gruppen von zu übertragenden Daten identifiziert:

Übertragung der Nutzerposition: Darunter fallen alle Daten, welche benötigt werden, um die Avatare des Gesprächspartners in ihrer Lage und Ausrichtung korrekt darzustellen.

Bidirektionale audiovisuelle Kommunikation: Neben dem gesprochenen Wort, welches essenziell zur Kommunikation zwischen Personen im Feld und der aus der Ferne unterstützenden Person ist, sind auch Videodaten zu übertragen. Diese können Detail-Informationen vermitteln, welche sich nur schwer verbal beschreiben oder nicht hinreichend gut als virtuelles Modell abbilden lassen.

Übertragung der Realumgebung: Zumindest die im Fokus der Fernunterstützung stehenden Objekte, wie Maschinen oder Maschinenteile, müssen in die VR-Umgebung des Experten übertragen werden.

Im letzten Punkt lag eines der Kernziele des Projekts TeleInteraction-XR. In drei aufeinander aufbauenden Umsetzungsvarianten wurden verschiedene Techniken zur Übertragung bzw. Bereitstellung von Daten zur Realumgebung erstellt und getestet.

In der *Basisvariante* kommen vormodellierte, dreidimensionale Modelle von Maschinen und Anlagen zum Einsatz. Dies ist die bisher weit verbreitete Vorgehensweise zur Erstellung von AR- und VR-Szenarien. Dabei werden alle

Objekte, die in der virtuellen Umgebung zu sehen sein sollen, im Vorfeld als dreidimensionale Computergrafik bereitgestellt. Dies kann anhand von Bildvorlagen, Zeichnungen, Beschreibungen oder auf Basis bestehender CAD-Modelle erfolgen. Auch letztere müssen in die für VR- und AR-Anwendungen eingesetzten, Vertex-basierten polygonalen Modelle transferiert werden.

Vorteil dieser Basisvariante ist, dass jedes benötigte Detail für eine erfolgreiche Fernunterstützung in entsprechender Qualität ausmodelliert und in der VR-Umgebung zur Verfügung gestellt werden kann. Nachteile sind die benötigte Vorlaufzeit und die durch Modellierung oder Konvertierung entstehenden Kosten. Auch kann diese Technik nicht bei hochdynamischen Objekten und Umgebungen angewandt werden.

In Fokus der *zweiten Umsetzungsvariante* stehen 2D- und stereoskopische bzw. 360°-Videostreams zur Übertragung der Umgebung des AR-Nutzenden in die virtuelle Welt des VR-Nutzenden. Übertragen werden dabei nicht allein die Objekte, die direkter Gegenstand der Fernunterstützung sind, sondern auch die erweiterte räumliche Umgebung. Dadurch sollen VR-Nutzenden den räumlichen Kontext – etwa bauliche Gegebenheiten sowie die Ver- und Entsorgungsinfrastruktur – erfassen können.

Durch das Streaming können auch dynamische oder bezüglich Größe und Detailvielfalt nur schwer als 3D-Modell umsetzbare Umgebungen schnell und kostengünstig in die VR übertragen werden. Limitierend wirkt sich die bei einfachen 360°-Aufnahmen fehlende oder nur aufwändig zu erfassende Tiefeninformation aus. Änderungen der Position von VR-Nutzenden erfordern eine identische Änderung der Position der aufzeichnenden Kameras, um die korrekte Perspektivendarstellung zu erreichen.

In der *dritten Variante* werden die Tiefensensoren der AR-Brille genutzt. Die so aus der Umgebung gewonnenen Daten werden dynamisch in die VR-Umgebung übertragen und dort in Quasi-Echtzeit abgebildet. Der Vorteil dieser Variante besteht in der schnellen, dreidimensionalen Erfassung und Übertragung der Realumgebung mit in der AR-Brille integrierter Technik. Limitierend sind die schnelle steigende Datenmenge und die nach aktuellem Stand der Technik eingeschränkte Detailgenauigkeit, da AR-Brillen mit den Tiefensensoren erfassten Daten ursprünglich nur zu Navigationszwecken erfassten.

4.2 Herausforderungen bei der Datensynchronisation

Hintergrund: Eine computergenerierte dreidimensionale Umgebung, wie sie in AR- und VR-Systemen zum Einsatz kommt, basiert auf polygonalen Modellen aus einer Anzahl Flächen, deren Ecken durch Scheitelpunkte (Vertices)

definiert sind. Diese Punkte sind die grundlegende visuelle Einheit im 3D-Raum. Sie besitzen keine Größe, sondern nur eine Position im Raum. [3]

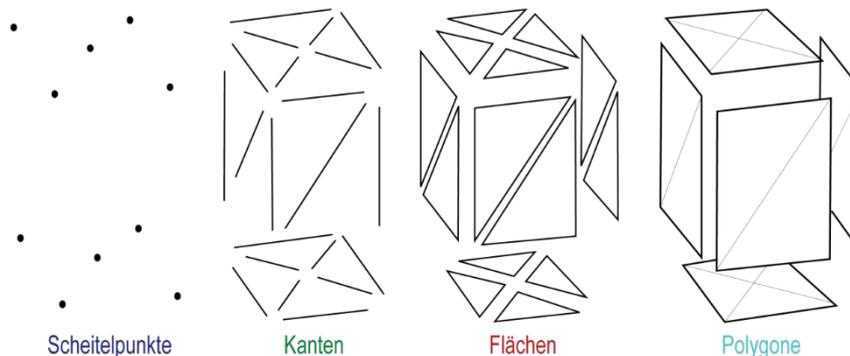


Abbildung 2: Zusammenhang zwischen Scheitelpunkten, Kanten, Flächen und Polygonen [3]

Für die Lagebestimmung der Scheitelpunkte kommen verschiedene Koordinatensysteme zum Einsatz. Die sogenannten Weltkoordinaten (World Space) ist das dreidimensionale Koordinatensystem für die gesamte Szene. Sein Ursprung (Origin) liegt in der Mitte der Szene. Die Achsen des World Space werden in der Regel mit x, y und z bezeichnet, wobei z meist die nach oben zeigender Achse darstellt. Der Objektraum (Object Space) ist das Koordinatensystem aus der Sicht eines Objekts. Die Position der Vertices eines Objektes werden im Koordinatensystem des Objektraum definiert. Der Ursprung des Objektraums besitzt eine Positionskomponente (x/y/z-Komponente) und Rotationskomponente (dargestellt als Eulerwinkel oder Quaternion) im World Space. Der Ursprung eines Objektraums ist somit auch Drehpunkt des Objekts im World Space. Der lokale Raum (Local Space) ist dem Objektraum ähnlich, verwendet jedoch den Ursprung und die Achsen des übergeordneten Knotens des Objekts in einer Gruppe von miteinander verknüpften Objekten. Dies ist nützlich, wenn ein Objekt nicht einzeln, sondern mit der übergeordneten Objektgruppe bewegt oder rotiert werden soll. [4]

Modellbereitstellung: Bevor die Lage von Objekten oder Avataren zwischen der AR- und der VR-Seite abgeglichen und synchronisiert werden kann, müssen die Daten zur Darstellung dieser Objekte zur Verfügung gestellt werden. Hierbei stehen grundsätzlich drei Ansätze zur Auswahl: die dynamische Übertragung, ein manueller Austausch oder die feste Hinterlegung von Daten im System selbst.

Eine *dynamische Übertragung zur Laufzeit*, also während der Nutzung des AR-/VR-Systems, ist sinnvoll, wenn bei jedem Einsatz unterschiedliche oder

stark dynamische Objekte dargestellt werden. Ein Beispiel für die dynamische Übertragung zur Laufzeit ist das oben genannte dynamisches Streaming von dreidimensionalen Modelldaten, die von den raumerfassenden Navigations-sensoren einer AR-Brille (HoloLens2) erzeugt werden.

Ein *manueller Abgleich* ist eine technisch weniger aufwändige Methode und eignet sich für Objekte, die nur selten verändert werden, wie beispielsweise Maschinen und Anlagen, die nur in längeren Abständen umgebaut oder modernisiert werden.

Die *feste Integration* in den Datenbestand einer Software ist für unveränderliche oder generische Daten möglich. Diese stehen dann in jeder Anwendungssession unverändert zur Verfügung. Für die Erprobungen im Projekt TeleInteraction-XR wurde dieser Ansatz für statischen Maschinen- und Anlagenbestandteile und die Avatar-Darstellung gewählt.

Abgleich der globalen Positionierung: AR- und VR-Systeme nutzen ihre jeweils eigenen Weltkoordinaten. Für die korrekte räumliche Positionierung von Objekten aus der AR-Umgebung im VR-System – und umgekehrt – ist daher eine initiale Kalibrierung der Lage von VR- und AR-Koordinaten notwendig. Hierzu haben sich verschiedene Methoden etabliert.

Für die prototypische Umsetzung im Projekt TeleInteraction-XR wurde auf eine manuelle Kalibrierung gesetzt. Das als Szenenursprung der VR-Seite dienende Maschinenmodell wurde manuell in der AR-Umgebung verschoben und manuell „über“ dem realen Objekt positioniert, so dass die Koordinaten von AR- und VR-Seite synchron liegen. Nach dieser initialen Kalibrierung war Positionsabgleich dann von dem Selbstverortungssystem der eingesetzten HoloLens2 abhängig. Bei längerer Nutzungsdauer zeigte sich in den Tests ein deutlicher Versatz zwischen realer und virtueller Welt. Daher war ein mehrfaches Nachkalibrieren erforderlich. Ein Versatz trat auch dann auf, wenn die internen Sensoren der HoloLens2 ihr räumliches Tracking verloren, etwa wenn der Raum keine deutlichen Kontraste aufwies.

Deshalb scheint für einen späteren kommerziellen Einsatz der AR-/VR-Kopplung der Einsatz von Markern sinnvoll. Die Marker werden auf vordefinierte Positionen in der realen Welt – beispielsweise an Außenflächen von Maschinen – angebracht, können vom AR-System erkannt und zur kontinuierlichen Verfolgung der relativen Lage der AR-Brille genutzt werden. Auf diese Weise können auch leichte Drifts der Koordinatensysteme kontinuierlich ausgeglichen werden.

Übertragung von Positionsdaten (Transforms): Neben den Weltkoordinaten sind auch die Positions- und Rotationsdaten aller darzustellenden Objekte zu synchronisieren. Dies betrifft die visuellen Modelle, aber auch die Position

und Lage der Avatare der AR- und VR-Nutzenden. Sollen Objekte in ihrer Form manipulier- oder änderbar sein, müssen auch die entsprechenden Koordinaten ihrer Vertices in den lokalen Koordinatensystemen der Objekte zwischen AR- und VR-System synchronisiert werden. Dabei sind unterschiedlich ausgerichtete Koordinatensysteme (linkshändig, rechtshändig, Lage der Z-Achse etc.) der eingesetzten Systeme zu beachten und bei Bedarf zu transformieren.

Avatar-Bewegung: Avatare unterscheiden sich von anderen Objekten nur dadurch, dass die Positionsdaten nicht nur durch eine Interaktion der Nutzenden verändert werden, sondern dass die AR- und VR-Systeme ihre jeweilige Sensorik nutzen, um Lage- und Positionsänderungen kontinuierlich zu erfassen. Je nach Ausprägung der Avatare müssen nur eigenständige 3D-Objekte für Rumpf, Hände und Kopf verändert werden, oder – bei realistischeren Darstellungen – auch die lokalen Koordinaten der Scheitelpunkte, die die dreidimensionale Form des Avatars definieren.

Interaktionsdaten: In Echtzeit zu synchronisieren sind auch Daten zu Interaktionen der Nutzenden mit der virtuellen und erweiterten Umgebung. Dazu zählen das Hinzufügen, Verändern und Löschen von Annotationen, das Einblenden von Verpixelungen oder Verwisch-Effekten – etwa um den Schutz von Persönlichkeitsrechten zu gewähren – sowie Material- und Beleuchtungsinformationen von Objekten, wenn diese optisch hervorgehoben oder ausgeblendet werden sollen.

Audio- und Kommunikationsdaten: Für ein starkes Telepräsenzempfinden müssen auch Audiodaten (Sprache und Umgebungsgeräusche) in Echtzeit übertragen und synchronisiert werden. Dies umfasst bestenfalls auch die räumliche Verortung der Audioquellen.

Umgebungsdaten: Um eine konsistente Darstellung der hybriden Umgebung zu gewährleisten, müssen Daten zur virtuellen (VR-System) und realen Umgebung (AR-System) synchronisiert werden. Dies umfasst auch Informationen zu Änderungen und Anpassungen in beiden Umgebungen.

Alle übertragenen Daten müssen konsistent und fehlerfrei sein, um eine korrekte Darstellung zu ermöglichen. Mechanismen zur Fehlererkennung und -korrektur sind hierbei entscheidend. Da personenbezogene und andere sensible Daten übertragen werden, müssen robuste Sicherheitsmechanismen implementiert werden. Eine der größten Herausforderungen besteht darin, die Latenzen bei der Datenübertragung so gering wie möglich zu halten, um ein flüssiges und immersives Erlebnis zu gewährleisten.

5 Fazit

Für die AR/VR-Kollaboration für eine telepräsenste Fernunterstützung ist es essenziell, die verschiedenen zu synchronisierenden Daten bereits im Vorfeld zu identifizieren und Herausforderungen bei der Übertragung zu erkennen.

Zentral ist der räumliche Abgleich der unterschiedlichen Koordinatensysteme, mit denen beide Seiten der Kommunikation (VR und AR) ihre 3D-Daten verwalten und darstellen. Für die Übertragung der Daten stehen unterschiedliche Methoden zur Verfügung, die sich nach Anwendungsspezifik (statische vs. dynamische Daten), eingesetzten Technologien (sensorische Objekterkennung vs. Vormodellierung) und Aufwand unterscheiden. Für das Empfinden von Telepräsenz und für die Kommunikation ist weiterhin die Übertragung von Audio- und Interaktionsdaten notwendig bzw. förderlich.

Auf Basis der identifizierten Bedarfe und Herausforderungen bei der Datensynchronisation wurden im Projekt TeleInteraction-XR eine Demonstrator-Lösungen entwickelt und in verschiedenen Anwendungsfällen erprobt. Die Details der Entwicklung werden in [5, 6, 7] und die Erprobung in [8] beschrieben.

Die Forschungs- und Entwicklungsarbeit wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ im Projekt TeleInteraction-XR (Förderkennzeichen 02L21B550-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt der Veröffentlichung liegt beim Autor.

