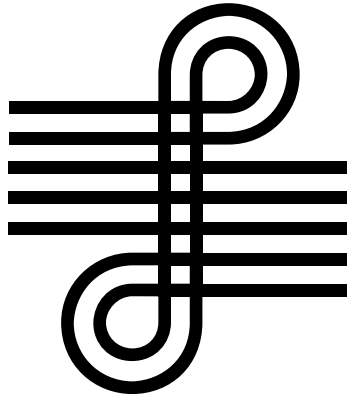


Hochschule für Musik Karlsruhe

IMWI - Institute für Musikinformatik und Musikwissenschaft



Bachelorarbeit zum Thema:

Interaktion mit Audio-Daten in VR mithilfe von Unity und der htc Vive

Zur Erlangung des Grades Bachelor of Arts

Vorgelegt von:

Manuel-Philippe Hergenröder, 12085

mail@manuelhergenroeder.de

Betreuender Dozent:

Prof. Dr. Damon T. Lee

Studiengang:

Musikinformatik (Hauptfach) / Musikwissenschaften (Nebenfach)

Abgabe:

xx.07.2020

1	Inhaltsverzeichnis	
2	Einleitung	3
3	Grundlagen	3
3.1	Virtual Reality	3
3.1.1	Historie und Einordnung zu verwandten Konzepten	3
3.1.2	Spezielle Anforderungen an VR-Software	4
3.2	Diskrete Fourier-Transformation	4
4	Vorstellung VrAudioSandbox	5
4.1	Architektur	5
4.1.1	Unity als Laufzeitumgebung	6
4.1.2	NAudio Bibliothek	6
4.1.3	FFTW3 für die DFT	6
4.2	Implementation	7
4.2.1	FFT und Audio Engine	7
4.2.2	Visualisierung der Spektrum-Daten	7
4.2.3	User Interface	7
5	Kritische Betrachtung der Ergebnisse	7
6	Fazit	7
7	Literatur	7
8	Sonstige Quellen	7
9	Abbildungsverzeichnis	8
10	Eidesstattliche Erklärung	9

2 Einleitung

Virtual Reality (VR) bietet das Potential unserem Geist direkten und erlebbaren Zugriff auf digitale Daten – losgelöst von der Umgebung in der wir uns aufhalten – zu geben. Dabei

In dieser Arbeit sollen die Möglichkeiten der Interaktion mit Audio-Daten im virtuellen Raum am Beispiel der Darstellung und Manipulation von FFT-Audio-Daten mithilfe der htc Vive und den dazugehörigen Motion-Tracking-Controllern untersucht werden. Gegenstand der Arbeit ist die praktische Implementation dieses Konzepts und **<Herausforderungen, Bewertung, Ausblick>**.

3 Grundlagen

3.1 Virtual Reality

3.1.1 Definition, Historie und Abgrenzung

Der Begriff Virtual Reality impliziert, dass eine virtuelle Realität mithilfe einer Simulation geschaffen wird, welche für den Nutzer als real wahrgenommen wird. Bereits 1965 beschrieb der Computergrafikpionier Ivan E. Sutherland mit dem „ultimate display“ einen Raum, in dem die Materie durch den Computer gesteuert wird. Auf einen Stuhl in diesem Raum könne man sich setzen, eine Pistolenkugel wäre tödlich.¹

Dieses Ziel der kompletten Immersion konnte bisher nicht umgesetzt werden. Ein weiterer Aspekt ist aber auch nicht nur der Versuch eine komplett reale und glaubwürdige Umgebung zu schaffen, sondern

Gerade im Bereich der Hardware spielt die Entwicklung des sogenannten „Head-Mounted-Display“ (engl. für „Am-Kopf-befestigter-Bildschirm“) – oder kurz HMD – eine wichtige Rolle. Ein HMD besteht aus einem oder mehreren Displays und stellt am Kopf befestigt eine Schnittstelle zwischen der visuellen Wahrnehmung des Menschen und der Grafikausgabe eines Computers her. Ivan E. Sutherland hat 1968 ein solches System mit dem Namen „The Sword of Damocles“ mithilfe seiner Studenten realisiert.² Dies legte die Grundlagen zum einen für die stereoskopische Darstellung des im Computer erzeugten Raums in Vektorgrafik, zum anderen für die notwendige schnellen Synchronisierung der angezeigten Bilder mit den Kopfbewegungen

