

# Orientierung und Navigation durch immersives 3D Audio in virtuellen Welten

Samantha Povolny



MASTERARBEIT

eingereicht am  
Fachhochschul-Masterstudiengang

Digital Arts

in Hagenberg

im September 2018

© Copyright 2018 Samantha Povolny

Diese Arbeit wird unter den Bedingungen der Creative Commons Lizenz *Attribution-NonCommercial-NoDerivatives 4.0 International* (CC BY-NC-ND 4.0) veröffentlicht – siehe <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

# Erklärung

Ich erkläre eidesstattlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benutzt und die den benutzten Quellen entnommenen Stellen als solche gekennzeichnet habe. Die Arbeit wurde bisher in gleicher oder ähnlicher Form keiner anderen Prüfungsbehörde vorgelegt.

Hagenberg, am 14. September 2018

Samantha Povolny

# Inhaltsverzeichnis

<b>Erklärung</b>	<b>iii</b>
<b>Kurzfassung</b>	<b>vii</b>
<b>Abstract</b>	<b>viii</b>
<b>1 Einleitung</b>	<b>1</b>
1.1 Motivation . . . . .	1
1.2 Fragestellung . . . . .	1
1.3 Struktur der Arbeit . . . . .	2
<b>2 Binaurales Audio</b>	<b>3</b>
2.1 Definition Binaurales Audio . . . . .	3
2.2 Aufnahmeverfahren . . . . .	3
2.3 Auditive Ortung des menschlichen Ohres . . . . .	4
2.3.1 Auditive Lokalisierung . . . . .	4
2.3.2 Prinzipien der Wahrnehmungsorganisation . . . . .	5
2.3.3 Räumliche Wahrnehmung . . . . .	5
<b>3 3D Audio und Virtual Reality</b>	<b>7</b>
3.1 Psychische Aspekte . . . . .	7
3.2 Atmo und Ambience Sounds . . . . .	8
3.3 Technische Aspekte . . . . .	9
3.4 Gestalterische Aspekte . . . . .	9
3.5 Head-Mounted Display . . . . .	10
3.6 Besonderheit Head Tracking . . . . .	10
3.7 Virtual Room . . . . .	11
<b>4 Lautsprecher und Kopfhörer</b>	<b>14</b>
4.1 Lautsprecher . . . . .	14
4.2 Kopfhörer . . . . .	16
<b>5 Ausgewählte Theorien zur Orientierung</b>	<b>18</b>
5.1 Sound Charakteristiken nach McMullen . . . . .	18
5.2 Navigation durch auditive Hinweise nach Lokki und Gröhn . . . . .	19
5.3 Fazit der Theorien . . . . .	22