6 Literaturverzeichnis

- [1] Bailenson, J.N. & Blascovich, Jim. (2014). Encyclopedia of Human-Computer Interaction. Avatars. 64. 64-68.
- [2] Brade, Jennifer & Kögel, Alexander & Fuchs, Christian & Klimant, Philipp. (2020). Impact of First Person Avatar Representation in Assembly Simulations on Perceived Presence and Acceptance. In Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2020) - Volume 1: GRAPP, pages 17-24 DOI: 10.5220/0008878700170024.
- [3] https://en.wikipedia.org/wiki/Polygon_mes
- [4] Nischwitz, A., Fischer, M., Haberäcker, P., & Socher, G. (2012). Computergrafik und Bildverarbeitung. Vieweg+Teubner Verlag. <https://doi.org/10.1007/978-3-8348-8323-0>

- [5] Kiesel, M., Kögel, A. (2025). Ich habe etwas vorbereitet – 3D-Modell-basierte AR-/VR-Kopplung. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenste Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [6] Wulf-Misaki, D. (2025). Kamera, Licht, Action! - Videostream-basierte AR/VR Kollaboration. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenste Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [7] Kiesel, M., Wulf-Misaki, D. (2025). Mehr als oberflächlich! – Live Volumetric Streamings für die AR/VR-Kollaboration. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenste Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [8] Berger, M., Holz, N., Ruffert, D., Kögel, A., (2025). TeleInterACTION! – VR/AR-Kopplung im Praxistest. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenste Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.

Autor



Kögel, Alexander

Alexander Kögel studierte Medientechnik (B.Eng.) und Medienmanagement (M.Eng) an der HTWK Leipzig. Seit 2017 ist er an der TU Chemnitz als wissenschaftlicher Mitarbeiter tätig. Schwerpunkte sind VR-Anwendungen, das Präsenz-Empfinden in virtuellen Umgebungen und der Erschließung der Potenziale der Digitalisierung, Vernetzung und Industrie 4.0.

Ich habe etwas vorbereitet – 3D-Modell-basierte AR/VR-Kopplung

Markus Kiesel¹, Alexander Kögel²

¹ CMC-Kiesel GmbH, ² TU Chemnitz – Professur Arbeitswissenschaft und Innovationsmanagement

Zusammenfassung

Eine Möglichkeit der Bereitstellung von Objektdaten für die Kollaboration zwischen Augmented-Reality (AR)-Anwendern im Feld und Experten, die in einer Virtuellen Umgebung (VR) agieren, sind vorab aus CAD-Daten erstellte 3D-Modelle. Der Beitrag beschreibt eine Vorgehensweise zur Modellerstellung sowie Anforderungen an den Backend-Server und skizziert einen Ansatz, um Abweichungen zwischen Modell und Realobjekt zu identifizieren.

1 Einleitung

Die naheliegende Grundlage für die Bereitstellung von 3D-Objektmodellen für AR/VR-Kollaborationen bilden vorhandene CAD-Daten, welche typischerweise in der Entwurfs- und Konstruktionsphase einer Maschine oder Anlage mit etablierten CAD-Programmen erstellt werden. CAD-Modelle dienen als Digitaler Zwilling bereits als zentrale Referenz für sämtliche nachfolgenden Produktionsschritte – angefangen bei der Fertigung einzelner Komponenten bis hin zur finalen Montage.

Vorteil der CAD-Daten bzw. Digitalen Zwillinge ist, dass sie in der Regel auf Basis vollständiger und präziser Stücklisten alle Komponenten einer Maschine, Anlage oder – allgemein gesprochen – eines Objekts umfassen. Auch die AR/VR-basierte Fernunterstützung von Fehleranalysen, Wartungen und Instandsetzungen profitiert von solchen detailgetreuen Abbildern.

Allerdings können CAD-Daten nicht direkt und im ursprünglichen Format für die Echtzeitvisualisierung in AR und VR genutzt werden. Im Folgenden werden die Gründe genauer erläutert und anschließend Lösungen aufgezeigt.

2 BREP und Mesh-Modelle. Was ist was?

CAD-Daten liegen grundsätzlich als mathematische Modelle vor, sogenannte Boundary Representation (BREP)-Daten. Hierbei wird die Geometrie eines Modells durch **exakte mathematische Beschreibungen seiner Oberflächen** dargestellt. Die BREP-Modelle entstehen mittels verschiedener Modellierungsalgorithmen, etwa durch Extrusions- und Subtraktionsverfahren (CSG-Algorithmen, Constructive Solid Geometry). Dabei werden eine oder mehrere zweidimensionale Kurven zu komplexeren 3D-Modellen verarbeitet.

Diese Vorgehensweise ist für Anwendungen geeignet, die eine hohe Genauigkeit, Maßhaltigkeit und numerische Stabilität erfordern – beispielsweise in der Konstruktion, Simulation oder Fertigung. Für die Darstellung in AR- und VR-Umgebungen ist dieses Verfahren jedoch ungeeignet, da die mathematische Beschreibung der Oberflächen in Echtzeit berechnet werden müsste, was zu erheblichen Performanceeinbußen führt.

Ein weiteres Hindernis für die direkte Nutzung von CAD-Daten in AR und VR ist die hohe **Komplexität typischer CAD-Gesamtmodelle**. Maschinen oder Anlagen bestehen häufig aus einer großen Anzahl einzelner Bauteile, die schnell 1.000 bis 10.000 erreicht, und die in unterschiedlicher Form miteinander verbunden sind. Auch diese Komplexität führt zu einem massiven Ressourcenbedarf bzw. Performanceproblemen bei Echtzeitvisualisierungen.

Um eine performante Echtzeitvisualisierung von CAD-Modellen zu ermöglichen, sind die BREP-Daten zwingend in effizient verarbeitbare, polygonale Repräsentationen – sogenannte Mesh-Daten – umzuwandeln. Dieser Prozess wird **Triangulierung** genannt. Dabei werden alle Oberflächen des ursprünglichen mathematisch definierten Modells durch Dreiecke angenähert. Dadurch werden beispielsweise Kreisflächen zu N-Ecken. Die Genauigkeit der Mesh-Daten lässt sich durch die Anzahl der verwendeten Dreiecke variieren: Eine hohe Anzahl an Dreiecken bietet eine bessere visuelle Annäherung an das ursprüngliche CAD-Modell.

Abbildung 1 veranschaulicht dies anhand eines Robotermodells. Beide Modelle basieren auf denselben CAD-Daten und wirken auf den ersten Blick sehr ähnlich. Das linke Modell besitzt jedoch fast doppelt so viele Dreiecke wie das rechte und erscheint bei näherer Betrachtung schärfer und realitätstreuer.

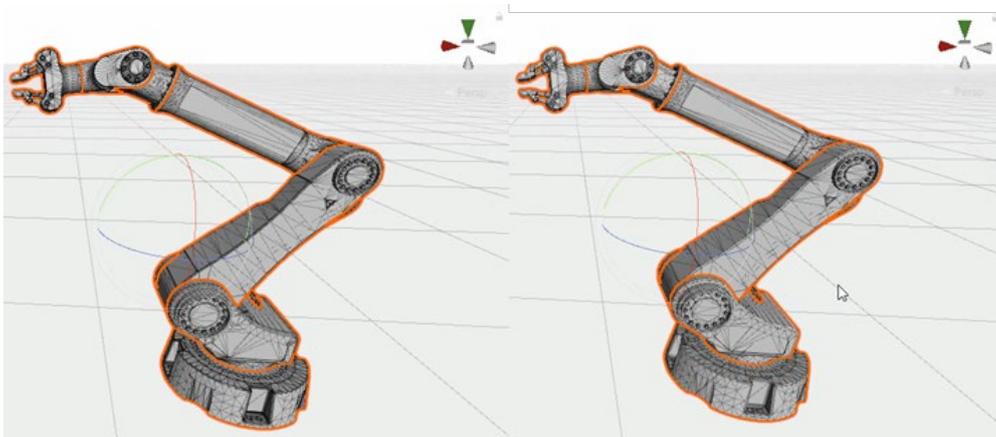


Abbildung 1: Beispielhafte Unterschiede bei der Triangulierung: Linkes Modell mit 50.000 und rechtes Modell mit 26.000 Dreiecken (eigene Darstellung)

Eine zu hohe **Anzahl von Dreiecken pro Szene** führt wiederum zu Performanceeinbußen. Tabelle 1 zeigt empfohlene Grenzwerte für deren maximale Anzahl bei ausgewählten Augmented- und Mixed-Reality-Geräten. Dabei wird für Praktiker schnell deutlich, dass diese Grenzwerte bei komplexen Anlagenmodellen schnell erreicht oder überschritten werden. Zudem sind Reservekapazitäten für weitere Berechnungen wie komplexe Shader, Beleuchtung oder Materialeigenschaften einzukalkulieren.

Tabelle 1: Übersicht maximale Anzahl an Dreiecken

Gerät	Maximale Anzahl an Dreiecken	Geräte-Art
Microsoft HoloLens 1	50.000 – 100.000	AR
Microsoft HoloLens 2	100.000 – 300.000	AR
Meta Quest 2	750.000 – 1.000.000	XR
Meta Quest 3	1.500.000 – 2.500.000	XR
Magic Leap 1	300.000 – 500.000	AR
Magic Leap 2	500.000 – 1.000.000	AR
HTC Vive XR Elite	1.500.000 – 3.000.000	XR
Apple Vision Pro	3.000.000 – 5.000.000	XR

Im Kontext von VR/AR-Anwendungen für die Fernunterstützung in der Industrie reduziert sich die effektiv für Maschinenmodelle nutzbare Anzahl von Dreiecken weiter, da meist noch zusätzliche Objekte wie Avatare, Hilfsgeometrie zur Bedienung sowie grafische Overlays dargestellt werden müssen.

Daher ist zur Gewährleistung ausreichender Performance in der Regel eine umfangreiche, meist manuelle Optimierung von Mesh-Daten für Maschinen und Anlagen notwendig. Dies erfolgt typischerweise in folgenden Schritten:

- Initiale Konvertierung des gesamten Maschinenmodells in Mesh-Daten unter Berücksichtigung einer definierten Genauigkeit. Diese hängt von der Größe und Komplexität der Anlage ab.
- Reduzierung der Dreiecksanzahl einzelner Elemente durch vereinfachende Verfahren wie Decimation und Mesh-Optimierung.
- Optionales Entfernen sehr kleiner, nicht relevanter Bauteile, beispielsweise Schrauben, Muttern oder Unterlegscheiben, die visuell kaum wahrnehmbar sind.
- Optionales Entfernen irrelevanter Geometrie-Features, etwa innerer Strukturen, die für die AR/VR-Darstellung irrelevant sind.

Auch die Anzahl der einzelnen Bauteile beeinflusst die Darstellungsleistung einer Echtzeitvisualisierung erheblich. Es böte sich daher an, Bauteile nach gemeinsamen Eigenschaften – beispielsweise gleicher Materialfarbe oder Zugehörigkeit zu einer Baugruppe – zusammenzufassen. Das schränkt allerdings die Flexibilität ein, da so einzelne Komponenten nicht mehr separat bewegt oder selektiv bearbeitet werden können. Für die Fernunterstützung via AR/VR-Kopplung sollte daher auf diese Form der Optimierung verzichtet werden, um eine hohe Interaktivität auf Komponentenebene zu gewährleisten.

In den Anwendungsfällen [1] des Projekts TeleInteraction-XR wurden Geräte, Maschinen und Anlagen mit einer moderaten Komplexität betrachtet. Deshalb konnte dort auf die oben genannten optionalen Optimierungsschritte verzichtet werden.

3 Datenkommunikation – Backend-Server als Drehscheibe

Um eine nahtlose Echtzeit-Kollaboration zwischen verschiedenen örtlich verteilten Systemen zu ermöglichen, ist ein möglichst flexibles und skalierbares Backend-System nötig. Das Backend soll heterogene Lösungen unter-

schiedlichster Art anbinden und mit wechselnden Zahlen von Nutzenden effizient umgehen können.

Im Zentrum des im Projekt TeleInteraction-XR entwickelten Backend-Systems steht ein leistungsfähiger, event-basierter NodeJS-Server, der die Kommunikation zwischen sämtlichen Teilnehmenden (Clients) koordiniert. Dessen primäre Aufgabe besteht darin, von den einzelnen Clients empfangene Datenpakete entgegenzunehmen und umgehend an alle anderen verbundenen Systeme weiterzuleiten. Darüber hinaus speichert der Server die Historie sämtlicher Nachrichten, um nachträglich hinzukommenden Clients ein vollständiges Abbild der Session bereitstellen zu können.

Die Kommunikation zwischen Backend und Clients erfolgt über zwei Hauptkanäle, die sich aufgrund ihrer unterschiedlichen technologischen Stärken gegenseitigen ergänzen:

gRPC (Google Remote Procedure Call) ist ein leistungsstarkes, plattformübergreifendes Open-Source-Framework zur Kommunikation zwischen verteilten Systemen. Es basiert auf Protocol Buffers (protobuf), einer standardisierten Methode zur kompakten und maschinenlesbaren Definition von Schnittstellen. Durch den Einsatz von HTTP/2 bietet gRPC eine besonders schnelle, zuverlässige und ressourcenschonende Datenübertragung. Ein weiterer Vorteil liegt darin, dass definierte protobuf-Schnittstellen automatisch und mit minimalem Aufwand in zahlreiche Programmiersprachen kompiliert werden können, was die Integration in diverse Applikationen erleichtert.

WebSockets wiederum stellen ein bidirektionales, vollduplexfähiges Kommunikationsprotokoll dar, welches über eine persistente TCP-Verbindung arbeitet. Im Gegensatz zu klassischen HTTP-Verbindungen müssen Clients dabei nicht kontinuierlich neue Anfragen senden, um aktuelle Informationen zu erhalten. Vielmehr ermöglicht eine einmal etablierte Verbindung einen kontinuierlichen, nahezu verzögerungsfreien Datenaustausch zwischen Server und Client in Echtzeit. Aufgrund der breiten Unterstützung in modernen Browsern und vielfältigen Anwendungen stellen WebSockets eine äußerst universelle Lösung für die Echtzeitkommunikation dar.

Das entwickelte Backend zeichnet sich insbesondere durch seine Fähigkeit aus, Informationen dynamisch zwischen diesen beiden Kommunikationskanälen zu übersetzen: Es ist möglich, eine Seite einer Kommunikationsverbindung via WebSockets anzubinden, während auf der Gegenseite die Kommunikation über gRPC erfolgt. Damit kann nahezu jedes System einfach und effizient an das Kollaborations-Backend angekoppelt werden.

4 Besonderheiten bei 3D-basierte AR/VR-Kollaboration

Um spezielle Anforderungen der AR/VR-Kollaboration mittels vorbereiteter 3D-Modelle zu genügen, wurden dem Backend-Server folgende Funktionalitäten hinzugefügt:

Session-Management: Um mehrere Kollaborationssitzungen parallel durchführen zu können, wurde ein Mechanismus implementiert, der Clients und deren übertragene Daten eindeutig einzelnen Sessions zuordnen und Datenpakete adäquat verteilen kann. Eine Session wird dazu durch eine eindeutige alphanumerische Kennung identifiziert. So lässt sich mit nur einem Server eine Skalierung auf zahlreiche parallele Sessions erreichen.