¹ Vgl. Sutherland, Ivan: „The Ultimate Display“ – Konferenzband: Information Processing 1965: proceedings of IFIP Congress / Wayne A. Kalenich [Hrsg.]. International Federation for Information Processing, Amsterdam u.a., Washington u.a. (1965), S. 508 – <http://worrydream.com/refs/Sutherland%20-%20The%20Ultimate%20Display.pdf>, letzter Abruf: 09.06.2020

² Vgl. Sutherland, Ivan: „A head-mounted three dimensional display“ – AFIPS '68 (Fall, part I): Proceedings of the December 9-11, 1968, fall joint computer conference, part I, December 1968 Pages 757–764, S. 757 – <http://www.medien.ifi.lmu.de/lehre/ss09/ar/p757-sutherland.pdf>, letzter Abruf: 10.06.2020

des Benutzers. Ein Sensor mit mechanischem Arm hat zunächst die Kopfposition ermittelt. Später wurde dies für bessere Bewegungsfreiheit mithilfe von Ultraschall-Emittern und -Empfängern umgesetzt.³ Dieses sogenannte „Head-Tracking“ und die zeitnahe Abstimmung der angezeigten Bilder an Kopfposition und -bewegungen ist auch heute noch ein wichtiger Aspekt bei HMDs um Übelkeit (sogenannte „Motion-Sickness“), die in unserem Gehirn entsteht, wenn widersprüchliche Reize – in diesem Fall die visuellen Informationen unseres Sehapparates und die körperliche Wahrnehmung bezüglich der Bewegung und Stellung des Kopfes verarbeitet werden – beim Benutzer vorzubeugen.

Die Entwicklung von Technologien für VR reicht weit zurück. <NASA Datenhandschuh, VR Hype 90er, 2. Welle Oculus Rift/Vive, Nintendo Virtual Boy, Gaming Industry, Professionelle Anwendungen Architektur, CAD, Ergonomie>

Ein mit der Virtual Reality verwandtes Feld ist die sogenannte „Augmented Reality“ bei der die echte Realität mit zusätzlichen Informationen angereichert wird. Im Gegensatz zur VR werden reales Umfeld und virtuelle Bestandteile vom Computer vermischt – z.B. durch ein transparentes Display oder mithilfe eines eingeblendeten Live-Kamera-Bildes. Bei einer starken Interaktion zwischen Bestandteilen realer und virtueller Welt wird zudem auch der Begriff „Mixed-Reality“ verwendet.

3.1.2 Spezielle Anforderungen an VR-Software

- Motion sickness, Latenz
- Refresh rate
- Bedienparadigma (2D UI vs intuitiver Umgang mit Objekten in VR)

3.2 Diskrete Fourier-Transformation

Die diskrete Fourier-Transformation ist ein wichtiger Algorithmus in der digitalen Signalverarbeitung und überführt Signal-Daten aus dem Zeitbereich in den Frequenzbereich. Fast-Fourier-Transform (kurz. FFT) beschreibt dabei eine Untermenge von Implementationen der DFT, die besonders effizient und schnell zu berechnen sind.

³ Ebd. S. 760-761

4 Vorstellung VrAudioSandbox

Im Rahmen dieser Arbeit wurde die Möglichkeit der Darstellung und Manipulation von Audio-Daten innerhalb der Virtual Reality praktisch implementiert. Dazu wurde neben den technischen Voraussetzungen ein Bedienkonzept zugeschnitten auf die Motion Tracking Controller der htc Vive entwickelt. Der Benutzer kann beliebiges vorhandenes Audiomaterial importieren, mithilfe des htc Vive HMDs und der dazugehörigen Motion-Controller modifizieren und das Ergebnis wieder als Audio-Datei exportieren. Die Bedienung findet dabei vollständig in VR statt.