<b>6</b>	<b>Orientierung in virtuellen auditiven Umgebungen</b>	<b>23</b>
6.1	Aufbau einer VAE . . . . .	23
6.2	Methoden der Wegfindung . . . . .	24
6.3	Auditiv Lokalisierung in VAE . . . . .	25
6.4	Erinnerungsvermögen, Diegese und Interaktivität . . . . .	28
6.4.1	Erinnerungsvermögen . . . . .	28
6.4.2	Diegese . . . . .	28
6.4.3	Interaktivität . . . . .	29
<b>7</b>	<b>Aufbereitung von auditiven Hinweisen</b>	<b>32</b>
7.1	Kategorisierung von Klangobjekten . . . . .	32
7.2	Technische Voraussetzungen der Audiodateien für VR-Anwendungen . . . . .	33
7.2.1	Vorbereitung . . . . .	33
7.2.2	Audioeinstellungen der Unreal Engine . . . . .	34
<b>8</b>	<b>Leitlinien</b>	<b>37</b>
8.1	Wahl des Ausgabemediums . . . . .	37
8.2	Technische Voraussetzungen einer Game Engine . . . . .	37
8.3	Vertrautheit akustischer Signale . . . . .	37
8.4	Ambience Sounds . . . . .	38
8.5	Vertonung sichtbarer und interaktiver Objekte . . . . .	38
8.6	Akustische Orientierungspunkte/Landmarks . . . . .	38
8.7	Führungspunkte setzen . . . . .	38
8.8	Zusätzliche Höheninformationen . . . . .	39
8.9	Reflektive Umgebungen . . . . .	39
8.10	Interaktivität und Wiederholungen . . . . .	39
8.11	Verständlichkeit . . . . .	39
8.12	Statische und dynamische Schallquellen . . . . .	40
8.13	Bedeutung von Klangobjekte beachten . . . . .	40
8.14	Zusätzliche Informationen über die Audioebene . . . . .	40
8.15	Wechselwirkung der visuellen und auditiven Ebene . . . . .	40
<b>9</b>	<b>Beispielanalyse</b>	<b>41</b>
9.1	<i>The Gallery – Episode 1: Call of the Starseed</i> . . . . .	41
9.1.1	Kurzbeschreibung . . . . .	41
9.1.2	Analyse . . . . .	41
9.1.3	Fazit . . . . .	46
9.2	<i>Elena</i> . . . . .	46
9.2.1	Kurzbeschreibung . . . . .	46
9.2.2	Analyse . . . . .	46
9.2.3	Fazit . . . . .	49
9.3	<i>Letzte Worte</i> . . . . .	50
9.3.1	Kurzbeschreibung . . . . .	50
9.3.2	Analyse . . . . .	50
9.3.3	Fazit . . . . .	55
<b>10</b>	<b>Fazit</b>	<b>56</b>

Inhaltsverzeichnis	vi
10.1 Zusammenfassung . . . . .	56
10.2 Ausblick . . . . .	56
<b>A Inhalt der CD-ROM/DVD</b>	<b>57</b>
A.1 PDF-Dateien . . . . .	57
A.2 Bildmaterial . . . . .	57
A.3 Online-Ressourcen . . . . .	57
<b>Quellenverzeichnis</b>	<b>58</b>
Literatur . . . . .	58
Software . . . . .	59
Online-Quellen . . . . .	60

# Kurzfassung

In der vorliegenden Arbeit wurden Leitlinien aufgestellt, die bei der Planung und Umsetzung eines Audiokonzepts zu Virtual Reality Spielen als Orientierungshilfe genutzt werden können. Hierzu wird auf die Funktionsweise des binauralen Audios eingegangen, sowie ein Überblick über die Merkmale von Lautsprechern und Kopfhörern gegeben, die unterschiedliche Probleme und Anwendungsbereiche besitzen. Durch Theorien, Studien und eigene Beobachtungen wird ein Überblick über den Aufbau einer virtuellen auditiven Umgebung zu geben und welche Aspekte der Audioebene modifiziert werden können, um eine Verbesserung der Lokalisierung von auditiven Hinweisen zu erhalten. Aus den Beobachtungen und Erkenntnissen der vorliegenden Arbeit werden Leitlinien zur Verwendung von 3D Audio in Virtual Reality Spielen abgeleitet. Am Ende der Arbeit werden drei explorative, storybasierte Virtual Reality Spiele, darunter das eigene Masterprojekt namens *Letzte Worte* (Matthias Patscheider, Samantha Povolny und Bianca Zankl. HTC Vive, Oculus Rift. 2018) im Hinblick auf die Leitlinien untersucht.

# Abstract

The following thesis tries to define guidelines to help with the development and implementation of an audio concept for virtual reality games. Therefore, the functionality of binaural audio is mentioned, as well as an overview of the properties of loudspeakers and headphones which differ in problems and field of application. It is attempted to create an overview of the structure of a virtual auditory environment and which aspects of the audio level can be modified to improve the localization of auditory cues through theories, studies and my own observations. Based on observations and results of the following thesis, I defined guidelines concerning the usage of 3D Audio for virtual reality games. At the end of the thesis, three explorative, story-based virtual reality games, including my own master thesis project called *Letzte Worte* (Matthias Patscheider, Samantha Povolny and Bianca Zankl. HTC Vive, Oculus Rift. 2018) are then analyzed using these guidelines.

