

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].

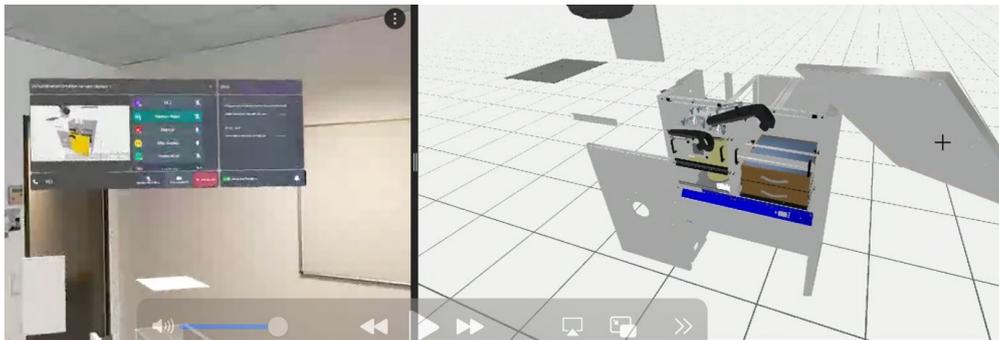


Abbildung 1: Demonstrator-Testsession mit der MR SHARE Applikation – AR-Nutzerperspektive (links) und korrespondierende VR-Ansicht (rechts).

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.

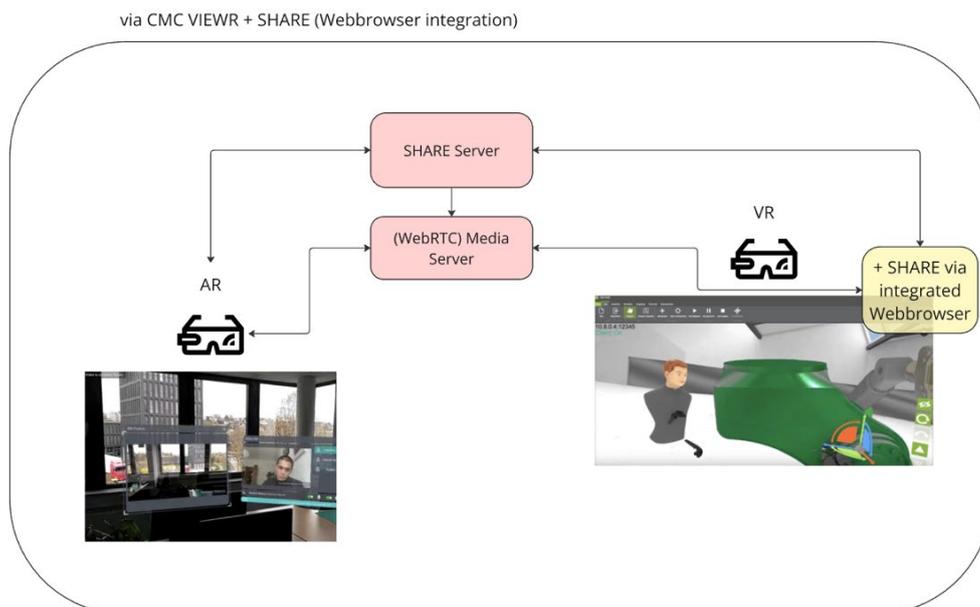


Abbildung 2: Technischen Infrastruktur für die Videostream-basierte VR/AR-Kollaboration

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.