Transformationsübertragung: Um sicherzustellen, dass sämtliche 3D-Modelle auf allen Plattformen konsistent und positionsgenau dargestellt werden, werden Positionsänderungen der einzelnen Komponenten kontinuierlich an alle Teilnehmer übertragen. Das entsprechende Datenpaket umfasst jeweils Informationen zur Position, Rotation sowie zur Skalierung eines spezifischen Bauteils im dreidimensionalen Raum.

Avatar-Übertragung: Für Avatare müssen neben Transformationsdaten (Position, Orientierung) zusätzliche Zustandsinformationen wie Benutzernamen, Statusinformationen (z. B. offene vs. geschlossene Hand) sowie weitere dynamische Eigenschaften übertragen werden.

Highlight-Funktionalität: Um die Aufmerksamkeit von Teilnehmenden während einer Kollaboration zu lenken, wurde eine zusätzliche Funktion implementiert, mit der einzelne oder mehrere Komponenten innerhalb eines komplexen Modells visuell hervorgehoben werden können.

Identifikation: Die eindeutige Identifikation sämtlicher Komponenten innerhalb des virtuellen Modells erfolgt über Global Unique Identifiers (GUIDs), die während der Datenaufbereitung und Optimierung automatisch für jedes einzelne Bauteil vergeben werden. Nach Abschluss dieses Prozesses wird das gesamte optimierte Modell als GLB-Datei exportiert.

GLB (Graphics Language Binary) ist ein von der Khronos Group entwickeltes offenes binäres Dateiformat für das standardisierte Grafikformat glTF (Graphics Language Transmission Format). Es wurde für die Speicherung und Übertragung komplexer 3D-Modelle optimiert und eignet sich hervorragend für Web-, AR- und VR-Anwendungen, da es kompakt und universell kompatibel ist. Außerdem können neben Geometrie- und Texturdaten auch beliebige nutzerdefinierte Metadaten integriert werden.

Die Verteilung der GLB-Dateien an die Teilnehmenden einer AR/VR-Kollaboration kann entweder direkt über das Backend oder über einen manuellen Datenaustausch erfolgen.

Da die Anwendungsfälle im Projekt TeleInteraction-XR auf spezifische, vorab bekannte Geräte, Maschinen bzw. Anlagen der beteiligten Industriepartner fokussiert waren, wurden die vorbereiteten 3D-Objektmodelle insbesondere auf der AR-Seite bereits im Vorfeld der Kollaboration auf der jeweiligen lokalen Applikation integriert. Dies gewährleistet eine hohe Performance und minimierte die initialen Ladezeiten beim Start der Kollaborationssitzung.

5 Verortung in Koordinatensystemen

Eine exakte und zuverlässige Ausrichtung zwischen realem Objekt und seinem virtuellen Abbild in der AR-Darstellung ist schon für allein betriebene AR-Anwendungen essenziell. Für die AR/VR-Kopplung sind zusätzlich noch die unterschiedlichen Weltkoordinaten der AR- und der VR-Seite zu synchronisieren (Grundlagen und Begriffe siehe [2]). Dafür stehen mehrere Methoden zur Verfügung, von denen zwei näher betrachtet werden:

Beim **image-basierten Verfahren** erfolgt die Verortung des virtuellen Modells durch Erkennung und räumliche Zuordnung eines zuvor definierten Referenzbildes oder Markers. Dieser Marker muss speziell für den Einsatzzweck konzipiert sein und über eindeutig erkennbare Muster und Kontraste verfügen. Nach erfolgreicher Bilderkennung durch die AR-Anwendung wird automatisch die räumliche Position und Orientierung des virtuellen 3D-Modells ermittelt und entsprechend an die reale Anlage angelegt.

Damit dieses Verfahren korrekt funktioniert, müssen sämtliche 3D-Daten innerhalb eines verschiebbaren, lokalen Koordinatensystems (Local Space) eines Hilfsobjekts positioniert sein, dessen Ursprung (Origin) exakt auf dem Referenzmarker liegt. Nach der initialen Ausrichtung auf Basis des Bildes erfolgt häufig ein Übergang zum sogenannten Sensorfusion-basierten Tracking. Hierbei werden zusätzliche Sensoren wie Kameras, Gyroskope oder Beschleunigungsmesser eingesetzt, um auch bei Bewegung eine kontinuierliche Positionsverfolgung zu ermöglichen. Allerdings besteht bei zunehmender Entfernung vom ursprünglichen Marker die Gefahr, dass Abweichungen auftreten und somit Ungenauigkeiten bei der Zuordnung zwischen virtueller und realer Welt entstehen.

Eine Alternative ist die **semi-manuelle Verortung** (Anchoring). Hierbei bestimmt der Nutzer einen Punkt, der sowohl im digitalen Modell als auch in der realen Anlage leicht identifizier- und erreichbar ist. Besonders geeignet

sind dafür definierte Eckpunkte von Geometrien. Abbildung 2 zeigt dafür ein Beispiel. Links ist das virtuelle 3D-Modell der Maschine zu sehen, während rechts die Platzierung des Ankerpunktes an der realen Anlage gezeigt wird.

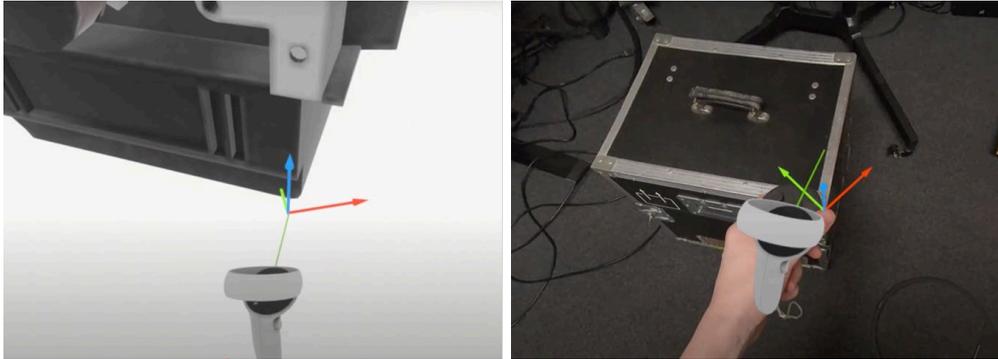


Abbildung 2: Ankerung, links am digitalen Modell im VR Modus, rechts an einem realen Modell in XR Modus

6 Digitaler Schatten – über den Produktlebensweg veränderliche 3D-Modelle

Nach erfolgreicher Inbetriebnahme einer Maschine oder Anlage folgt typischerweise eine lang andauernde Abfolge von Betriebs-, Wartungs- und ggf. Reparatur-, Umbau- bzw. Modernisierungsphasen. Dabei erfolgte Modifikationen werden nicht automatisch im ursprünglichen CAD-Modell berücksichtigt, wodurch es über die Zeit zu Abweichungen zwischen der real existierenden Anlage und ihrem virtuellen, digitalen Zwilling kommt. Wird das digitale Modell nicht regelmäßig aktualisiert, wachsen diese Abweichungen mit der Zeit immer weiter an, was die Präzision und Aussagekraft der AR/VR-basierten Kollaboration zunehmend beeinträchtigt.

Um trotz dieser – nicht immer vermeidbaren – Diskrepanzen zwischen Realität und virtuellem Modell eine akzeptable AR/VR-Kollaboration zu gewährleisten, kann es helfen, diese Abweichungen sicht- und nachvollziehbar zu machen.

Im Projekt TeleInteraction-XR wurde mit der **Integration von Punktwolken zur Visualisierung von Abweichungen** eine Methode zur Darstellung dieser Unterschiede beispielhaft erprobt. Die Punktwolke besteht dabei aus einer Vielzahl von räumlich erfassten Messpunkten, die mittels 3D-Scanning-Technologien (z. B. LiDAR-Scanner oder photogrammetrische Verfahren)

aus der realen Umgebung aufgenommen wurden. Diese Punktwolken erfüllen mehrere Aufgaben:

- **Visualisierung der tatsächlichen Umgebung:** Die aufgenommenen Punktwolken ermöglichen es, eine detaillierte und realitätsgetreue Abbildung der tatsächlichen Umgebung zu erzeugen. Dieser Kontext verbessert gerade auf der AR-Seite die Orientierung erheblich.
- **Identifikation von Abweichungen:** Durch den Vergleich der Punktwolke mit dem 3D-Modell können Abweichungen erkannt und analysiert werden.
- **Erleichterte Entscheidungsfindung für Experten:** Auf Basis der identifizierten Anforderungen kann entschieden werden, ob die Modelle und die zugrundeliegenden CAD-Daten nachzupflegen oder unzulässige Abweichungen in der Realität zu korrigieren sind.

Als beispielhafte Erprobung dieses Konzepts wurde durch einen 3D-Scan mit einem iPad eine Punktwolke eines realen Objekts aufgenommen und anschließend in das bestehende digitale Modell integriert. Aus der gemeinsamen Darstellung von digitalem Modell und real erfasster Punktwolke ließen sich erwartungsgemäß die Diskrepanzen erkennen.

7 Fazit

Die AR/VR-Kopplung auf Basis vorbereiteter 3D-Modelle, die aus CAD-Daten aufbereitet wurden, konnte in Verbindung mit einem spezifisch konfigurierten Backend-Server erfolgreich implementiert, getestet und praktisch validiert werden.

Die Vorgehensweise, Objektdaten für die AR/VR-Kollaboration mit vorbereiteten 3D-Modellen bereit zu stellen, verspricht insbesondere für neue Maschinen und Anlagen einen hohen Nutzwert. Für diese sind in der Regel vollständige und präzise CAD-Daten verfügbar, die zuverlässig und relativ aufwandsarm in hinreichend detailreiche 3D-Modelle gewandelt werden können.

Um Änderungen an Maschinen und Anlagen, die über den Lebenszyklus entstehen und die nicht ohne weiteres in den CAD-Modellen nachgepflegt werden, in der AR/VR-Kollaboration zu berücksichtigen, wurde beispielhaft das Erfassen einer Punktwolke vom realen Objekt und dessen Integration in die AR/VR-Umgebungen mit Erfolg erprobt. Es sind jedoch weitere Entwicklungsschritte nötig, um eine industrietaugliche Anwendung zu ermöglichen.

Die Forschungs- und Entwicklungsarbeit wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ im Projekt TeleInteraction-XR (Förderkennzeichen 02L21B550-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt der Veröffentlichung liegt beim Autor.

8 Literaturverzeichnis

- [1] Ruffert, D., Kögel, A., Löffler, T. (2025). So fern und doch so nah – Telepräsenes Arbeiten in der Industrie. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenes Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.
- [2] Kögel, A. (2025). Wer und was ist wo? - Synchronisation von Virtueller und Erweiterter Realität. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenes Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis.

Autoren



Markus Kiesel

Markus Kiesel ist studierter Maschinenbauingenieur (M.Eng.) Seit 2016 ist er als Unternehmer im Bereich der industriellen grafischen Simulation tätig. Mit seinem Team entwickelt er Anwendungen welche weltweite Verbreitung finden.



Kögel, Alexander

Alexander Kögel studierte Medientechnik (B.Eng.) und Medienmanagement (M.Eng) an der HTWK Leipzig. Seit 2017 ist er an der TU Chemnitz als wissenschaftlicher Mitarbeiter beschäftigt, mit dem Fokus auf VR-Anwendungen, Einflussgrößen auf das Presence-Empfinden in virtuellen Umgebungen und der Erschließung der technologischen und wirtschaftlichen Potenziale der Digitalisierung, Vernetzung und Anwendung von Industrie 4.0.

Kamera, Licht, Action! - Videostream-basierte AR/VR-Kollaboration

Daniel Wulf-Misaki¹

¹oculavis GmbH

Zusammenfassung

Um Augmented Reality (AR) und Virtual Reality (VR) zu koppeln, ist Videostreaming eine interessante Alternative zum Datenaustausch auf Basis vorgefertigter 3D-Modelle. Der Artikel beschreibt, wie der Einsatz moderner Streaming-Technologien, kombiniert mit spezialisierter Hardware und einer eigens weiterentwickelten Softwareplattform, dazu beiträgt, die Fernunterstützung bei Wartungsarbeiten effizienter, flexibler und präziser zu gestalten.

1 Einleitung

Die Kopplung von VR- und AR-Technologien eröffnet neue Möglichkeiten zur Fernunterstützung insbesondere in Bereichen wie Inbetriebnahme, Wartung und Diagnose technischer Anlagen. Nicht immer stehen aktuelle CAD-Modelle der betreffenden Maschinen, Anlagen und ihrer Peripherie zur Verfügung, um auf deren Basis entsprechende 3D-Modelle für die VR/AR-Kollaboration zu generieren. Zudem erfordert die Aufbereitung der 3D-Modelle für die VR/AR-Kopplung spezifische Expertise [1], die nicht zwingend zu den Kernkompetenzen von Industrieunternehmen zählt.

Im Forschungsprojekt TeleInteraktion-XR wurde diese Herausforderung aufgegriffen und eine Videostream-basierte VR/AR-Kollaboration entwickelt. Dabei wird mithilfe moderner Technologien wie der Microsoft HoloLens2, hochauflösender 360-Grad-Kameras sowie des Echtzeit-Kommunikationsstandards WebRTC (Web Real-Time Communication) eine immersive, interaktive und standortunabhängige Zusammenarbeit ermöglicht. Diese technische Infrastruktur versetzt aus der Ferne aus einer VR-Umgebung unterstützende Expertinnen und Experten in die Lage, unmittelbar, intuitiv und in Echtzeit an Wartungs- und Reparaturprozessen mitzuwirken.

2 VR/AR-Kollaboration mittels Videostreaming

Der entscheidende Mehrwert der Videostream-basierten Kollaboration besteht in der gegenseitigen Echtzeit-Visualisierung zwischen AR- und VR-Nutzenden. VR-Nutzende erhalten kontinuierlich detailgetreue visuelle Informationen aus der realen Umgebung, was ihnen erlaubt, präzise und effektiv technische Anweisungen zu geben [2]. Gleichzeitig haben AR-Nutzende, ausgestattet mit der Microsoft HoloLens2, die Möglichkeit, die virtuellen Interaktionen und Anweisungen der VR-Experten unmittelbar in ihrer realen Umgebung wahrzunehmen. Diese wechselseitige Integration der Realität und virtueller Informationen fördert eine präzise, intuitive und kollaborative Arbeitsweise [3].