# Kapitel 1

## Einleitung

Audio ist vielseitig einsetzbar und kann audiovisuelle Projekte mit Emotionen bereichern oder sogar zur Orientierung genutzt werden. Vor allem mit dem Erscheinen von modernen Virtual Reality Technologien spielt Audio eine äußerst wichtige Rolle. Binaurales bzw. 3D Audio bieten vor allem für virtuelle Welten einen wesentlichen Grundstein zur Lokalisierung von Elementen und der eigenen Position innerhalb einer solchen Umgebung. 3D Audio ist auf kein Sichtfeld begrenzt, sondern bietet zu jeder Zeit einen 360° Eindruck der auditiven Soundkulisse und somit des virtuellen Ortes. Trotz all dieser Eigenschaften gibt es keine einheitlichen Theorien oder auch Richtlinien, die für solch neue Technologien wie Virtual Reality eingesetzt werden können.

In dieser Arbeit wurde auf genderneutrale Sprache verzichtet. Jede Formulierung impliziert weibliche wie männliche Formen. Bilder wurden zur besseren Lesbarkeit in Helligkeit oder Form bearbeitet.

### 1.1 Motivation

Durch die Planung und Erstellung des eigenen praktischen Masterprojekts, bei dem es sich um ein Virtual Reality Spiel handelt, wurde versucht die Problematik der Orientierung durch den Einsatz von Audio zu lösen. Erst bei der Konzeption wurde klar, dass trotz der Bedeutung von Audio für die Orientierung kaum Richtlinien oder Studien existieren, die sich mit der Rolle des Audios für virtuelle Welten beschäftigen. Vor diesem Hintergrund wollte ich mich in dieser Arbeit genauer mit dem Thema Orientierung und Audio im Zusammenhang mit Virtual Reality auseinandersetzen.

### 1.2 Fragestellung

Orientierung in virtuellen Welten stellt eine besondere Herausforderung dar, die unter anderem durch Audio gelöst werden kann. Eine große Rolle spielt die Orientierung vor allem in explorativen, storybasierten Virtual Reality Spielen, da sich diese aus mehreren Umgebungen zusammensetzen in denen sich der Spieler zurechtfinden soll. Die Geschichte des Spiels leitet ihn durch die unterschiedlichen Umgebungen und bietet ihm Anreize zur genaueren Untersuchung der einzelnen Bestandteile eines Ortes. Dadurch stellt sich die Frage, wie die Audioebene genutzt werden kann, um die Orientierung in solch kom-

plexen Welten zu steigern und dem Spieler die Navigation darin zu erleichtern. Durch die Analyse der Funktionsweise von binauralem Audio und der auditiven Ortung des menschlichen Gehörs, sowie der Miteinbeziehung von existierenden Studien soll versucht werden, diese Fragestellungen zu beantworten und Leitlinien bezüglich der Orientierung durch 3D Audio aufzustellen.

### 1.3 Struktur der Arbeit

Zu Beginn der Arbeit wird ein Überblick über die Funktionsweise von binauralem Audio und der räumlichen Wahrnehmung gegeben. Ebenso wird auf das Thema Virtual Reality und die Unterschiede zwischen Kopfhörer und Lautsprecher eingegangen. Des Weiteren werden zwei Studien zur Orientierung und ihre Erkenntnisse genauer beschrieben, die hilfreich zur Erklärung des Aufbaus einer virtuellen auditiven Umgebung und der Verwendung von auditiven Hinweisen sind. Durch die exemplarische Anwendung der Unreal Engine werden die Eigenschaften und Voraussetzungen an die Audiodateien, sowie die unterschiedlichen Audioeinstellungen zur Simulation von 3D Audio erklärt. Dank der zuvor gewonnenen Erkenntnisse war es möglich Leitlinien aufzustellen, die bei der Erstellung eines Audiokonzepts für Virtual Reality Spiele hilfreich sein können. Am Ende der Arbeit werden drei explorative, storybasierte Virtual Reality Spiele mit den vorliegenden Leitlinien untersucht, darunter auch das eigene Masterprojekt *Letzte Worte* [24].

