



netidee

PROJEKTE

Cognitive AR

Zwischenbericht | Call 15 | Projekt ID 5121

Lizenz: CC-BY-SA

Inhalt

1. Einleitung.....	3
2. Status der Arbeitspakete.....	3
2.1. Arbeitspaket 1 – AP1: Detailplanung und Formales am Projektstart.....	3
2.2. Arbeitspaket 2 – AP2: Design & Technologie Evaluation.....	3
2.3. Arbeitspaket 3 – AP3: Computer Vision / Artificial Intelligence -Komponenten.....	4
2.4. Arbeitspaket 4 – AP4: Plattform Implementierung.....	4
2.5. Arbeitspaket 5 – AP5: Augmented Reality Apps.....	4
2.6. Arbeitspaket 6 – AP6: Demo-Deployment.....	5
2.7. Arbeitspaket 7 – AP7: Dissemination.....	5
3. Umsetzung Förderauflagen.....	5
4. Zusammenfassung Planaktualisierung.....	6
5. Öffentlichkeitsarbeit/ Vernetzung.....	6
6. Eigene Projektwebsite.....	7

1. Einleitung

Der folgende Bericht erläutert die Fortschritte von „Cognitive AR“, das erste Projekt der größeren *Cognitive XR* Initiative. Das Ziel von Cognitive AR ist der Aufbau einer Software Plattform und Client-Server Struktur, für das Integrieren von Augmented Reality Brillen mit einer dezentralen Edge Computing Infrastruktur. Zuerst möchten wir uns für die Unterstützung der Netidee bedanken die es uns ermöglicht dieses spannende Projekt und in unseren Augen Zukunftsweisende Thema zu bearbeiten. Die Netidee Förderung ermöglicht uns sowohl die Open Source Entwicklung der Software, als auch das Publizieren unsere Ergebnisse in Blogs und auf Konferenzen. Wir liegen im Projektplan etwa drei Monate zurück, und haben deshalb um eine Projekt- verlängerung angesucht, die Sie uns dankenswerterweise genehmigt haben. Wir sind mit dem Fortschritt aber insgesamt sehr zufrieden, und sind überdies zuversichtlich den im Antrag beschriebenen Projekterfolg mit Ende des Jahres erreicht zu haben. Bei etwa 40% der Ergebnisse werden wir den definierten ausgezeichneten Erfolg erreichen, und bei 60% den minimalen Erfolg.

2. Status der Arbeitspakete

Zusammenfassung: Unser wichtigstes Ziel war es, bis März 2021 das Design der Plattform zu finalisieren, die ersten Technologieevaluierungen abzuschließen (AP2), und erste Prototypen für die kritischen Systemkomponenten entwickelt (AP3, AP4) zu haben. Diesen Meilenstein haben wir erreicht, und auch eine Publikation aus unserem Design und ersten Resultaten verfasst, die im März auf der IEEE VR erschienen ist (AP7). Wir haben anschließend die Prototypen genutzt um eine simple Augmented Reality Proof of Concept Anwendung zu entwickeln (AP5). Seit April sind wir nun damit beschäftigt die Systemkomponenten zu erweitern bzw. robuster zu machen (AP3, AP4), weitere Demo Anwendungen (AP5) zu entwickeln, und in unserem Labor ein Testbett für die Edge-Infrastruktur auf zu setzen (AP6).

2.1. Arbeitspaket 1 – AP1: *Detailplanung und Formales am Projektstart*

Der Projektstart wurde rasch und ohne weitere Probleme abgewickelt.

2.2. Arbeitspaket 2 – AP2: *Design & Technologie Evaluation*

Unsere initiale Fragestellungen waren: a) gibt es existierenden KI-basierten Videoanalyse Technologien und wenn ja wie sind diese nutzbar? b) braucht man wirklich Edge Computing, oder könnte man unsere Anwendungsfälle auch mit Cloud-Basierten Ansätzen lösen? c) Welche Netzwerkprotokolle eignen sich für KI-basierte Videoanalyse? d) Welche Lösungen gibt es zum tracken von Objekten im 3D-Raum anhand von 2D Kamerabildern?