Im Vergleich zur 3D-Modell-basierten Kollaboration, die auf vorab erstellten digitalen Modellen basiert, zeichnet sich die Videostream-basierte Kollaboration durch ihre hohe Flexibilität und direkte Einsetzbarkeit aus. Sie erfordert keine zeitaufwendige und kostenintensive Erstellung und Pflege digitaler Zwillinge, wodurch sie besonders für heterogene oder ältere Maschinenparks vorteilhaft ist. Zusätzlich werden reale Zustände, wie Verschmutzungen oder Beschädigungen, ohne Abweichungen exakt wiedergegeben [2].

Auch gegenüber dem Live Volumetric Streaming, welches räumliche Daten in Form von Punktwolken oder Meshs in Echtzeit überträgt, weist die Videostream-basierte Methode Vorteile auf. Diese umfassen eine geringere technische Komplexität, niedrigeren Bandbreitenbedarf sowie eine höhere Systemstabilität. Volumetrische Systeme erfordern hingegen spezialisierte und kostspielige Hardware sowie erheblichen Rechenaufwand, der in vielen industriellen Umgebungen nicht praktikabel ist [3].

Dennoch existieren auch Herausforderungen und limitierende Faktoren beim Videostreaming. Geringe Bandbreite, Netzwerkausfälle oder hohe Latenzzeiten können die Effizienz und Qualität des Videostreams und damit der Kollaboration mindern. Technische Limitierungen der verwendeten Hardware, wie eingeschränkte Akkulaufzeiten und begrenzte Sichtfelder, sowie schwierige Umweltbedingungen wie komplexe räumliche Gegebenheiten und schlechte Lichtverhältnisse müssen berücksichtigt und adressiert werden, um einen optimalen Einsatz sicherzustellen [4].

3 Technische Realisierung mittels AR und 360-Grad-Technologien

Die technische Realisierung der VR/AR-Kollaboration erfolgte mithilfe der Microsoft HoloLens2. Diese AR-Brille bietet eine hochauflösende Mixed Reality-Visualisierung mit einer Bildauflösung von 2048 x 1080 Pixeln pro Auge und einem diagonalen Sichtfeld von 52 Grad [4]. Fortschrittliches Inside-Out-Tracking mittels integrierter Tiefensensoren ermöglicht eine präzise räumliche Orientierung. Die intuitive Handgesten- und Blicksteuerung erlaubt eine natürliche Integration in die bestehenden Arbeitsprozesse [2].

Ergänzend kommen fortschrittliche 360-Grad-Kameras zum Einsatz, die typischerweise mit mehreren Objektiven ausgestattet sind, um sphärische Videoaufnahmen in Auflösungen von bis zu 8K zu erzeugen. Die Auswahl der spezifischen Hardware erfolgt gezielt, um unterschiedliche Videostream-Typen zu ermöglichen, die dann verschiedene Formen der Abbildung in der VR-Umgebung erlauben. Die Kamera Sensoren der HL wurden verwendet, um die Perspektive der AR-Nutzenden in der VR zu visualisieren. Hochauflösende 360-Grad-Kameras bieten zusätzlich eine immersive Rundumsicht und eignen sich hervorragend zur umfassenden räumlichen Darstellung von Umgebungen und Szenarien in der VR-Anwendung. Die VR-Nutzenden erhalten so einen realistischen Eindruck der realen Situation vor Ort, können sich schnell und intuitiv in die Arbeitsumgebung hineinzusetzen, was die Qualität und Geschwindigkeit der Fehlerdiagnose und Problemlösung deutlich verbessert [3].

Die HoloLens2 hingegen ermöglicht eine Überlagerung der realen Umgebung mit ergänzenden virtuellen Informationen. Diese Kombination unterschiedlicher Videostream-Typen gewährleistet, dass für jedes spezifische Anwendungsszenario die optimale visuelle Unterstützung und realitätsnahe Darstellung gewährleistet ist.

4 MR SHARE Applikation: Videokommunikation aus AR-Nutzerperspektive

Die speziell für die Microsoft HoloLens2 entwickelte MR SHARE Applikation ermöglicht eine immersive, Hands-Free Videokommunikation und Kollaboration zwischen AR- und VR-Nutzenden. Im Zentrum der Anwendung steht die Echtzeitübertragung von Video- und Audiodaten mittels WebRTC, wobei der hocheffiziente Videocodec H.264 genutzt wird. Aus Sicht der AR-

Nutzenden eröffnet sich hierdurch die Möglichkeit, Kommunikation visuell und akustisch unmittelbar in seiner realen Arbeitsumgebung zu integrieren, ohne dafür weitere Geräte bedienen zu müssen [4].

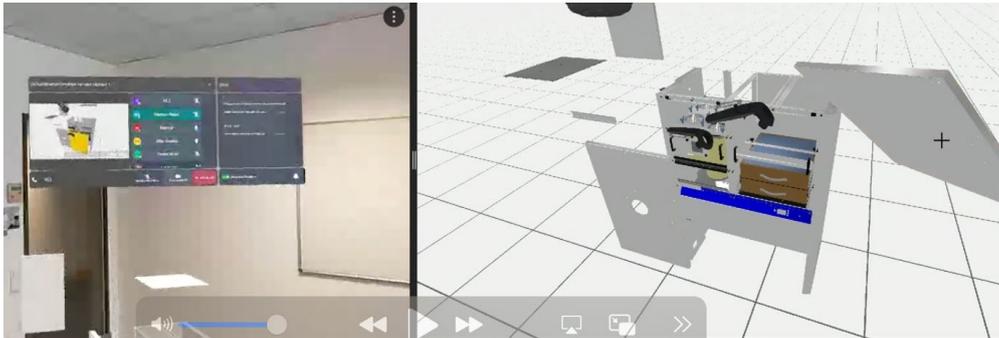


Abbildung 1: Demonstrator-Testsession mit der MR SHARE Applikation – AR-Nutzerperspektive (links) und korrespondierende VR-Ansicht (rechts).

Ein entscheidender Mehrwert der MR SHARE Applikation liegt darin, dass virtuelle Benutzeroberflächen und Videostreams stabil in der realen Umgebung platziert werden können. Die HoloLens2 nutzt dabei ihre integrierte Tiefensensorik, um Oberflächen und räumliche Strukturen zuverlässig zu erkennen [4]. Nutzende können somit virtuelle Bildschirme an geeigneten Positionen innerhalb ihres Arbeitsbereichs verankern, beispielsweise direkt neben einer Maschine oder an ergonomisch sinnvollen Stellen, wodurch Hände und Sichtfeld frei bleiben [3].

Ein beispielhaftes Szenario zeigt Abbildung 1: Links ist die Sicht des AR-Nutzenden durch die HoloLens2 dargestellt. Zu sehen ist ein virtueller Bildschirm, auf dem sowohl die Videoübertragung als auch die Benutzeroberfläche der SHARE-Plattform angezeigt werden. Die dargestellten Informationen beinhalten die Live-Ansicht aus der VR-Umgebung, die zusätzlich im rechten Bildteil zu sehen ist. Dort führt eine Expertin oder ein Experte virtuell an einem 3D-Modell der betreffenden Maschine bestimmte Maßnahmen vor. Die AR-Nutzenden können diese dann unmittelbar in der realen Umgebung nachvollziehen.

Die Interaktion innerhalb der MR SHARE Applikation erfolgt primär über intuitive Handgesten wie den „Air Tap“ oder Greifgesten. Nutzende können

damit virtuelle Bildschirme und Videofenster flexibel positionieren, skalieren oder anderweitig interaktiv bedienen.

Eine weitere Funktion der MR SHARE Applikation liegt in der Visualisierung von Mixed-Reality-Annotationen, die vom VR-Nutzenden erzeugt werden, direkt im realen Arbeitsumfeld des AR-Nutzenden. Technisch umgesetzt wird dies mithilfe des Umgebungs-Trackings der HoloLens 2 [4].

5 Technische Infrastruktur der Videostreaming Plattform und WebRTC

Die Abbildung 2 veranschaulicht den Aufbau des Videostreaming-Systems: Alle Clients – ob AR-Brille, VR-System oder Webbrowser – sind mit dem SHARE-Server und dem WebRTC-Mediaserver verbunden. Im Folgenden werden die einzelnen Komponenten und ihre Funktionen erläutert.

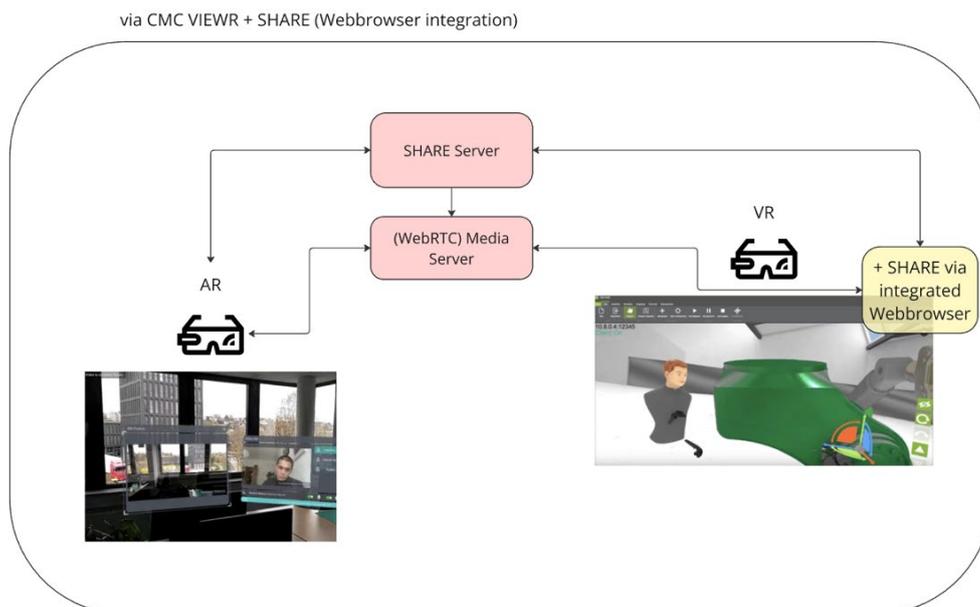


Abbildung 2: Technische Infrastruktur für die Videostream-basierte VR/AR-Kollaboration

Der SHARE Server stellt hierbei die zentrale Steuereinheit dar und übernimmt die Verwaltung der Nutzer, Sitzungen sowie Schnittstellen zu den beteiligten Komponenten. Konkret steuert er die Authentifizierung von Nutzern, verwaltet deren Zugriffsrechte und koordiniert virtuelle Meeting-Räume. Beim Beitritt zu einer Sitzung generiert und verwaltet der SHARE Server eine Session-ID, die alle Teilnehmer eindeutig identifiziert und miteinander verbindet. Der SHARE Server verarbeitet dabei keine Medieninhalte und leitet sie auch nicht selbst weiter; vielmehr koordiniert er den Verbindungsaufbau zwischen Teilnehmern und Media Server. Zudem stellt er Schnittstellen in Form von REST-APIs oder WebSocket-Endpunkten bereit, über die AR-, VR- und Browser-Clients Sitzungen beitreten und Informationen abrufen können.

Der angebundene WebRTC Media Server übernimmt Verarbeitung und Weiterleitung der Videostreams zwischen AR- und VR-Teilnehmern. Insbesondere in Multi-Teilnehmer-Konferenzen ist die Nutzung eines Media Servers essenziell, um die Skalierbarkeit zu gewährleisten. In reinen Peer-to-Peer-Systemen würde die Anzahl direkter Verbindungen bei mehreren Teilnehmern exponentiell wachsen und eine übermäßige Belastung der Endgeräte verursachen. Zudem ermöglicht der WebRTC Media Server eine dynamische Anpassung der Stream-Qualität entsprechend den aktuellen Netzwerkbedingungen, indem er etwa Teilnehmern mit schwächeren Verbindungen niedriger aufgelöste Videostreams bereitstellt. Technologisch nutzt WebRTC dazu UDP-basierte Protokolle für einen schnellen, latenzarmen Datenaustausch sowie NAT-Traversierungsmechanismen (ICE, STUN und TURN), um auch Teilnehmer hinter Firewalls oder NATs zuverlässig miteinander zu verbinden. Weiterhin unterstützt der Media Server die gängigen Videocodecs wie H.264 (der u.a. von der HoloLens2 verwendet wird) sowie Audiocodecs wie Opus oder AAC.

Die AR-Komponente, hier die Microsoft HoloLens2, integriert sich über die oben beschriebene MR SHARE Applikation, die mithilfe der MixedReality-WebRTC-Bibliothek und Unity3D entwickelt wurde. Beim Verbindungsaufbau authentifiziert sich die HoloLens zunächst beim SHARE Server, erhält Session-Daten wie Token und Media Server-Adressen, und baut anschließend eine WebRTC-Verbindung auf. Über diese Verbindung sendet sie den eigenen Kamera-Stream (Blickfeld des AR-Nutzers) und empfängt zugleich Videostreams anderer Sitzungsteilnehmer. Die empfangenen Videoinhalte werden räumlich präzise als virtuelle Bildschirme innerhalb der realen Umgebung des AR-Nutzers visualisiert, wobei die Positionierung stabil mittels der Sensorik der HoloLens erfolgt. Interaktionen der Nutzenden mit den eingeblendeten Inhalten, wie etwa das Platzieren oder Verschieben von Annotationen, erfolgen intuitiv über Handgestensteuerung.

Die VR-Komponente basiert auf einer integrierten Webbrowser-Implementierung innerhalb der VR-Anwendung, beispielsweise der CMC VIEWR. Über diesen integrierten Browser greifen VR-Nutzende direkt auf die SHARE-Plattform zu, empfangen Videostreams und nehmen an der Sitzung interaktiv teil, ohne die VR-Umgebung verlassen zu müssen. Der Webbrowser agiert dabei wie ein gewöhnlicher Client der SHARE-Plattform, was eine unkomplizierte und flexible Integration ermöglicht. Teilnehmende können die eingehenden Videostreams auf frei platzierbaren virtuellen Bildschirmen innerhalb ihrer VR-Umgebung darstellen und interaktiv bedienen, etwa über VR-Controller. Damit erleichtert diese Umsetzung die Einbindung von WebRTC-basierten Videostreams in komplexe immersive Szenarien, etwa bei VR-basierten Design Reviews oder Trainingssituationen, in denen ein fließender Übergang zwischen virtueller Umgebung und realer Videokommunikation gefordert ist [3].