## Kapitel 2

# Binaurales Audio

### 2.1 Definition Binaurales Audio

Das Wort „binaural“ selbst bedeutet „beide Ohren betreffend“ oder auch „mit beiden Ohren“ [27]. Bei einer binauralen Aufnahme wird versucht Schallsignale so einzufangen, dass der Zuhörer bei der Wiedergabe über Kopfhörer einen möglichst natürlichen Höreindruck erhält, über den er Annahmen zur Richtungslokalisierung treffen kann.

Um diesen Effekt des menschlichen Hörens nachahmen zu können, wurden spezielle Mikrofone bzw. Köpfe entwickelt, die den Aufbau und die Struktur des Kopfes nachempfinden. Besonders für Virtual Reality Anwendungen spielt die Besonderheit des binauralen Audios eine große Rolle.

### 2.2 Aufnahmeverfahren

Zur Aufnahme von binauralem Audio werden spezielle „Kunstköpfe“ eingesetzt, bei denen sich Mikrofone mit Kugel-Richtcharakteristik im künstlichen Gehörgang befinden [16, S. 70f.]. Binaurales Audio greift auf die Wahrnehmungsmechanismen des Menschen zurück und kann dadurch einen möglichst natürlichen Höreindruck liefern. Durch den Kopf werden bestimmte Schallsignale gedämpft und kommen so mit niedrigerer Intensität beim abgewandten Ohr an. Diese Wahrnehmungsmechanismen werden in Abschnitt 2.3 noch genauer untersucht. Des Weiteren wird versucht, dass der Kunstkopf die Maße eines durchschnittlichen, menschlichen Kopfes misst, um die Zeitdifferenz zwischen dem Eintreffen des Schallsignals in beiden Ohren nachzuempfinden. Ein Kunstkopfmodell der Firma Neumann wird in Abbildung 2.1 beispielhaft dargestellt. Normalerweise werden für Virtual Reality Anwendungen keine binauralen Aufnahmen, sondern Audiodateien, die mit Filtern innerhalb einer Game Engine<sup>1</sup> bearbeitet werden verwendet. Dadurch ist der Umgang mit 3D Audio in Virtual Reality Spielen für Spieleentwickler einfacher.

---

<sup>1</sup>Eine Game Engine ist ein Framework für Videospiele, durch das audiovisuellen Inhalte und Spielmechaniken erstellt werden können.



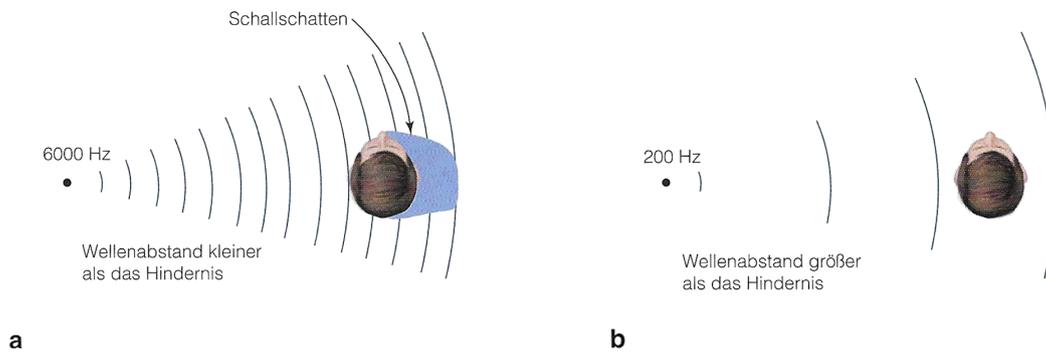
**Abbildung 2.1:** Kunstkopf KU 100 der Firma Neumann. Der Kunstkopf stellt ein binaurales Stereomikrofon dar und ist besonders für die Kopfhörerwiedergabe geeignet [32].

