

Zweiseitige Kurzfassung des wesentlichen fachlichen Inhalts - HIVAM

Das Verbundprojekt HIVAM verfolgt das Ziel, eine immersive Telemedizin-Plattform zu entwickeln und zu erforschen, die auf Mixed-Reality-Technologien (MR) basiert und die Expertise realer Gesundheitsexpert:innen mit hybriden Avatar- und Assistent-Modellen verbindet.

Die Hochschule Düsseldorf (HSD) konzentrierte ihre Arbeiten im Vorhaben HIVAM hauptsächlich auf die Entwicklung, Erprobung und Dokumentation eines vollständigen 3D-Avatar-Workflows, der den Weg von der Datenerfassung bis zur Echtzeiteinbindung in gängige Game-Engines abdeckt. Ausgangspunkt waren Photogrammetrieaufnahmen realer Personen in einem Multi-Kamera-Setup (zunächst Full-Body, später ergänzend hochauflösende Kopfskans). Die daraus abgeleiteten 3D-Daten wurden mit eigenen und etablierten Werkzeugen zu animierbaren, für die Telemedizin geeigneten digitalen Personenmodellen aufbereitet. Ziel war eine technisch robuste, reproduzierbare Pipeline mit klar definierten Arbeitsschritten, die nachgelagerte Interaktions- und Evaluationskomponenten mitberücksichtigen.

Für die Verarbeitung der 3D-Scans implementierte die HSD spezialisierte Tools, die wiederkehrende Arbeitsschritte automatisieren und qualitative Konstanz sichern. Dazu zählen Verfahren zur Übertragung von Deformationen zwischen topologiegleichen Modellen, automatisierte Gelenkplatzierung, Transfer von Weight-Paintings sowie die Generierung von ARKit-kompatiblen Blendshapes und Visemes. Ein Alignment-Verfahren verbessert die Anpassung zwischen Basis-Mesh und Scan-Geometrie und schafft die Grundlage für präzise, lippen-synchrone Gesichtsanimationen. Ergänzend wurden 3D-Wrapping und Textur-Rückprojektion eingesetzt, um eine animierbare Topologie mit realitätsnahen Oberflächentexturen zu verbinden.

Auf Engine-Ebene wurden zwei Pfade systematisch verfolgt. In Unreal wurde der Mesh-to-MetaHuman-Prozess getestet und für Realtime-Einsatzbedingungen angepasst; darauf aufbauend wurde eine Audio-basierte LipSync-Umsetzung für MetaHuman-Gesichter eingesetzt. Parallel dazu wurden Avatare in Unity über die DigitalHuman-Pipeline integriert. Shader-, Material- und Modellanpassungen zielten auf VR-Tauglichkeit und Performance; die spätere Portierung von HDRP auf URP ermöglichte die Darstellung der 3D-Avatare auf mobilen Endgeräten. Damit liegen verschiedene Bereitstellungswege für unterschiedliche Zielplattformen vor, ohne Kernfunktionen wie Mimik- und Gestikanimationen einzuschränken.

Zusätzlich wurde eine weitere Pipeline mit bilddatengetriebenen Komponenten erarbeitet. Ein trainiertes Generative-Adversarial-Network ermöglichte ein DeepFace-Overlay einer realen Person zur Echtzeitübertragung von Gesichtsdeformationen auf einem 3D-Avatar. Ergänzend prüfte die HSD multiple Verfahren zur schnellen Avatarerzeugung aus 2D-Bildern (u. a. SMPL/SMPL-X, PIFuHD, Avaturn, In3D, ReadyPlayerMe) und band die Meshcapade-API ein, um 3D-Modelle aus Einzelfotos zu

generieren. In Summe entstand ein Methodenbaukasten, der je nach Datenlage, Qualitätsanforderung und Laufzeitbudget unterschiedliche Erzeugungspfade erlaubt.

Zur Erhöhung des Ausdrucksrepertoires wurden Motion-Capturing-Aufnahmen durchgeführt und die Inverse Kinematik des HIVAM Demonstrators für stabile Interaktionsposen optimiert. Ein vorhandenes Verhaltensmodell eines Projektpartners wurde so angepasst, dass provisorische Gesten durch realistischere Animationen ersetzt werden konnten. Ein ergänzender MetaHuman-Workflow ermöglicht, die qualitativ hochwertigen Gesichtsanimations-Werkzeuge aus dem MetaHuman-Ökosystem auch mit photogrammetrie-basierten Körpermodellen zu kombinieren.

Die Erstellung der 3D-Avatare enthielt umfangreiche Capturing-Sessions (u. a. Full-Body-Capturing mit 106 Kameras und Head-Capturing mit 466 Bildern). Eigens entwickelte Maya-Skripte beschleunigten die Topologieanpassungen sowie den Transfer von Blendshapes und Visemes. Die eingesetzten Rigging- und LipSync-Verfahren verbesserten die Animationsqualität gegenüber einem manuellen Workflow mit ähnlichen Zeitbeschränkungen. Ein zusätzlicher 3D-Avatar wurde für die telemedizinische Plattform eines Projektpartners erstellt und im Projektkontext integriert. Darüber hinaus wurden Drittsysteme (z. B. D-ID) für KI-gestützte Gesprächs-Avatare getestet, ohne diese in den Kernworkflow zu übernehmen.

Die Ergebnisse wurden in Demonstratoren und Veröffentlichungen sichtbar gemacht. Es wurde ein Paper auf der „Mensch und Computer 2023“ zu HIVAM veröffentlicht, an welchem die HSD mit beteiligt war. Ebenso konnte eine in HIVAM von der HSD entwickelte 3D-Avatar-Pipeline für generische XR-Workflows auf einer internationalen Konferenz eingereicht werden. Diese Publikation wurde mit dem Best Paper Award ausgezeichnet (XR Salento 2025, Best Paper). Begleitend erfolgten regelmäßige Projekttreffen und thematische Workshops (ELSA, Value-Sensitive-Design), ein gemeinsames Konsortialtreffen im HIVE-Lab, die Teilnahme am BMBF-Vernetzungssymposium 2024, ein Konsortialtreffen mit VDI/VDE Anfang 2024, die Präsentation eines Demonstrators auf der GI VR/AR (September 2024) sowie die Teilnahme am BMBF-Abschlusstreffen (Februar 2025). Eine Masterarbeit (12/2024) dokumentierte und verglich zwei Verfahren zur Avatarerzeugung.

Insgesamt liegt mit dem HIVAM-Beitrag der HSD eine modulare, dokumentierte Ende-zu-Ende-Kette zur Erzeugung, Aufbereitung und Echtzeiteinbindung digitaler Menschen vor. Sie berücksichtigt die für telemedizinische Interaktionen zentralen Anforderungen wie Mimik, Gestik, lippen-synchrone Sprachanimation und plattformübergreifende Laufzeit-Performance. Durch den Einsatz von Photogrammetrie, Engine-spezifischen Pipelines und bei Bedarf selektiv eingebundenen Machine-Learning-Verfahren entsteht eine technische Grundlage, die den Einsatz in stationären VR-Umgebungen ebenso wie auf mobilen Endgeräten unterstützt und sich für nachgelagerte Studien zu Embodiment, sozialer Präsenz und Akzeptanz eignet.