Ergänzend bietet die Plattform eine umfassende Webbrowser-Integration, sodass auch Teilnehmer ohne spezifische AR- oder VR-Hardware unmittelbar über einen Standard-Webbrowser (z.B. Chrome, Firefox oder Edge) an den Sitzungen teilnehmen können. Dadurch werden Zugänglichkeit und Reichweite deutlich erhöht, da nahezu jedes internetfähige Gerät als Einstiegspunkt dienen kann. Neben der einfachen Nutzung profitieren Entwickler zudem von einer einheitlichen Weboberfläche, die konsistente Nutzererfahrungen über verschiedene Geräte hinweg ermöglicht. Dies steigert die Flexibilität der SHARE-Plattform erheblich und trägt maßgeblich zur Nutzerakzeptanz und breiten Anwendbarkeit bei.

6 Zusammenfassung und Ausblick

Die vorgestellte Videostream-basierte VR/AR-Kollaboration eröffnet erhebliche Potenziale für die Fernunterstützung bei der industriellen Wartung und vergleichbaren Services. Durch Echtzeitübertragung visueller Informationen zwischen AR- und VR-Nutzenden werden die Unterstützungsprozesse beschleunigt, Fehler reduziert und die Effizienz der Zusammenarbeit erhöht. Im Vergleich zu anderen Ansätzen wie 3D-Modell-basierter oder volumetrischer Streaming-Kollaboration bietet der Videostream-Ansatz entscheidende Vorteile hinsichtlich Flexibilität, Kosteneffizienz und leichter Implementierung. Die technische Umsetzung mittels Microsoft HoloLens2 und hochauflösenden 360-Grad-Kameras ermöglicht sowohl immersive Darstellungen in VR als auch die intuitive Integration virtueller Inhalte in reale Arbeitsumgebungen der AR-Nutzenden.

Eine zentrale Rolle spielt dabei die entwickelte MR SHARE Applikation, welche durch Hands-Free-Bedienbarkeit und präzise platzierbare AR-Interfaces sowie Echtzeit-Videostreams besonders auf die Bedürfnisse der Nutzenden abgestimmt ist. Ergänzt wird diese durch die robuste SHARE-Plattform, welche mittels WebRTC-Technologie und spezifischen Infrastrukturkomponenten – darunter SHARE Server, WebRTC Media Server sowie Webbrowser- und VR-Integration – eine zuverlässige und flexible Kommunikation zwischen unterschiedlichsten Endgeräten ermöglicht.

Dennoch müssen praktische Herausforderungen wie Netzwerkbeschränkungen, limitierte Hardware-Ressourcen und Beleuchtungssituationen berücksichtigt und gezielt adressiert werden, um das volle Potenzial der Technologie auszuschöpfen. Zukünftige Entwicklungen könnten die bestehenden technischen Einschränkungen weiter minimieren und die Integration zusätzlicher Technologien wie künstlicher Intelligenz ermöglichen, wodurch die Videostream-basierte VR-Kollaboration weiter an Bedeutung gewinnen wird.

Die Forschungs- und Entwicklungsarbeit wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ im Projekt TeleInteraction-XR (Förderkennzeichen 02L21B550-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt der Veröffentlichung liegt beim Autor.

7 Literaturverzeichnis

- [1] Kiesel, M., Kögel, A. (2025). Ich habe etwas vorbereitet – 3D-Modell-basierte AR-/VR-Kopplung. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenste Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis. DOI
- [2] Billinghurst, M., Clark, A., & Lee, G. (2015). Augmented reality. *Foundations and Trends® in Human-Computer Interaction*, 8(2-3), 73-272. <https://doi.org/10.1561/11000000049>
- [3] Flavián, C., Ibáñez-Sánchez, S., & Orús, C. (2019). The impact of virtual, augmented and mixed reality technologies on the customer experience. *Journal of Business Research*, 100, 547-560.
- [4] Microsoft. (2020). Microsoft HoloLens 2 Technical Specifications. <https://docs.microsoft.com/en-us/hololens/hololens2-hardware>

Autor



Wulf-Misaki, Daniel

Daniel Wulf-Misaki ist studierter Volkswirt (B.Sc.) und hat im Anschluss Mensch-Computer-Interaktion (M.Sc.) studiert. Seit 2022 ist er als Produkt Manager für die oculavis GmbH tätig und ist u.a. für die operative Leitung Gestaltung von Forschungsprojekten zuständig.

Mehr als oberflächlich: Live Volumetric Streaming für die AR/VR-Kollaboration

Markus Kiesel¹, Daniel Wulf-Misaki²

¹ CMC-Kiesel GmbH, ² oculavis GmbH

Zusammenfassung

Live Volumetric Streaming ist eine von mehreren Technologien, mit denen eine AR/VR-Kollaboration umgesetzt werden kann. Der Beitrag beschreibt Systemarchitektur, technologische Herausforderungen sowie eine prototypische Umsetzung. Die Ergebnisse zeigen, dass Live Volumetric Streaming insbesondere zur besseren Orientierung und Kontexterfassung bei der AR/VR-Kollaboration in der Industrie beitragen kann.

1 Einleitung

Im Forschungsprojekt TeleInteraction-XR wurde ein System zur interaktiven, telepräsenten Zusammenarbeit zwischen Expertinnen oder Experten in einem Backoffice und Servicepersonal im Einsatzfeld entwickelt. Ziel war ein möglichst natürlicher und unterbrechungsfreier Wissensaustausch über räumliche Distanz hinweg. Das Projekt untersuchte hierzu verschiedene technologische Varianten, die sich hinsichtlich Datenquellen, Interaktionsmöglichkeiten und Immersionsgrad unterscheiden [1].

Im Fokus dieses Beitrags steht der Ansatz des Live Volumetric Streamings, der die reale Umgebung des Servicepersonals im Feld mithilfe von Tiefensensoren dreidimensional erfasst, in ein digitales Volumenmodell umwandelt und in Echtzeit an den Experten in der VR-Umgebung überträgt. Im Unterschied zur klassischen Videoübertragungen [2] oder zu statischen 3D-Modellen [3] ermöglicht diese Technologie eine dynamische, echtzeitfähige und perspektivisch freie Sicht in die Einsatzsituation.

Anstatt Informationen mühsam über Sprache, Fotos oder Videos zu übermitteln, nutzt die Person im Feld ein mobiles AR- oder XR-Gerät mit Tiefensensorik. Das damit generierte 3D-Modell wird in Echtzeit an den entfernten Experten bzw. die Expertin gesendet, der oder die die Szene virtuell aus unterschiedlichen Blickwinkeln analysiert, Anomalien erkennt und präzise Handlungsanweisungen erteilen kann – etwa zur Demontage oder Justierung einzelner Komponenten. Der Experte bzw. die Expertin erhält somit ein vollständiges räumliches Verständnis, das über die Möglichkeiten herkömmlicher Videoübertragung hinausgeht. Zudem können digitale Markierungen, Annotationen und digitale Handlungsanweisungen die Kommunikation ergänzen.

2 Grundlagen und Stand der Technik

2.1 Sensorik

Zur Erfassung volumetrischer Daten kommen verschiedene Sensortechnologien zum Einsatz, die geometrische und visuelle Informationen in Echtzeit liefern. Gängige Systeme umfassen RGB-D-Kameras, Time-of-Flight-(ToF)-Sensoren, strukturierte Lichtsysteme und LiDAR-Technologien [4] [5].

Ein etabliertes – und für die Entwicklungen im Projekt TeleInteraction-XR als Vorzugsvariante ausgewähltes – System im Bereich Augmented Reality ist die Microsoft HoloLens, die mittels integriertem Tiefensensor und Spatial-Mapping-Algorithmen ein kontinuierlich aktualisiertes Mesh der Umgebung erzeugt. Dieses dient der Objekterkennung, dem Handling von Verdeckungen (Occlusion) und der präzisen Platzierung virtueller Inhalte. Trotz geringerer Auflösung im Vergleich zu spezialisierten Tiefenkameras bietet die HoloLens durch autarkes Inside-Out-Tracking eine robuste, mobile 3D-Erfassung [5] [6].

Weitere relevante Systeme zur mobilen volumetrischen Erfassung sind:

- **Azure Kinect:** Hochauflösende RGB-D-Kamera mit Weitwinkelobjektiv, besonders für präzise Nahbereichserfassung geeignet.
- **Intel RealSense:** Portable Kameraserie mit Varianten wie D435i und L515, geeignet für flexible Anwendungsszenarien.
- **LiDAR-Sensoren in Consumer-Geräten** (z. B. iPad Pro, iPhone Pro): Erlauben eine mobile Datenerfassung ohne Zusatzhardware.
- **Stereoskopische Verfahren:** Nutzung von zwei oder mehr RGB-Kameras zur Tiefenerkennung, etwa durch Googles ARCore Depth API.

Neben der Hardware ist auch die Software-Pipeline entscheidend. Während Geräte wie die HoloLens die komplette Rekonstruktion direkt auf dem Gerät ausführen (On-Device), übertragen andere Systeme Rohdaten oder Punktwolken zur externen Weiterverarbeitung. Abhängig vom Einsatzszenario werden die Daten entweder als Punktwolken oder als weiterverarbeitete Oberflächen-Meshs genutzt.

Die Auswahl geeigneter Sensortechnologie richtet sich nach Anwendungskontext, erforderlicher Mobilität, Datenqualität, Energieverbrauch und Echtzeitanforderungen. Für industrielle Szenarien sind Systeme notwendig, die kontinuierlich aktualisierte volumetrische Daten bereitstellen und nahtlos mit AR-/VR-Umgebungen zu synchronisieren sind.

2.2 Mesh-Erzeugung und Datenverarbeitung

Die Echtzeit-Erzeugung dynamischer 3D-Meshes aus Sensordaten stellt hohe Anforderungen an Verarbeitungsgeschwindigkeit und Stabilität. Der Prozess erfolgt in mehreren Schritten:

1. **Erfassung von Punktwolken:** Aus den Eingangsdaten werden kontinuierlich Punktwolken generiert.
2. **Vorverarbeitung:** Rauschreduktion, Glättung und Normalenschätzung verbessern die Qualität und Stabilität der späteren Geometrie.
3. **Mesh-Erzeugung:** Algorithmen wie Marching Cubes, Poisson Surface Reconstruction oder Delaunay-Triangulierung erzeugen daraus stabile 3D-Modelle.

Die Microsoft HoloLens 2 zeigt eine effiziente Verarbeitung: Ihr Spatial Mapping rekonstruiert die Umgebung laufend auf dem Gerät. Neue Bereiche werden automatisch integriert, bestehende aktualisiert. Das System basiert auf voxelbasierten Partitionierungen und sogenannten Spatial Surface Observers, was gezielte und ressourcenschonende Aktualisierungen ermöglicht [4] [5] [6].

3 Systemarchitektur und technische Umsetzung

3.1 Gesamtarchitektur

Die im Projekt TeleInteraction-XR entwickelte Systemarchitektur besteht aus mehreren miteinander verknüpften Kernkomponenten:

- **Kollaborations-Backend-Server**
Das zentrale Backend empfängt die vom AR-Nutzer erfassten volumetrischen Daten, verarbeitet sie in Echtzeit und stellt sie dem VR-Client zur Verfügung. Neben dem Mesh-Routing synchronisiert der Server zusätzliche Datenquellen, darunter Transformationsdaten aus 3D-Geometrien und Avatare. Dadurch entsteht eine konsistente, synchrone Datenbasis für alle Beteiligten.
- **CMC ViewR als VR-Client**
Zur Darstellung und Interaktion dient die VR-Plattform CMC ViewR. Sie erlaubt die dreidimensionale Visualisierung und Transformation (Skalierung, Rotation, Translation) empfangener oder importierter 3D-Geometrien. Die Navigation erfolgt durch Raumbewegung (Roomscale), Teleportation oder einen Flugmodus – letzterer eignet sich besonders für großflächige oder komplexe Szenarien.
- **Oculavis Share auf der AR-Seite**
Die AR-Komponente basiert auf der Plattform oculavis Share und wird in Form der MR Share-App auf der Microsoft HoloLens 2 eingesetzt. Diese dient sowohl als Sensorplattform zur volumetrischen Datenerfassung als auch zur AR-Ausgabe von Informationen und Annotationen. Die erfassten Mesh-Daten werden kontinuierlich in Echtzeit an den Backend-Server und von dort an die VR-Anwendung weitergeleitet.

Die Architektur ist offen konzipiert, um verschiedene VR-Systeme zu integrieren. Im Projekt kamen beispielsweise die HTC Vive Pro und die Meta Quest 3 zum Einsatz, je nach spezifischem Nutzungsszenario – von stationärer industrieller Anwendung bis hin zu mobilen Einsatzkontexten.

Ein entscheidender Vorteil des Systems besteht darin, dass beide Nutzer – AR- und VR-Seite – dieselbe volumetrische Datenbasis verwenden. Dies erlaubt exakte, positionsgenaue Annotationen und eine unmissverständliche Kommunikation.

3.2 Erfassungssystem

Die Auswahl eines geeigneten Erfassungssystems fokussierte sich frühzeitig auf AR/XR-Headsets, da alternative Technologien aufgrund gravierender Nachteile für den Einsatz zur Fernunterstützung in der Industrie ausschieden:

- **LiDAR-Sensoren**, wie sie zum Beispiel im iPad Pro verbaut sind, liefern sehr präzise Tiefendaten, benötigen aber externe Verarbeitungseinheiten für die Mesh-Erzeugung, was ihre Mobilität einschränkt.
- **RGB-D-Kameras** (z. B. Azure Kinect) bieten eine höhere Auflösung, sind aber ebenfalls auf externe Rechner angewiesen und somit nicht autark.
- **Stereoskopische Systeme** wie die ARCore Depth API sind preiswert, aber haben Probleme bei schwierigen Lichtverhältnissen und reflektierenden Oberflächen.

Die Eignung gängiger AR/XR-Headsets wurde anhand von AR-Fähigkeit, Kamera-Zugriffsrechten, Leistungsfähigkeit und Marktverfügbarkeit analysiert (Tabelle 1). Obwohl Geräte wie Oculus Quest 3 und Apple Vision Pro technisch leistungsfähig sind, zeigten sie im Projektkontext erhebliche Einschränkungen. Meta- und Apple-Geräten schieden aufgrund der zum Zeitpunkt der Analyse restriktiven Zugriffsrechte auf Kamera- und Tiefensensor-Daten aus.

Gerät	AR-fähig	Verfügbarkeit	Kamera-Stream	Leistungsfähigkeit
HoloLens 2	✓	✓	✓	+
Magic Leap 2	✓	(✓)	✓	+
Oculus Quest 3	(✓)	✓	(X)	+++
Pico 4	(✓)	✓	X	+++
Vision Pro	(✓)	✓	✓	+++
Pico 4 Ultra Enterprise	(✓)	✓	✓	+++

Tabelle 1: Gerätevergleich

Im Projekt TeleInteraction-XR wurde daher für die volumetrische Datenerfassung und Darstellung auf die Microsoft HoloLens 2 gesetzt. Der Fokus lag auf

einer mobilen, autarken Lösung, die ohne externe Hardware auskommt und sich direkt in industriellen Anwendungsszenarien einsetzen lässt. Die HoloLens 2 erfüllt als einziges getestetes Gerät alle relevanten Anforderungen für den produktiven Einsatz: Sie bietet integrierte Tiefensensorik, On-Device-Verarbeitung, eine robuste Spatial-Mapping-Funktionalität und direkten Zugriff auf die Kameras.