## 2.3 Auditive Ortung des menschlichen Ohres

Die Funktionsweise von binauralem Audio basiert auf den auditiven Wahrnehmungsmechanismen des Menschen. Sie spielen somit eine bedeutende Rolle, um das Prinzip hinter der Arbeitsweise von binauralem Audio besser verstehen zu können. Im Folgenden werden daher die Funktionen der auditiven Lokalisierung, sowie der räumlichen Wahrnehmung erläutert.

### 2.3.1 Auditive Lokalisierung

Goldstein beschreibt in [7, S. 286–289] den Mechanismus der auditiven Lokalisierung, der uns die Bestimmung der Position von Schallquellen ermöglicht. Zur Bestimmung dieser Position verwendet die auditive Lokalisierung vor allem binaurale Positionsreize, die durch das Einwirken von Kopf und Ohren auf eintreffende Schallwellen entstehen. Dabei versuchen binaurale Positionsreize die links-rechts Position einer Schallwelle (Azimuth) zu bestimmen. Hierfür werden die interaurale Zeitdifferenz und die interaurale Pegeldifferenz verwendet. Unter der interauralen Zeitdifferenz versteht man den „Zeitunterschied zwischen dem Eintreffen eines Schallsignals in beiden Ohren“ [7, S. 286]. Diese zeitliche Differenz entsteht dadurch, dass sich eine Schallquelle näher an einem der beiden Ohr befindet und dadurch früher dort ankommt. Umso weiter seitlich sie sich befindet, desto eindeutiger lässt sich die Position einer Schallquelle links oder rechts einordnen. Die interaurale Pegeldifferenz hingegen „basiert auf dem Unterschied der Schalldruckpegel des Schalls, der die beiden Ohren erreicht“ [7, S. 287]. Der Kopf selbst fungiert hier als eine Art „Dämpfer“ und stellt somit ein Hindernis für die Schallausbreitung dar. Es entsteht auf der abgewandten Seite ein „Schallschatten“. Durch ihn wird das Schallsignal reduziert und ein weniger intensives Schallsignal wird im abgewandten Ohr wahrgenommen. Dabei ist zu beachten, dass diese Reduktion des Schalldrucks nur bei hochfrequenten Signalen auftritt und nicht bei niederfrequenten [7, S. 287]. Wie in Abbildung 2.2 dargestellt, besitzen niederfrequente Schallwellen größere Wellenabstän-



**Abbildung 2.2:** (a) Hochfrequente Schallwellen, die beim Treffen mit dem Kopf einen Schallschatten erzeugen, durch den Schallsignale auf der abgewandten Seite abgedämpft werden. (b) Niederfrequente Schallwellen erzeugen durch ihre großen Wellenabstände keinen Schallschatten und werden daher auch nicht abgedämpft [7].

de und sind damit größer als der Durchmesser des Kopfes, wodurch sie nicht von ihm aufgehalten werden. Hochfrequente Schallwellen besitzen kleinere Wellenabstände und treffen dadurch auf den Kopf, der ihren Schalldruck abschwächt.