Die Antworten auf diese Fragen konnten wir durch eine wissenschaftliche und systematische

Herangehensweise gut erarbeiten, und haben diese sogar als Paper auf der IEEE VR Publiziert (siehe Öffentlichkeitsarbeit)

2.3. Arbeitspaket 3 – *AP3: Computer Vision / Artificial Intelligence -Komponenten*

Viele Anwendungsfälle in unserer Problemdomäne erfordern das Lokalisieren von Objekten im 3D Raum anhand von statischen Kameras. Etwa zum Positionieren von Autos um deren Positionen an AR Geräte in der Nähe zu senden. Diesen wesentlichen Teil von AP3 haben wir durch unser System Cyber-Physical Object Positioning fertig. Wir verwenden dafür OpenCV und diverse KI Modelle für die Tiefenerkennung. Wir können Objekte in Echtzeit erkennen und die Positionsdaten von 2D in 3D Positionen übersetzen. Diese werden über das Netzwerk an umliegende Geräte gesendet. Ein Demonstration dazu haben wir in unserer “Proof of Concept Anwendung”. Obwohl damit mit AP3 Ende im Plan waren, hat CPOP unter gewissen Annahmen funktioniert (z.B. dass Objekte am Boden stehen). Wir haben CPOP seit April nun erweitert, dass es auch frei-fliegende Objekte im Raum tracken kann.

2.4. Arbeitspaket 4 – *AP4: Plattform Implementierung*

Ein wesentlicher Bestandteil von AP3 und AP4 ist ein Server für KI-basierte Echtzeit- Videoanalyse zum verarbeiten von Kamera-Feeds (sowohl statische Kameras als auch der AR Geräte), der sich auf kleinen Edge-Computing Geräten (wie etwa dem NVIDIA Jetson Xavier NX) betreiben lässt. Unsere Plattform CogStream implementiert eine Plugin- Architektur, mit der sogenannte Engines verwaltet werden können. Engines kapseln die Funktionalität eines oder mehrere KI Modelle um die Bilder der Kamera zu analysieren. Sowohl das Framework, der Server, als auch die Engines befinden sich im selben Repository, dabei sind Framework und Server dem AP4, und die Engines dem AP3 zuzuordnen. Wir haben bereits drei Engines implementiert: Objekterkennung, Gesichtserkennung, und Facial Expression Recognition (Emotionserkennung). Die HoloLens kann sich bereits zum Server verbinden um den Kamera-Feed der Brille analysieren zu lassen.

2.5. Arbeitspaket 5 – *AP5: Augmented Reality Apps*

Als Projektergebnis 9 haben wir ein Client SDK definiert, mit dem es AR EntwicklerInnen leicht fällt mit unserem System zu kommunizieren. Diese Entwicklung hat im wesentlichen in AP5 stattgefunden. Der meiste C# client Code befindet sich noch in unserer Unity Demo App, wird aber nun im Rahmen von AP5 weiter in eigene Repositories herausgeholt um die als Bibliotheken für andere AR EntwicklerInnen zur Verfügung zu stellen. Die Herangehensweise hat sich für uns als Produktiv herausgestellt, führt aber auch dazu, dass wir noch nicht in der Lage sind mit unserem Client SDK zu werben, und daher ist unsere open source traction noch sehr niedrig.

Mittels CPOP haben wir bereits eine simple Demo Anwendung aus AP5 umgesetzt: Augmented Reality Objekt-Tracking mittels Cyber-Physical Object Positioning (CPOP). Video:

<https://www.youtube.com/watch?v=nY3PLUTVSbw>

Fast vollständig implementiert ist eine weitere Anwendung die das Video offloading mit AP4 demonstriert. Die HoloLens sendet dabei ihren Kamerafeed zu einem Edge Knoten, der Gesichter und Emotionen erkennt, und im Display die Information entsprechend Anzeigt. Diese Anwendung funktioniert bereits, und ein Demo Video ist in Arbeit

Eine Herausforderung war das Anschaffen zusätzlicher AR Brillen neben der HoloLens. Viele Anbieter haben Lieferschwierigkeiten bzw sind viele Brillen noch in Entwicklung und gar nicht Verfügbar, obwohl das im Marketing oft suggeriert wird. Wir werden daher auch den minimalen Erfolg erzielen (HoloLens Unterstützung)

2.6. Arbeitspaket 6 – AP6: Demo-Deployment

Um Projektergebnis 8 (Infrastruktur Prototyp) und 11 (End-To-End Demo) zu erreichen, haben wir in den Räumlichkeiten der TU Wien ein kleines Labor einrichten können. Dort haben wir diverse Hardware die wir von der TU Wien nutzen können aufgebaut (etwa eine HoloLens, und Hardware für Edge Knoten). Für die Bauanleitung des Edge Knotens haben wir bereits zwei Blogeinträge geschrieben (Siehe Publikationen und PR), und ein Repository angelegt das sowohl die Anleitung als auch Scripts zum Automatisieren der Konfiguration enthält. Eine erste Version davon ist online, und wir werden diese nach und nach erweitern.