Die HoloLens 2 nutzt ein integriertes Time-of-Flight-(ToF)-System, das mittels Infrarotimpulsen Tiefeninformationen erfasst. Zusammen mit RGB-Kameras entsteht so ein adaptives Mesh, dessen Polygonauflösung automatisch an die Komplexität der Umgebung angepasst wird. Die Mesh-Generierung, Verarbeitung und Aktualisierung erfolgen vollständig Onboard – externe Recheneinheiten sind überflüssig. Änderungen in der Szene werden kontinuierlich erkannt und in das Modell integriert [4] [5] [6].

Im praktischen Einsatz zeigten sich jedoch auch folgende Einschränkungen:

- Die Auflösung des generierten Meshs ist durch die Hardware begrenzt.
- Das Sichtfeld von ca. 52° erschwert die vollständige Erfassung großer Maschinen oder komplexer Szenen.
- Die Akkulaufzeit liegt bei etwa 2 bis 3 Stunden und limitiert die Einsatzdauer.

Speziell bei der Datenerfassung traten folgende Herausforderungen auf:

- Mesh-Lücken bei schlechten Lichtverhältnissen oder stark reflektierenden Oberflächen,
- Latenzen bei hoher Rechenlast auf dem Gerät und
- die Notwendigkeit manueller Kalibrierungsschritte, um die Genauigkeit über die standardmäßig integrierten Routinen hinaus zu verbessern.

Dennoch hat sich die HoloLens 2 als derzeit einzig praxistaugliche Lösung für die mobile volumetrische Echtzeit-Erfassung bewährt – insbesondere in Verbindung mit gängigen Frameworks wie Unity AR Foundation.

3.3 Datenverarbeitung und technische Umsetzung des Volumetric Streaming

Die von dem AR/XR-Headset generierten Meshs werden über eine eigens entwickelte WebSocket-Schnittstelle an den Backend-Server übertragen. Die WebSockets erlauben eine latenzarme, stabile und plattformunabhängige

Übertragung in Echtzeit. Der Server verarbeitet die eingehenden Daten direkt weiter und verteilt sie über eine optimierte gRPC-Schnittstelle an alle aktiven VR-Teilnehmenden einer Kollaborationssession. Die gRPC-Technologie gewährleistet minimale Kommunikationslatenz und geringen Overhead. Das System unterstützt mehrere parallele Sessions. Alle jeweils registrierten Nutzenden können – unabhängig vom Standort – auf die identischen volumetrischen Daten zugreifen. Eine Übersicht ist in Abbildung 1 dargestellt.

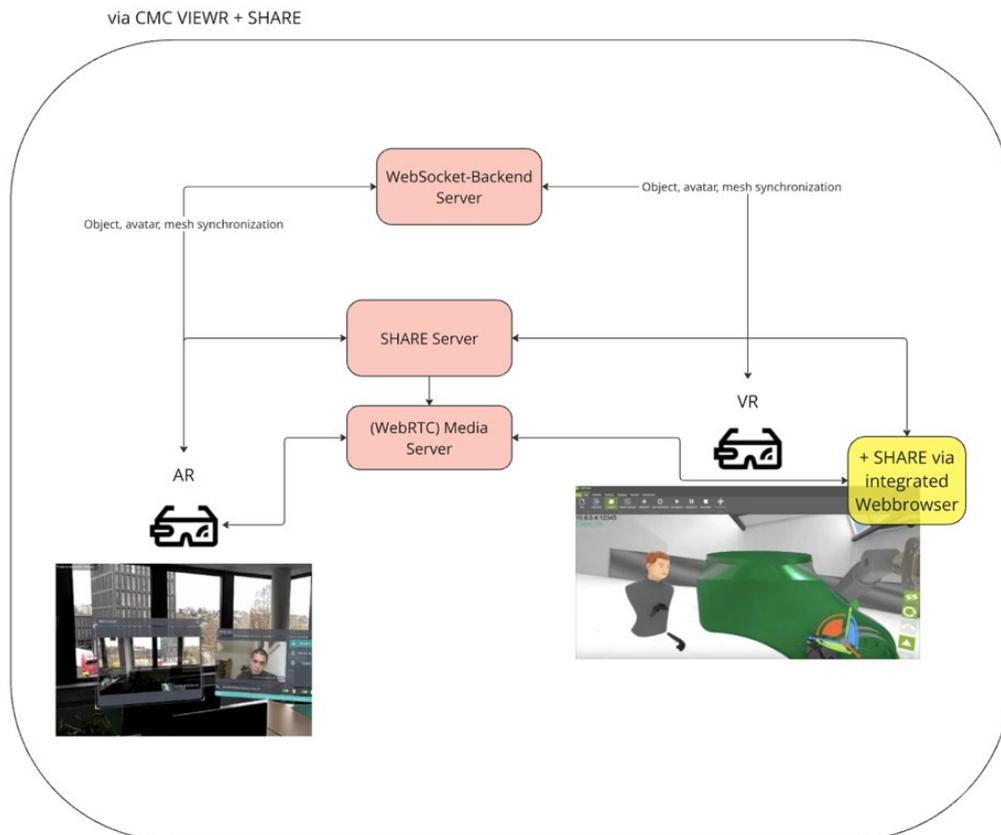


Abbildung 1: Übersicht der Systeme

Optimierungen zur Reduktion von Latenz und Ressourcenverbrauch: Um eine möglichst stabile und angenehme Erfahrung der Nutzenden zu gewährleisten – auch in Hinblick auf VR-typische Herausforderungen wie Motion Sickness – wurden zwei zentrale Optimierungsmaßnahmen implementiert:

- **On-Demand-Refresh:**
Mesh-Updates werden nicht in festen Intervallen, sondern ausschließlich bei signifikanten Veränderungen im Umgebungsmodell gesendet. Zusätzlich wird die Anzahl der Aktualisierungen pro Minute begrenzt. Dadurch sinken sowohl Datenvolumen als auch Server- und Clientlast erheblich.
- **Asynchrones Mesh-Loading und segmentierte Verarbeitung in ViewR:**
Innerhalb der ViewR-VR-Plattform erfolgt das Laden und Konvertieren der Mesh-Daten asynchron in separaten Verarbeitungsthreads. Dies führt zu einer stabileren Bildrate und reduziert wahrgenommene Latenzen deutlich.
Die Meshs sind in **Segmenten** organisiert, was eine inkrementelle Aktualisierung erlaubt: Neue Raumbereiche werden ergänzt, ohne bereits verarbeitete Mesh-Bereiche erneut zu laden. Das spart Rechenleistung und verbessert die Skalierbarkeit des Systems.

Latenz und visuelle Konsistenz: Trotz aller Optimierungen lässt sich eine geringe Verzögerung zwischen realer Umgebungsveränderung und VR-Darstellung technisch nicht vollständig vermeiden. Diese Latenz bleibt jedoch in einem akzeptablen Rahmen und beeinträchtigt die Anwendung nicht spürbar, da das virtuelle Modell persistent erhalten bleibt und kontinuierlich erweitert wird. Das Ergebnis ist ein stabiles, aktuelles 3D-Abbild der realen Arbeitsumgebung, das eine präzise Entscheidungsunterstützung, Annotation und Interaktion im Kontext der Fernkollaboration ermöglicht.

3.4 Anwendungsszenarien

Im Projekt wurden zwei reale Szenarien mit Industriepartnern umgesetzt welche im Folgenden kurz beschrieben werden. Eine vollständige Beschreibung der Anwendungsfälle findet sich in [7].

- **Remote-Fehlerdiagnose an einer industriellen Filteranlage**
Der Experte unterstützte den Techniker vor Ort bei der Analyse und Behebung einer Störung. Obwohl die Mesh-Auflösung der HoloLens nicht ausreichte, um die Ursache visuell eindeutig zu identifizieren, bot die volumetrische Übertragung wertvolle Kontextinformationen – insbesondere zur räumlichen Orientierung im VR-System (Abbildung 2).
- **Wartungseinsatz an einem XR-CAVE-System**
Auch hier wurde die Umgebung volumetrisch erfasst und in Echtzeit übertragen. Die Mesh-Qualität genügte nicht für eine vollständige Fehleranalyse, erleichterte jedoch das räumliche Verständnis und die Kommunikation zwischen den Beteiligten.

In beiden Fällen wurde deutlich, dass volumetrische Daten derzeit nicht als alleinige Informationsquelle ausreichen, jedoch eine sinnvolle Ergänzung klassischer Remote-Werkzeuge darstellen.

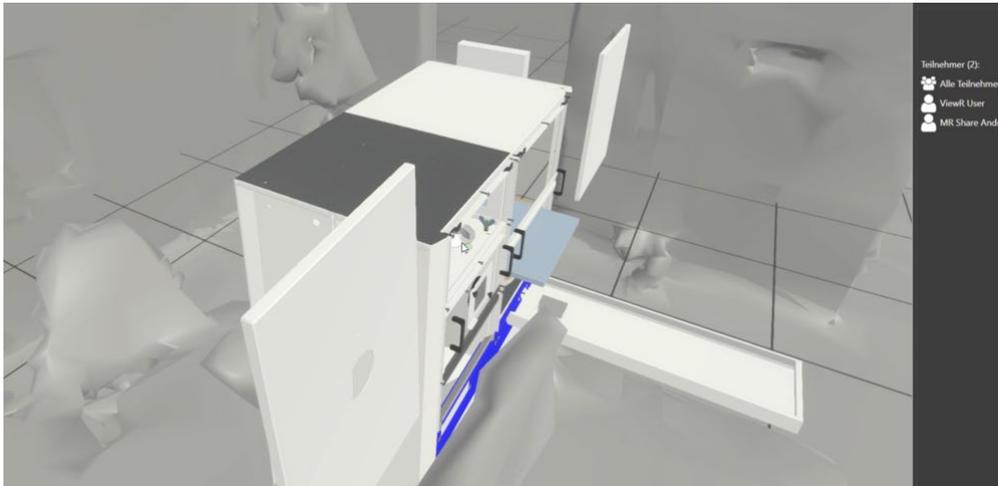


Abbildung 2: 3D-Mesh innerhalb eines volumetrischen Meshs im Anwendungsszenario Filteranlage

4 Diskussion und Ausblick

Die Ergebnisse zeigen, dass volumetrische Echtzeitdaten einen Beitrag insbesondere bei der Bereitstellung räumlicher Kontextinformationen leisten können, die über klassische 2D-Videostreams oder statische 3D-Modelle hinausgehen. Dies ist insbesondere für komplexe Wartungs- und Schulungssituationen von hohem Nutzen.

Technologische Grenzen: Die betrachteten AR-Systeme – einschließlich der Microsoft HoloLens 2 – weisen erhebliche Einschränkungen auf hinsichtlich Rechenleistung, Übertragungskapazität, Mesh-Qualität und allgemeiner Robustheit im industriellen Einsatz.

Die HoloLens 2 erwies sich als das derzeit bestgeeignete Gerät, das ohne externe Komponenten auskommt und eine zuverlässige On-Device-Mesherstellung bietet. Dennoch bilden die begrenzte räumliche Auflösung und Detailtreue der erzeugten Meshs eine erhebliche Hürde für den Industrieeinsatz.

Herstellerseitige Einschränkungen: Ein weiteres Hindernis stellen restriktive Sicherheits- und Zugriffsvorgaben führender Hersteller dar. Insbesondere bei Geräten von Meta und Apple war zum Zeitpunkt des Forschungsprojektes der direkte Zugriff auf Rohdaten wie Tiefensensorwerte oder Kamera-Streams nicht möglich.

Praktische Erkenntnisse aus Anwendungsszenarien: In beiden untersuchten Szenarien – Fehlerdiagnose an einer Filteranlage und Serviceeinsatz in einem XR-CAVE-System – wurde deutlich, dass volumetrische Daten allein derzeit nicht ausreichen, um komplexe Probleme zu lösen. Dennoch ermöglichten sie eine verbesserte räumliche Orientierung für den VR-Experten. Die zusätzliche Kontextualisierung trug dazu bei, die Kommunikation zu erleichtern und die Interaktion zu strukturieren.

Perspektiven für Forschung und Entwicklung: Zukünftige Entwicklungen sollten sich auf die AR-Hardware sowie Software- und Streaming-Architekturen konzentrieren. Besonders die Auseinandersetzung mit folgenden Aspekten erscheint vielversprechend:

- Leistungsfähigere AR-Geräte mit besserer Mesh-Auflösung,
- offene Schnittstellen für direkten Sensorzugriff,
- effizientere inkrementelle Datenübertragung,
- KI-gestützte Mesh-Optimierung und Objektsegmentierung zur Reduktion von Latenzen und Rechenlasten.

5 Fazit

Live Volumetric Streaming bietet ein hohes Innovationspotenzial für die AR/VR-Kollaboration. Der aktuelle Stand der Technik erlaubt bereits sinnvolle Einsätze zur räumlichen Kontextunterstützung. Für eine vollumfängliche Nutzung in der industriellen Praxis sind jedoch gezielte Weiterentwicklungen auf Hardware-, Software- und Systemebene erforderlich. Mit zunehmender Reife der Technologie wird es möglich sein, Live Volumetric Streaming als integralen Bestandteil digitaler Fernunterstützungslösungen zu etablieren.

Die Forschungs- und Entwicklungsarbeit wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ im Projekt TelemInteraction-XR (Förderkennzeichen 02L21B550-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt der Veröffentlichung liegt beim Autor.