### 2.3.2 Prinzipien der Wahrnehmungsorganisation

Eine weitere Hilfestellung bei der auditiven Lokalisation sind laut Goldstein die Prinzipien der Wahrnehmungsorganisation, die Schallwellen nach bestimmten Merkmalen gruppieren [7, S. 301–304]. Hierbei werden auf den Herkunftsort, die Ähnlichkeit in Klangfarbe und Tonhöhe, die zeitliche Nähe, den guten Verlauf und die Erfahrung geachtet. Beim Herkunftsort werden Schallereignisse aus derselben Position oder einer sich leicht veränderten Position meist als eine Quelle angesehen. Bei dem Prinzip Ähnlichkeit in Klangfarbe und Tonhöhe werden Töne mit ähnlicher Klangfarbe oft der gleichen Quelle zugeordnet. Dieses Prinzip wird beispielsweise auch auf Stimmen angewendet. Auch hier werden Unterschiede in der Tonhöhe und Klangfarbe untersucht und wahrgenommen. Bei der zeitlichen Nähe werden Schallereignisse, die kurz nacheinander auftreten, ebenfalls einer einzelnen Schallquelle zugeordnet. Ist die Zeitdifferenz zwischen den Tönen zu lange, spricht dies meist für unterschiedliche Quellen. Der gute Verlauf achtet auf Schallsignale, deren Verlauf meist gleichbleibend oder nur langsam verändert wird. Hierbei spielt die Kontinuität eine wichtige Rolle. Auch die Erfahrung trägt zur Gruppierung einzelner Geräusche und Melodien bei, da das Gedächtnisschema angesprochen wird. Dieses wird normalerweise für Melodien verwendet. All diese Prinzipien bieten Informationen über die auditive Umwelt, allerdings müssen laut Goldstein meist mehrere Prinzipien kombiniert werden, um ausreichend Informationen zur Position und Anzahl bekommen zu können.

### 2.3.3 Räumliche Wahrnehmung

Um sich in einer Umgebung zurechtfinden und Annahmen über die Art des Ortes treffen zu können, greift der Spieler unter anderem auf den Hörsinn und das räumliche Bewusst-

sein zurück. Normalerweise liefern Umgebungen eine Vielzahl an auditiven Hinweisen auf den drei Achsen Azimuth (Winkel von links nach rechts), Elevation (Höhenwinkel) und Entfernung [7, S. 286]. Menschen benutzen normalerweise ihren Sehsinn, um Dinge, die sich vor ihnen befinden wahrzunehmen und zu erfassen. Für Hinweise aus der Umgebung, die über oder hinter ihnen sind, muss er sich auf den Hörsinn verlassen [16, S. 1]. Dieser ist, im Gegensatz zum Sehsinn, nicht auf ein bestimmtes Sichtfeld begrenzt, sondern hat Zugriff auf alle  $360^\circ$ .

Umgebungen beinhalten meist mehrere Schallquellen, wobei jede einzelne eigene Eigenschaften und Positionen besitzt. Durch das Vorhandensein mehrerer Schallquellen gruppiert unsere Wahrnehmung einige davon, was es schwierig macht einzelne Sounds darin zu orten. Diese Geräusche tragen meist zur Schaffung einer Atmosphäre bei, die wiederum Aufschluss auf die vorhandene Räumlichkeit der Umgebung geben [16, S. 1f.]. Ein typisches Beispiel wäre eine Stadt, bei der Verkehrslärm, Reifenquietschen, Autohupen, Stimmen und Schritte von Passanten, sowie öffentliche Verkehrsmittel und Einsatzfahrzeuge häufig vorkommen. Durch die unterschiedlichen Entfernungen zum Zuhörer wird eine gewisse Tiefe erzeugt, die in uns die Wahrnehmung „Außen“ auslöst [16, S. 2]. Durch Reflexionen innerhalb eines Raumes existieren mehrere Signale von einer einzigen Schallquelle, die zur Wahrnehmung „Innen“ beitragen und somit die räumliche Wahrnehmung der Umgebung mitformen. Je nach Räumlichkeit gelangt ein gewisser Teil auf direktem Weg von der Schallquelle zum Ohr, während der restliche Anteil zuvor von Wänden, Decke und Boden des Raums reflektiert wird und erst danach im Ohr eintrifft. Dabei unterscheidet man zwischen Direktschall, also der Schall auf direktem Weg zum Ohr, und indirektem Schall bzw. Raumschall, der zuvor durch verschiedene Oberflächen reflektiert wurde [7, S. 297]. Durch die unterschiedlichen Richtungen aus denen das Schallsignal eintrifft, kommt es zu Problemen bei der Lokalisation. Die Lokalisation von einzelnen Sounds kann dadurch in einer reflektiven Umgebung schwieriger sein, allerdings werden durch das Timing der Reflexionen zahlreiche Informationen über die Entfernung und Tiefe geliefert, die Rückschlüsse auf die Position der Schallquelle liefern [16, S. 6f.]. Frühe Reflexionen, die von der nächstgelegenen Oberfläche reflektiert werden, tragen zur Orientierung der Raumgröße bei. Später auftretende Reflexionen hingegen übertragen das Gefühl der Räumlichkeit. Auf die Lokalisation in reflektiven Umgebungen wird in Abschnitt 6.3 noch genauer eingegangen.