Outdoor Umsetzung (z.B. erkennen von Autos, oder Montierung von Kameras auf Ampeln) hat sich im Laufe der Entwicklung durch technische Herausforderungen als nicht im Projektumfang möglich herausgestellt. Hier werden wir daher den minimal Erfolg mit einem Indoor Prototypen erreichen.

2.7. Arbeitspaket 7 – AP7: Dissemination

Siehe Öffentlichkeitsarbeit/Vernetzung.

3. Umsetzung Förderauflagen

Die Förderauflagen zu unserem Projekt wurden bei Projektanfang mit einer Ergänzung zum Projektantrag welche Inhaltliche und Personelle Unklarheiten auf geklärt hat, und am 23.10.2020 an Dr. Andreas Koman gesendet wurde erfüllt.

4. Zusammenfassung Planaktualisierung

Neues Projektende: 31.12.2021 (von 1.10.2021)

Durch die beruflichen Situation von Thomas Rausch (TRa) konnten die ursprünglich geplanten Stunden nicht mehr im selben Ausmaß geleistet werden. Insgesamt 207 Stunden von Thomas wurden auf die Praktikanten Silvio Vasiljevic (SVa), Christian Stippel (CSt), und Matthias Hürbe (Mhü) aufgeteilt. Da sich diese jedoch noch in Ausbildung befinden, haben wir entsprechend den Aufwand der Arbeitspakete neu geschätzt und etwas erhöht.

Desweiteren waren unsere Praktikanten ursprünglich bei AP7 (Dissemination) nicht eingeplant, haben sich aber im Zuge der ersten wissenschaftlichen Publikation viel eingebracht und Stunden geleistet, und haben wesentlich bei einem technischen Blogartikel mitgeholfen. Bei AP8 werden unsere Praktikanten ebenfalls mithelfen (ursprünglich nicht geplant). Bei beiden APs aben sich daher die Stunden von SVa, CSt, und MHü deutlich erhöht.

Die **Projektkosten** haben sich jedenfalls **nicht wesentlich geändert**.

5. Öffentlichkeitsarbeit/ Vernetzung

Beschreibung der bereits erfolgten Öffentlichkeitsarbeit oder Vernetzung, bzw. Beschreibung des Plans künftiger Aktivitäten

Die folgenden Publikationen, Blog Artikel, und PR Materialien lassen sich unmittelbar für die strukturierte Umsetzung der Projektergebnisse 3-6 nutzen:

- [Radio Interview zu Cognitive AR auf “Start Me Up” Radio Magazin](#)
- **Drei** Inhaltliche Blogartikel auf der **Netidee Community** Website:
 - [Announcing CognitiveXR](#)
 - [Systemdesign, erste Prototypen und Ergebnisse](#)
 - [Unsere Fortschritte im CognitiveAR Labor](#)
- **Zwei** Blogartikel mit Technischen Inhalten auf der CognitiveXR Website
 - [Computer Vision with PyTorch on the Jetson Xavier NX](#)
 - [Setting up an Intel RealSense Camera on an NVIDIA Jetson Xavier NX](#)
- [Ein Vortrag und Paper zu CognitiveAR auf dem IEEE VR DISCE 2021 Workshop](#)
 - Dieser Vortrag hat zu einem Meeting für eine potentielle Kollaboration mit der UCLA und dem [OpenPTrack Team](#) geführt
- Social Media:
 - 163 follower auf [LinkedIn](#)
 - 37 follower auf [Twitter](#)
 - [YouTube channel](#) mit insgesamt 556 Video views und 14 subscribers

Des weiteren planen wir in Kooperation mit dem VRVis aus der zweiten AR App eine Wissenschaftliche Publikation auf der IEEE VR 2022 ein zu reichen.

6. Eigene Projektwebsite

<https://cognitivexr.at>