6 Literaturverzeichnis

- [1] Kögel, A. (2025). Wer und was ist wo? - Synchronisation von Virtueller und Erweiterter Realität. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenste Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis. DOI
- [2] Wulf-Misaki, D. (2025). Kamera, Licht, Action! - Videostream-basierte AR/VR Kollaboration. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenste Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis. DOI
- [3] Kiesel, M., Kögel, A. (2025). Ich habe etwas vorbereitet – 3D-Modell-basierte AR-/VR-Kopplung. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenste Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis. DOI
- [4] Microsoft (2023). HoloLens 2 hardware specifications. Online verfügbar unter: <https://learn.microsoft.com/de-de/hololens/hololens2-hardware> (letzter Zugriff: 31.03.2025).
- [5] Hübner, P., Clintworth, K., Liu, Q., Weinmann, M., Wursthorn, S. (2020). Evaluation of HoloLens Tracking & Depth Sensing for Indoor Mapping Applications. Sensors, 20(4), 1021.
- [6] Microsoft Corporation (2022). Spatial Mapping overview. <https://learn.microsoft.com/de-de/windows/mixed-reality/design/spatial-mapping>
- [7] Berger, M., Holz, N., Ruffert, D., Kögel, A., (2025). TeleInterACTION! – VR/AR-Kopplung im Praxistest. In: Bullinger-Hoffmann, A.C. (2025). Telepräsenste Fernunterstützung in der Industrie. aw&I – Wissenschaft und Praxis. DOI

Autoren



Kiesel, Markus

Markus Kiesel ist studierter Maschinenbauingenieur (M.Eng.) Seit 2016 ist er als Unternehmer im Bereich der industriellen grafischen Simulation tätig. Mit seinem Team entwickelt er Anwendungen welche weltweite Verbreitung finden.



Wulf-Misaki, Daniel

Daniel Wulf-Misaki ist studierter Volkswirt (B.Sc.) und hat im Anschluss Mensch-Computer-Interaktion (M.Sc.) studiert. Seit 2022 ist er als Produkt Manager für die oculavis GmbH tätig und ist u. a. für die operative Leitung und Gestaltung von Forschungsprojekten zuständig.

TeleInterACTION! – VR/AR-Kopplung im Praxistest

Madeleine Berger¹, Nicole Holz², Danny Rüffert³, Alexander Kögel³

¹ ULT AG

² IMSYS Immersive Systeme GmbH & Co. KG

³ Professur Arbeitswissenschaft und Innovationsmanagement, TU Chemnitz

Zusammenfassung

Die im Projekt TeleInteraction-XR entwickelte Kollaboration zwischen einem VR- und einem AR-Nutzenden wurden in zwei Anwendungsfällen mit den Projektpartnern ULT AG und IMSYS Immersive System GmbH & Co. KG getestet. Die Ergebnisse bestätigen die grundsätzliche Eignung des Konzepts und zeigen Möglichkeiten für Weiterentwicklungen auf.

1 Einleitung

Die beiden Anwendungspartner im Projekt TeleInteraction-XR ULT AG und IMSYS GmbH & Co. KG stehen wie viele andere Unternehmen vor der Herausforderung, in einem schwierigen Umfeld ihre Wettbewerbsfähigkeit zu erhalten. Ihre Geschäftsmodelle basieren zu einem erheblichen Teil auf technikbezogene Services wie Inbetriebnahmen und Wartung. Diese Services müssen effizient und ohne lange Wartezeiten für die Kundinnen und Kunden erbracht werden. Zugleich gilt es, Fachkräfte für solche Aufgaben zu binden bzw. neu zu gewinnen. Auf beide Zielsetzungen – reaktionsschnelle, effiziente Services und Arbeitgeberattraktivität – zahlt das Vermeiden von Dienstreisen ein [1].

Vor diesem Hintergrund wurde – nachdem die Anwendungspartner bereits bei der Anforderungserhebung mitwirkten – der im Projekt entwickelte Demonstrator eines VR/AR-Kollaborationssystemen in je einem Anwendungsfall getestet.

2 Anwendungsfälle

2.1 Filterwechsel bei lufttechnischen Anlagen (ULT)

Die ULT AG befasst sich seit vielen Jahren mit der Produktion, dem Vertrieb sowie mit Forschung und Entwicklung im Bereich von lufttechnischen Anlagen. Die Vermarktung des vielfältigen Portfolios erfolgt neben dem nationalen und europäischen Markt auch stark in Asien und den USA. Das heißt auch, dass Kundinnen und Kunden mit Aftersales-Services weltweit zu betreuen sind.

Sowohl in den Bereichen Inspektion, Wartung und Instandsetzung als auch für Anlernprozesse besteht ein großer Bedarf für eine bessere Fernunterstützung durch Telepräsenz und Teleinteraktion. Dies betrifft unter anderem den Austausch von Filterelementen an den von der ULT AG hergestellten lufttechnischen Anlagen. Eine fehlerhafte Reparatur oder Wartung von Abluftanlagen – z. B. durch einen unzureichend unterwiesenen Betreiber – kann schwerwiegende Folgen nach sich ziehen (z. B. Brände, chronische Lungenschäden).

Ziel der ULT AG war deshalb die Entwicklung und Erprobung einer VR/AR-Kollaboration bei der eine in einer VR-Umgebung agierende Expertin bzw. ein Experte (Remotearbeitsplatz bei der ULT AG) einen Techniker oder eine Technikerin bzw. Servicemitarbeitende bei der Kundschaft unterstützen kann. Abbildung 1 zeigt beispielhaft eine Filteranlage, wie sie für die Absaugung von Laserstaub und -rauch eingesetzt wird.

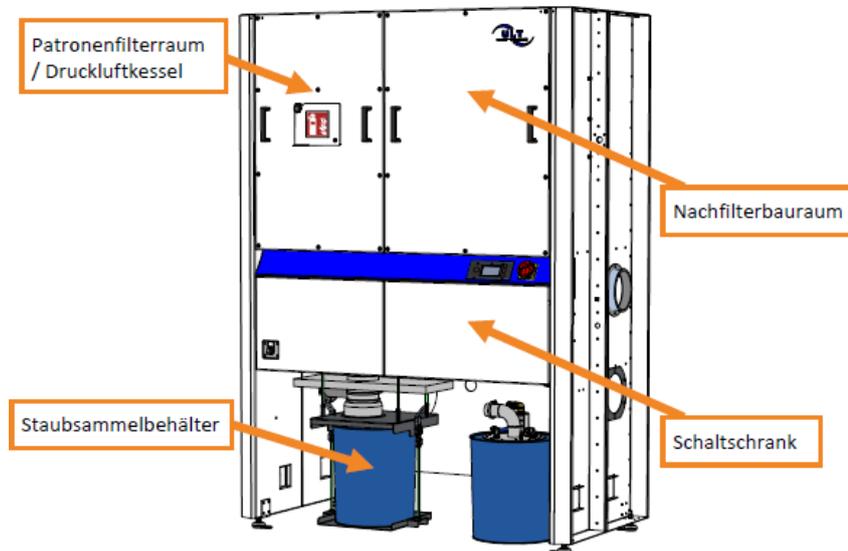


Abbildung 1: Filteranlage der Firma ULT AG

Für die Testung der VR/AR-Kollaboration wurde eine typische Kundensituation bei der Wartung an einer Filteranlage nachgestellt: Ein Techniker oder eine Technikerin ist beauftragt, die Filterelemente einer Anlage zu tauschen. Ihm oder ihr fehlen jedoch die notwendigen Kenntnisse zur korrekten Handhabung der spezifischen Anlage. Die Person vor Ort benötigt Unterstützung durch einen Experten oder eine Expertin. Beide verbinden sich über die entwickelte VR/AR-Teleinteraktion. Der Experte bzw. die Expertin in der VR-Umgebung weist den Techniker oder die Technikerin über das AR-Headset an und erklärt verbal sowie durch Manipulation des virtuellen Umgebungsmodells die Handlungsschritte. Die Technikerin oder der Techniker setzt diese Schritte am realen Objekt um.

2.2 Projektoren-Montage in einer 3-Seiten-Cave (IMSYS)

Die IMSYS GmbH & Co. KG ist seit über 20 Jahren eine der führenden Agenturen für Beratung, Planung und Projektierung von digitalen Arbeitsumgebungen, die auf Mixed Reality (XR) basieren. Eines ihrer Aufgabenfelder liegt im Aufbau kundenspezifischer Systeme wie Mehrseiten-Caves, Powerwalls oder LED-Projektionswände. Obwohl IMSYS eine hohe Expertise im Bereich XR-Technologien besitzt, war bisher die XR-basierte Fernunterstützung

eigener Mitarbeitender oder von Kundinnen und Kunden bei der Montage solcher Systeme noch nicht möglich.

Ziel der IMSYS war die Entwicklung und Erprobung einer VR/AR-Kollaboration, mit der ein in einer VR-Umgebung agierender Experte oder eine Expertin (Remotearbeitsplatz bei IMSYS) einem vor Ort bei der Kundschaft agierenden Servicemitarbeitenden unterstützen kann.

Für den Test der VR/AR-Kollaboration wurde der Anwendungsfall Aufbau einer 3-Seiten-Cave gewählt (Abbildung 2). Das Hauptaugenmerk liegt dabei auf der korrekten Ausrichtung der Projektoren (Abstand zur Wand, Bildgröße, Trapezverzerrung etc.). Das Servicepersonal vor Ort benötigt dafür die Unterstützung durch einen Experten oder eine Expertin aus der Firmenzentrale.



Abbildung 2: 3-Seiten-Cave der IMSYS GmbH & Co. KG

3 Erprobung

Für beide Anwendungsfälle wurde die VR/AR-Kollaboration in einer Probandenstudie getestet. Bei den Probanden handelte es sich um Mitarbeitende der jeweiligen Unternehmen (ULT AG: $n = 6$, IMSYS: $n = 4$). Den Probanden waren die eigenen Produkte bekannt, aber die Expertise hinsichtlich der jeweiligen Prozesse im Anwendungsfall fehlte. Die Rolle der in der VR-Umgebung agierenden Expertin bzw. des Experten übernahm jeweils eine mit dem

Prozess vertraute Person. Der technische Aufbau und das prinzipielle Vorgehen waren für beide Anwendungsfälle identisch.

Mithilfe der entwickelten VR/AR-Kopplung konnte die mit einer VR-Brille Meta Quest3 ausgerüstete Fachkraft die Situation, wie sie sich Mitarbeitenden vor Ort darstellte, in der VR-Umgebung realitätsnah miterleben. Dabei standen in der VR zur Verfügung:

- die vorgefertigten Modelle der jeweiligen Testumgebung (Basisvariante),
- zusätzliche 360°-Videostreams (zweite Umsetzungsvariante) und
- die in Echtzeit dreidimensional erfasste, erweiterte Umgebung (dritte Umsetzungsvariante) sowie
- zusätzlich ein Videofeed mit der Sicht der vor Ort agierenden Person.

Die Person an der Anlage nutzte ein HoloLens2 Seethrough-AR-Headset. Darin waren ein Avatar der in der VR agierenden Person und das von ihr manipulierbare Modell der Testumgebung (Filterschrank oder Projektor) zu sehen. Die Lage und Ausrichtung der Objekte und Avatare wurde in Echtzeit bidirektional zwischen der VR-Umgebung des Experten bzw. der Expertin und der AR-Umgebung des technischen Personals synchronisiert und lagegerecht an der Realumgebung der Anlage ausgerichtet. Beide Personen konnten mittels Gesten auf Objekte zeigen und die aus der Ferne unterstützende Person konnte Handlungen direkt am (virtuellen) Objekt vorführen. Darüber hinaus stand ein bidirektionaler Sprachkanal zur direkten Kommunikation zur Verfügung.

Je nach Anwendungsfall war das Ziel der Teleinteraktion, den Filter am realen Objekt zu wechseln (ULT AG) bzw. den Projektor korrekt auszurichten (IMSYS). Eine Besonderheit des Tests bestand darin: Nach dem ersten Durchgang wurden die Rollen getauscht, so dass auch der vormalige AR-User, nun mit dem vermittelten Wissen über den Vorgang ausgestattet, sich in der Expertenrolle wiederfand. Auf diese Weise wurde implizit der Erfolg des Wissenstransfers geprüft.

Abbildung 3 zeigt schematisch den Versuchsablauf. Abbildung 4 zeigt beispielhaft eine Real- und die zugehörige VR-Szene aus der Erprobung im Anwendungsfall Filterwechsel. Für den Anwendungsfall 3-Seiten-Cave illustriert Abbildung 5 die Situation vor Beginn des Einbaus des Projektors aus VR-Sicht.

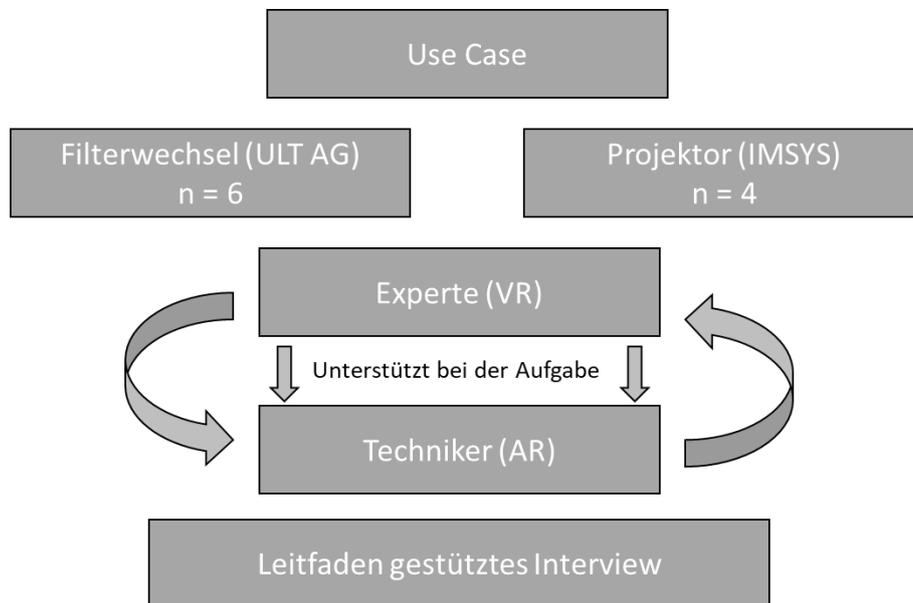


Abbildung 3: Versuchsablauf

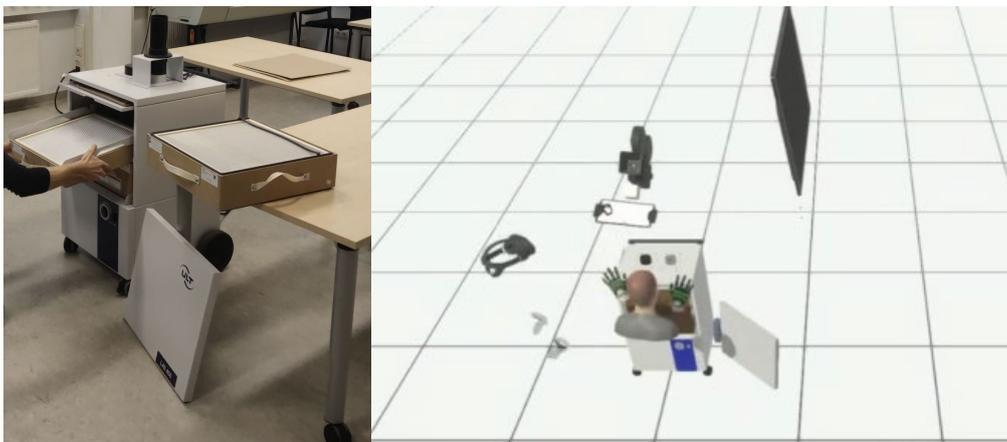


Abbildung 4: Real- und VR-Szene aus der Erprobung im Anwendungsfall Filterwechsel

