

# Sprachkommunikation in Virtual Reality

Studiengang: BSc in Informatik | Vertiefung: Computer Perception and Virtual Reality

Betreuer: Prof. Markus Hudritsch

Experte: Dr. Harald Studer ( Integrated Scientific Services AG)

Um die Einsatzmöglichkeiten von Virtual Reality anschaulich zu demonstrieren, hat das CPVR-Labor der BFH eine Demonstrationswelt erstellt. In dieser VR-Welt können die Spieler gemeinsam mit Objekten interagieren und kleine Spiele spielen. Für eine effiziente und sinnvolle Kollaboration fehlte aber noch die sprachliche Kommunikation. Diese wurde nun durch die Umsetzung dieser Bachelorthesis ermöglicht.

## Einleitung

Vor dem Beginn dieser Bachelorthesis konnten die Spieler nur miteinander kommunizieren, indem sie Gesten wie Zeigen, Winken oder Hüpfen verwendet haben. Diese Einschränkung machte es sehr umständlich zu kollaborieren und führte dazu, dass auf Drittanwendungen wie TeamSpeak oder Skype ausgewichen werden musste. Werden diese VoIP-Programme verwendet, können die Sprecher aber nicht in der VR-Welt durch 3D-Soundquellen oder Mundanimationen hervorgehoben werden.

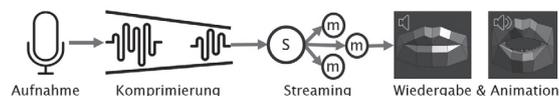
## Ziele

Das Hauptziel dieser Bachelorthesis ist es, Gespräche zwischen den Spielern zu ermöglichen und die Kommunikation zu verbessern. Das Hauptziel beinhaltet folgende drei Teilziele:

1. Aufnahme & Komprimierung
2. Audio verteilen an Mitspieler
3. Gesichtsanimation

## Aufnahme & Komprimierung

Für dieses Teilziel wurden verschiedene, existierende VoIP-Lösungen für Unity3D evaluiert. Mit der besten Lösung aus der Evaluation wurde eine gute Balance zwischen Qualität und Datenrate gefunden. Das eingesetzte Modul «Voice Chat» hat eine gekapselte Struktur und verwendet bestehende Unity3D-Objekte. Durch den Einsatz dieser Objekte war auch die Umsetzung des Teilzieles «3. Gesichtsanimation» möglich.



Ablauf von der Aufnahme bis zur Wiedergabe mit der Mundanimation

## Audio verteilen an Mitspieler

Nach dem die Daten vom Mikrophon aufgenommen und für den Transport komprimiert wurden, müssen sie an die anderen Mitspieler in der VR-Welt geschickt werden. Hier wird der Audio-Stream vom Client an den Server geschickt. Der Server leitet dann die Daten weiter an die einzelnen Clients. Diese Lösung wurde ohne Erweiterung umgesetzt.

## Gesichtsanimation

Damit man im Virtuellen Raum erkennen kann welche Person gerade spricht, muss das Gesicht des Sprechers dynamisch zum erhaltenen Audiosignal animiert werden. Nach der Evaluationsphase wurde die Unity-Erweiterung «SALSA» vom Entwickler «Crazy Minnow Studio» integriert.

«SALSA» wertet beim Ausgeben des empfangenen Audiosignals die Daten aus und berechnet anhand der Amplitude die korrekte Mundöffnung. Die Mundöffnungen werden mit Morph Targets definiert und durch die Erweiterung automatisch interpoliert.

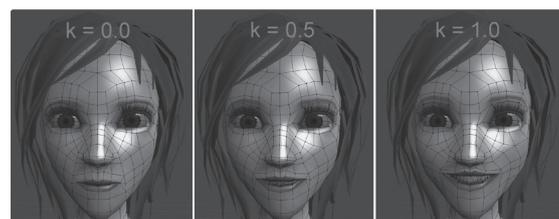
## Fazit

Mit Hilfe der geschriebenen Erweiterung kann das CPVR-Labor endlich eine VR-Welt vorstellen welche eine robuste und performante Lösung für die Kommunikation unter den Mitspielern bietet.

Die geschaffene, solide Codebasis und Dokumentation kann von kommenden Studenten verwendet werden um auf dem Projekt aufzubauen und es zu erweitern.



Martin Käser



Beispiel eines angewendeten Morph Targets auf den Mundbereich