4.1 Architektur

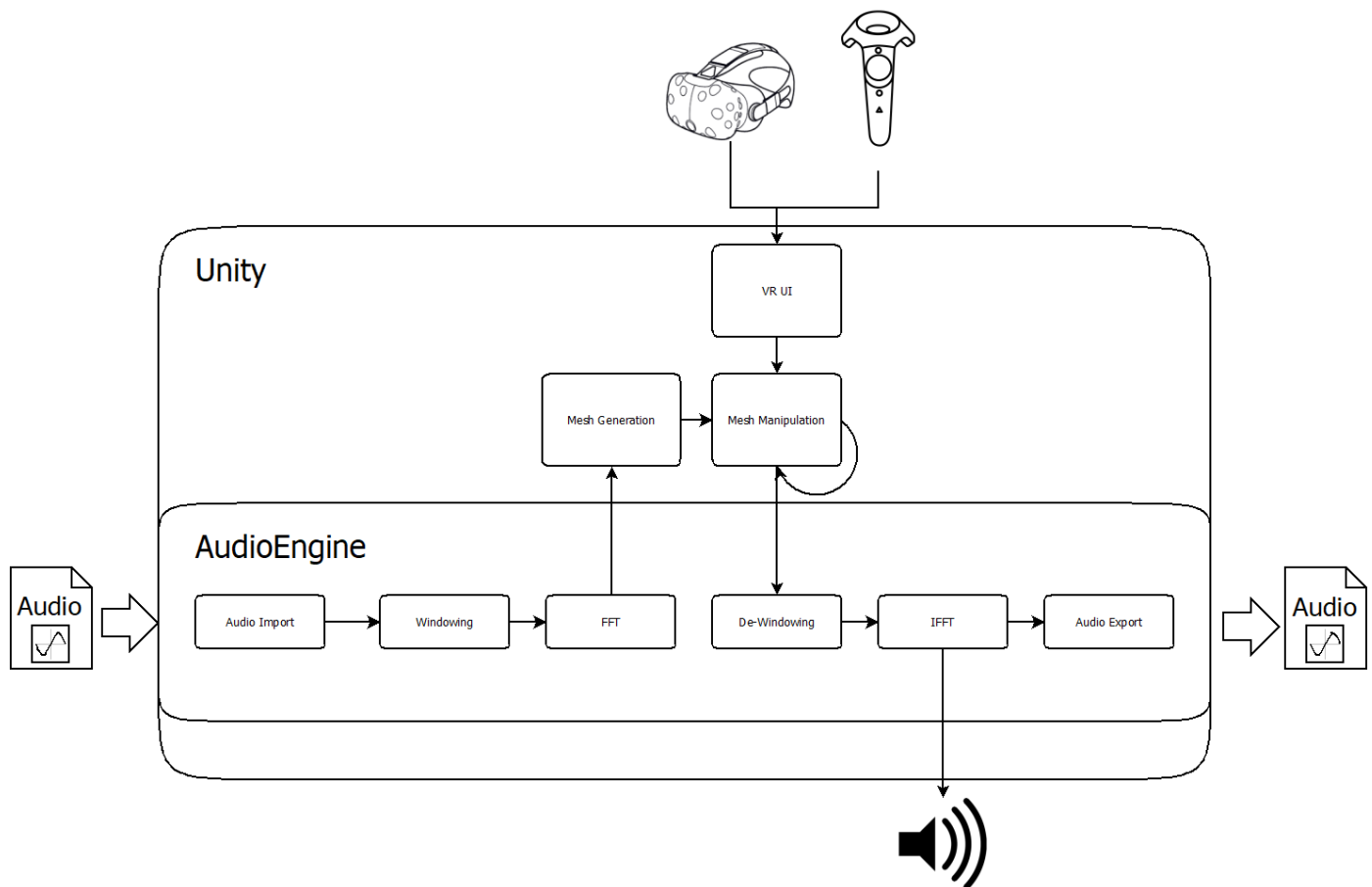


Abbildung 1 – Eine abstrakte Darstellung der Architektur - Eigene Darstellung

Das Herzstück für das Rendering und die Einbindung von SteamVR bildet dabei die Laufzeitumgebung Unity.

Im Folgenden werden die genutzten Bestandteile und externe Bibliotheken näher erläutert.

4.1.1 Unity als Laufzeitumgebung

Die Laufzeit- und Entwicklungsumgebung Unity ist eine weitverbreitete Spiele-Engine mit etwa 50% Marktanteil für den Bereich PC/Konsole/Mobile Games.⁴ Das dazugehörige Unternehmen „Unity Technologies“ (ehemals „Over the edge“) wurde 2004 von David Helgason, Nicholas Francis und Joachim Ante gegründet. Auch außerhalb der kommerziellen Spiele-Entwicklung ist Unity sehr beliebt – u.a. auch da es eine für nicht-kommerzielle Zwecke kostenlose „Personal“-Lizenz gibt, welche den Hauptteil des Funktionsumfangs von Unity bereitstellt.⁵

Unity bildet das Herzstück für das Rendering, <Erklärung Zusammenhang C#, Unity, etc>

4.1.2 NAudio Bibliothek

4.1.3 DFT mit FFTW3

FFTW (Akronym für „The Fastest Fourier Transform in the West“) ist eine freie FFT-Bibliothek zur Berechnung der diskreten Fourier-Transformation und wurde von Matteo Frigo und Steven G. Johnson am Massachusetts Institute of Technology entwickelt.⁶ Die in C und OCaml geschriebene Bibliothek ist im Quelltext verfügbar, sehr portabel und unterstützt daher viele Plattformen. Außerdem bietet sie laut den Autoren im Vergleich zu anderen Implementationen eine sehr gute Performance.⁷