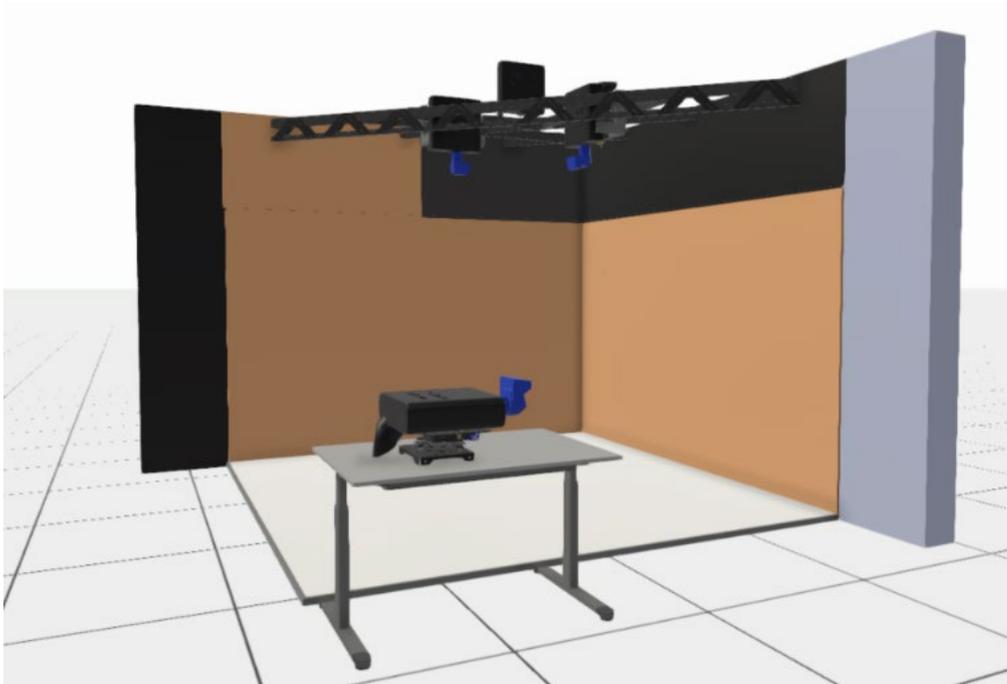


Abbildung 5: Projektor vor dem Einbau in die 3-Seiten-Cave in der VR-Sicht

Nach der Erprobung wurde mit den Probanden der beiden Anwendungsfälle ein leitfadengestütztes Interview zu deren Erfahrungen durchgeführt. Die Leitfragen betrafen

- die Zufriedenheit mit dem Prozessablauf mit VR/AR-Kollaboration,
- die Vor- und Nachteile der VR/AR-Kollaboration,
- gewünschte zusätzliche Funktionalitäten sowie
- die Zufriedenheit mit Stil und Detailtreue der Darstellung.

Aus den Interviews konnten Aussagen zur Eignung des VR/AR-Kollaborationssystems, zu Herausforderungen und Hemmnissen sowie zu gewünschten Zusatzfunktionen abgeleitet werden.

4 Ergebnisse

4.1 Validierung des Kollaborationskonzepts

Ein wesentliches Ergebnis der Praxistest ist, dass die vor Ort agierenden Service-Mitarbeitenden eine auf einer VR/AR-Kopplung basierende Unterstützung durch einen remote agierenden Experten bzw. eine Expertin als sehr hilfreich empfinden. Die Arbeit vor Ort wird erleichtert und Fehler werden reduziert.

Besonders die direkte visuelle Unterstützung in Echtzeit wurde von den Anwendenden als große Hilfe empfunden. Der Vorteil der Verwendung von AR-Brillen gegenüber anderen Lösungen (Video-Call mit Smartphones etc.) liegt in der „Hands Free“-Nutzung bei gleichzeitiger Sichtbarkeit aller Informationen im jeweils aktuellen Arbeitsbereich.

Besonders hervorgehoben wurde, dass die Expertinnen und Experten durch den Wegfall von Reisezeiten flexibler einsetzbar sind und mehrere Unterstützungsaufträge pro Zeiteinheit erledigen können. Dies führt zu einer besseren Ressourcennutzung und schnellerer Reaktionsfähigkeit des Service.

Eine Novität bei der Fernunterstützung im industriellen Umfeld war die Darstellung des jeweiligen Gegenübers als Avatar. Hier bestätigten die Testpersonen, dass die stilisierte Darstellung von Kopf, Oberkörperansatz und Händen ausreicht bzw. gegenüber einer realitätsnahen Darstellung sogar bevorzugt wird. Auch andere Studien bestätigen, dass aufwandsärmere Teilkörpervisualisierungen genügen, um hohe Präsenz und Akzeptanz zu erzeugen [2]. Wichtig sei vor allem, dass die Position des Gegenübers, seine Blickrichtung und die Orientierung der Hände erkennbar sind.

4.2 Herausforderungen und Hemmnisse

Trotz der Zustimmung zur prinzipiellen Tauglichkeit der VR/AR-Kollaboration zeigen die Interviews zu den Tests auch eine Reihe noch bestehender Herausforderungen:

Technische Infrastruktur: In vielen Fällen fehlt es an grundlegender Netzwerkinfrastruktur und ausreichender Hardware, um den reibungslosen Einsatz der Technologie zu gewährleisten.

Datensicherheit und Datenschutz: Die sichere Übertragung von Daten sowie die Einhaltung von Datenschutzvorgaben sind kritische Punkte, die

organisatorische Anpassungen oft auch auf Kundenseite erfordern. Dies betrifft nicht nur die Absicherung der über das öffentliche Internet übertragenden Daten, sondern auch grundlegende Fragen zum Einsatz Kamera-basierter Systeme in Hinblick auf Geschäftsgeheimnisse und die Wahrung der Persönlichkeitsrechte.

Realitätstreue, Detailliertheit und Aktualität der in die VR-Umgebung übertragenen Daten:

- *CAD generierte 3D-Modelle*: Eine korrekte Unterstützung durch VR-Experten oder Expertinnen setzt voraus, dass digitale 3D-Modelle mit den realen Gegebenheiten übereinstimmen. In der Praxis können jedoch Bauabweichungen oder nachträglich ein- und angebrachte Zubehörteile, die nicht dokumentiert wurden, Probleme verursachen. Hierbei können die im Projekt eingesetzten Umsetzungsvarianten 360°-Videostream und Live Volumetric Streaming der Realumgebung Abhilfe schaffen. Diese unterliegen jedoch eigenen, nachfolgend genannten Beschränkungen.
- *360°-Videostream*: Durch die fixe, nicht mit der Position des Experten oder der Expertin gekoppelte Perspektive entstehen Erkennungs- und Verdeckungs-Probleme, die eine häufige Umpositionierung der aufnehmenden Kamera nach sich ziehen. Außerdem fehlen Tiefeninformationen.
- *Live Volumetric Stream*: Durch den Einsatz der ursprünglich für Navigationszwecke entwickelten Sensorik der HoloLens2 waren die so erfassten dreidimensionalen Abbilder der Realumgebung einfarbig, niedrig auflösend und nicht in Echtzeit übertragbar. Sie helfen, sich in der Gesamtsituation vor Ort zu orientieren, befriedigen jedoch nicht, wenn sie als alleinige Datenquelle genutzt werden.
- Die Kombination aus vorgefertigten und live erfassten 3D-Daten mit 360°-Kamerastreams kann einige der individuellen Schwächen der Umsetzungsvarianten abmildern, jedoch nicht vollständig aufheben.

Geschäftsmodell und Kundenbindung: Die VR/AR-Kopplung kann (auf der AR-Seite vor Ort) auch vom Kunden oder der Kundin selbst genutzt werden, um ohne Hilfe eines Mitarbeitenden des Ausrüsters bestimmte Serviceaufgaben assistiert durchzuführen. Die trägt zu einem Wissensaufbau auf Kundenseite bei und kann langfristig den Bedarf an Servicepersonal reduzieren, was potenziell Auswirkungen auf die Erlöse durch das Servicegeschäfts und die Kundenbindung haben kann.

Arbeitsumweltbedingungen: Örtliche Gegebenheiten wie Beleuchtung, Lärm und Platzverhältnisse spielen eine entscheidende Rolle für die Qualität der bei der AR/VR-Kollaboration übertragenen Bilder und Sprache.

Zusätzliche Einschränkungen liegen vor, wenn persönliche Schutzausrüstung (z. B. Schutzbrille, Gehörschutz, Atemschutzmaske) getragen werden muss. Das Tragen eines AR-Headsets ist in diesen Fällen kaum kombinierbar.

4.3 Zusätzlich gewünschte Funktionalitäten

Bereits während der Entwicklungsarbeiten im Projekt TeleInteraction-XR wurden Wünsche der Nutzenden, die im Rahmen der Anforderungsanalyse erhoben wurden, aufgegriffen und umgesetzt. Dies betraf Basisfunktionalitäten wie das Markieren und Hervorheben von Bauteilen in der VR-Umgebung.

In den Interviews nach den Tests wurden von den Testnutzenden noch folgende weitere Zusatzfunktionen gewünscht:

- Zusätzliche **Annotationsmöglichkeiten** in der VR- bzw. AR-Umgebung (für technische Parameter, Prüfkriterien etc.).
- Möglichkeit eines **Videomitschnitts** zur Prozessprotokollierung (in Verbindung mit Annotationen)
 - für die Beweisführung bei Reklamationen oder Regressforderungen und
 - als Wissensspeicher für künftige Einweisungen etwa im Fall des Wechsels von Mitarbeitenden (Wissensmanagement von teils eher implizitem Wissen).
- **Physikalische Simulation** des 3D-Modells bezüglich
 - der Körpergrenzen, um Kollision zu erkennen.
 - der Verbindungen und Gelenke, um Freiheitsgrade beim Bewegen nachvollziehen zu können (für Instruktionen wie „Die Abdeckungen nach oben und dann nach vorn ziehen, bis es einrastet“.)
- **Vorgefertigte Animationen**, um die Wiederholgenauigkeit bei der Demonstration von Vorgängen zu verbessern (auch in Kombination mit der Simulation der Objektphysik).
- **Multi-User-Modus**, mit dem auch mehr als zwei Personen (z. B. zusätzliche Trainees) die VR/AR-Kollaboration nutzen können (Basistechnologien sind vorhanden, wurden aber bisher nicht explizit getestet).

5 Fazit

Tests in zwei Anwendungsfällen zeigen, dass das im Projekt TeleInteraction-XR als Demonstrator entwickelte VR/AR-Kollaborationssystem geeignet ist, Servicetätigkeiten durch Experten bzw. Expertinnen aus der Ferne zu unterstützen. Die Servicekräfte vor Ort empfinden die Unterstützung

insbesondere durch die gegenüber Telefonaten zusätzlichen visuellen Möglichkeiten und die „Hands Free“-Arbeitsweise als hilfreich. Die Unterstützung leistenden Expertinnen und Experten sparen durch vermiedene Geschäftsreisen wertvolle Zeitressourcen und können schneller und effizienter reagieren.

Zukünftig sollten die Systeme weiterentwickelt werden, um bestehende Hemmnisse, wie die Infrastruktur- und Sicherheitsanforderungen, zu überwinden und zusätzliche Funktionen, wie das Dokumentieren durch Videomitschnitt, zu implementieren.

Die Forschungs- und Entwicklungsarbeit wurde vom Bundesministerium für Bildung und Forschung im Fachprogramm „Zukunft der Wertschöpfung“ und der Fördermaßnahme „Innovative Arbeitswelten im Mittelstand“ im Projekt TeleInteraction-XR (Förderkennzeichen 02L21B550-4) gefördert und vom Projektträger Karlsruhe (PTKA) betreut. Die Verantwortung für den Inhalt der Veröffentlichung liegt beim Autor.

6 Literaturverzeichnis

- [1] Rüffert, D., Kögel, A., Löffler, T. (2025) So fern und doch so nah – Telepräsenes Arbeiten in der Industrie. In: Bullinger-Hoffmann, A.C. (2025) Telepräsenes Fernunterstützung in der Industrie - Abschlusspublikation der Projekte PraeRI und TeleInteraction-XR
- [2] Brade, J; Kögel, A.; Fuchs, C.; Klimant, P.: Impact of first person avatar representation in assembly simulations on perceived presence and acceptance. Proceedings of the 15th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 1: GRAPP, 27.02.-29.02.2020, Valletta, Malta, p. 17-24. - Science and Technology Publications Lda, Scitepress digital library, 2020. DOI: 10.5220/0008878700170024.

Autoren

Madeleine Berger

Madeleine Berger arbeitet seit 2017 im Bereich Forschung und Entwicklung bei der ULT AG Umwelt-Lufttechnik.

Aufgabengebiete sind die Netzwerkkoordination zu Partnern aus Forschung, Industrie und Projektträgern für öffentlich geförderte Forschungsprojekte und industrielle Entwicklungsprojekte. Die ULT AG entwickelt und fertigt Anlagen zur

effektiven Beseitigung luftgetragener Schadstoffe oder zur Reduzierung der Luftfeuchte.



Nicole Holz

Dipl.-Ing. Nicole Holz ist Teamleiterin für Consulting und Customer Success bei IMSYS und beschäftigt sich seit 20 Jahren mit der Entwicklung und Implementierung innovativer Softwarelösungen sowie der digitalen Transformation. Ihr Fokus liegt auf Prozessoptimierung, Technologieintegration und nachhaltigen Lösungen in den Branchen Automotive, Aerospace und Machinery.



Rüffert, Danny

Danny Rüffert studierte Sports Engineering an der Technischen Universität Chemnitz. Seit 2014 arbeitet er als wissenschaftlicher Mitarbeiter an der Professur Arbeitswissenschaft und Innovationsmanagement im Cluster Industrial Engineering. Seine Aufgaben umfassen die Themenfelder Alterssimulation und Ergonomie sowie die Forschung an Mensch-Maschine-Schnittstellen.



Kögel, Alexander

Alexander Kögel studierte Medientechnik (B. Eng.) und Medienmanagement (M. Eng) an der HTWK Leipzig. Seit 2017 ist er an der TU Chemnitz als wissenschaftlicher Mitarbeiter tätig. Schwerpunkte sind VR-Anwendungen, das Präsenz-Empfinden in virtuellen Umgebungen und der Erschließung der Potenziale der Digitalisierung, Vernetzung und Industrie 4.0.