## Kapitel 3

# 3D Audio und Virtual Reality

Wie bereits zuvor erwähnt spielt binaurales bzw. 3D Audio eine große, nicht mehr wegzudenkende Rolle für Virtual Reality (VR) Anwendungen. Im Folgenden werden auf die Besonderheiten von VR, sowie auf die unterschiedlichen Anwendungen von Audio im Zusammenspiel mit Head-Mounted Displays (HMD) eingegangen.

### 3.1 Psychische Aspekte

Um eine möglichst realistische Umgebung abbilden zu können, müssen so viele Sinne des Nutzers wie nur möglich angesprochen werden. Die Wahrnehmungsprozesse des Menschen spielen daher eine essenzielle Rolle für die Planung und Ausführung einer virtuellen Umgebung und den darin enthaltenen Interaktionen. Unter einem VR-System definieren Dörner et al. Folgendes [4, S. 21]:

[Ein] Computersystem, welches die Aktionen von Nutzern erkennt, unter dieser Beeinflussung die Welt simuliert und über eine entsprechende Reizerzeugung die Nutzer eine Virtuelle Welt wahrnehmen lässt.

Das Ziel einer Mensch-Maschine-Interaktion besteht darin, dass der Nutzer in der virtuellen Welt genau gleich handelt, wie er es in der realen Welt auch tun würde. Ebenso soll die Interaktion mit dem Computer möglichst ausgeblendet werden und sich natürlich und normal anfühlen. Für dieses Akzeptieren der dargestellten Welt ist die sogenannte *Suspension of Disbelief* verantwortlich [4, S. 29]. Für eine kurze Zeit akzeptieren die Nutzer willentlich die dargestellte Welt mit all ihren darin befindlichen Interaktionen und Inhalten. Hierbei gerät das Verständnis der Realität in den Hintergrund, um die Welt und ihre Inhalte akzeptieren zu können. Zwei weitere wichtige Begriffe sind die Immersion und die Präsenz [3, S. 46]. Die Immersion bezeichnet den Grad des „Eintauchens“ in die virtuelle Welt und stellt damit die „technologische Qualität des Mediums“ [17, S. 175] dar. Sie kann durch Aspekte wie ein breiteres Sichtfeld, stereoskopische Bilder und ein verbessertes User-Tracking gesteigert werden. Hierfür helfen die vom Computer generierten Sinneseindrücke. Umso mehr Sinneseindrücke aus der virtuellen Welt stammen, umso schneller verdrängt der Nutzer die reale Welt und taucht tiefer in die simulierte Umgebung ein. Die Präsenz wiederum gibt im weiteren Sinne das „Gefühl