Für die DFT und inverse DFT in VrAudioSandbox wird auf die „1d Discrete Fourier Transform (DFT)“ aus der FFTW3-Bibliothek mit Double-Precision zurückgegriffen, welche durch folgende mathematische Summenformel beschrieben wird:⁸

$$Y_k = \sum_{j=0}^{n-1} X_j e^{-2\pi j k \sqrt{-1/n}}$$

Die Eingabe X ist dabei ein ein-dimensionales Array komplexer Zahlen der Größe n und Y das Ausgabe-Array, wobei das k -te Element der Frequenz k/n entspricht.

⁴ Vgl. Unity Technologies: „Sie fragen sich was Unity ist? Entdecken Sie, wer wir sind, wo wir angefangen haben und wohin wir uns entwickeln | Unity“ – <https://unity.com/de/our-company>, letzter Abruf: 11.06.2020

⁵ Vgl. Unity Technologies: „Powerful 2D, 3D, VR, & AR software for cross-platform development of games and mobile apps.“ – <https://store.unity.com/#plans-individual>, letzter Abruf: 11.06.2020

⁶ Vgl. Frigo, Matteo; Johnson, Steven G.: „The Design and Implementation of FFTW3“ – Proceedings of the IEEE, Volume 93, Number 2, 2005, S. 216-231, S. 231 – <http://www.fftw.org/fftw-paper-ieee.pdf>, letzter Abruf: 13.06.2020

⁷ Vgl. FFTW: „FFT Benchmark Results“ – <http://www.fftw.org/speed/>, letzter Abruf: 13.06.2020

⁸ Vgl. FFTW: 3.3.8 Manual, Kapitel 4.8.1 „The 1d Discrete Fourier Transform (DFT)“ – http://www.fftw.org/fftw3_doc/The-1d-Discrete-Fourier-Transform-_0028DFT_0029.html, letzter Abruf: 13.06.2020

Bei der inversen Funktion handelt es sich um die gleiche Funktion ohne negatives Vorzeichen im Exponenten der DFT:

$$Y_k = \sum_{j=0}^{n-1} X_j e^{2\pi j k \sqrt{-1/n}}$$

FFTW berechnet eine unnormalisierte Transformation ohne Koeffizienten vor der Summe der DFT, d.h. eine vorwärtsgerichtete DFT gefolgt von der rückwärtsgerichteten DFT ergibt als Ergebnis die Eingabe multipliziert mit n .

4.2 Implementation

4.2.1 FFT und Audio Engine

- Implementation, Windowing, etc.
- Codebeispiele

4.2.2 Visualisierung der Spektrum-Daten

- Vertices, Triangles, Mesh
- Codebeispiele

4.2.3 User Interface

- Motion-Controller, Actions
- Konzept: Interaktion ohne verschachtelte 2D-Menüs

5 Kritische Betrachtung der Ergebnisse

- Verbesserungsvorschläge
- Ideen für weitere Werkzeuge / Funktionen

6 Fazit

7 Literatur

8 Sonstige Quellen

9 **Abbildungsverzeichnis**

Abbildung 1 – Eine abstrakte Darstellung der Architektur - Eigene Darstellung.....	5
--	---

10 Eidesstattliche Erklärung

Hiermit versichere ich, dass die vorliegende Arbeit von mir selbstständig und ohne unerlaubte Hilfe angefertigt worden ist, insbesondere dass ich alle Stellen, die wörtlich oder annähernd wörtlich aus Veröffentlichungen entnommen sind, durch Zitate als solche gekennzeichnet habe.

Ich versichere auch, dass die von mir eingereichte schriftliche Version mit der digitalen Version übereinstimmt. Weiterhin erkläre ich, dass die Arbeit in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde/Prüfungsstelle vorgelegen hat. Ich erkläre mich damit nicht einverstanden, dass die Arbeit der Öffentlichkeit zugänglich gemacht wird.

Ich erkläre mich damit einverstanden, dass die Digitalversion dieser Arbeit zwecks Plagiatsprüfung auf die Server externer Anbieter hochgeladen werden darf. Die Plagiatsprüfung stellt keine Zurverfügungstellung für die Öffentlichkeit dar.

Oberboihingen,
der xx.xx.2020

Manuel-Philippe
Hergenröder



Ort, Datum

Vorname Nachname

Unterschrift