an, sich innerhalb einer virtuellen Umgebung zu befinden“ [4, S. 18] und die „psychologische [Wirkung] der individuellen User Experience“ [17, S. 175] an. Wird dieses Gefühl durch beispielsweise vom Nutzer ausgeführte Bewegungen nicht korrekt auf die virtuelle Welt übertragen, führt dies zu einem Präsenzbruch. Es gibt laut Salomoni et al. noch drei weitere Unterteilungen der Präsenz in räumliche Präsenz, Selbstpräsenz und soziale Präsenz [17, S. 175]. Die räumliche Präsenz steht für das Gefühl, sich physisch wirklich in einem Raum zu befinden und wird durch die Qualität der dreidimensionalen Bilder und der GUI beeinflusst. Bei der Selbstpräsenz wiederum handelt es sich um die Fähigkeit des Spielers, sich vollkommen mit dem Avatar zu identifizieren, was vor allem durch natürliche Interaktionen mit der Umwelt und deren Auswirkungen zusammenhängt. Die soziale Präsenz setzt auf das Gefühl andere Spielern, ob wahre Mitspieler oder NPCs wahrzunehmen. Dies hängt hauptsächlich davon ab, wie realistisch sich diese bewegen bzw. aussehen und wie natürlich man selbst mit diesen kommunizieren kann. Dörner und Steinicke sprechen in [3, S. 46] davon, dass Präsenz vor allem dann eintritt, wenn der Immersionsgrad hoch ist. Ebenso sei Präsenz etwas Individuelles, dass bei jedem User unterschiedlich sein kann. Die beiden Aspekte können noch weiter durch einen großen Sichtbereich in dem verwendeten Headset, aktives Head Tracking und reales Gehen unterstützt werden. Des Weiteren spielen die Plausibilitätsillusion und die Involviertheit ebenfalls mit der Immersion zusammen [4, S. 18f.]. Bei der Plausibilitätsillusion geht es darum, dass Ereignisse aus der virtuellen Welt so wahrgenommen werden, als ob sie wirklich geschehen würden. Diese Ereignisse werden laut Dörner et al. jedoch nicht von dem Nutzer selbst ausgelöst, sondern beruhen auf Inhalten der dargestellten Welt und betreffen ihn nur. Ein Beispiel hierfür wäre das Projektil einer Waffe, das an dem User vorbeifliegt. Bei dem Begriff der Involviertheit handelt es sich um den „Grad der Aufmerksamkeit bzw. des Interesses des Nutzers an der simulierten Welt“ [4, S. 19]. Dieses Interesse wird hauptsächlich durch Inhalte aus der virtuellen Welt ausgelöst. Es ist ebenso möglich, dass der Nutzer zwar an die Illusion des Ortes glaubt, sich aber trotzdem langweilt und dadurch weniger involviert ist.

### 3.2 Atmo und Ambience Sounds

Bei der Gestaltung einer auditiven Umgebung sollte auf die unterschiedlichen Typen von Sounds geachtet werden. Bei der Wahrnehmung einer Umgebung spielen vor allem *Ambience Sounds* und *Atmo* eine wichtige Rolle. Unter *Ambience Sounds* versteht man Hintergrundgeräusche, die zusammen zu einer Soundkulisse verschmelzen und den Charakter des Ortes definieren und dem Spieler ein gewisses Gefühl vermitteln. Beispiele für *Ambience Sounds* einer Schule wären somit Stimmen und Schritte von Schülern, sowie das Klingeln der Schulglocke. Allerdings kann die Art der *Ambience Sounds* die Stimmung und dadurch das Gefühl des Ortes verändern. Würden die Schüler alle leise miteinander flüstern, würde dies eine andere Stimmung vermitteln als wenn sie miteinander laut lachen und fröhlich diskutieren. Die *Atmo* wiederum bezeichnet die Kombination mehrerer *Ambience Sounds*, die erst zusammen dem Zuhörer eine Ahnung über die Art des Ortes geben. Darunter können sich Informationen zur Art des Raumes, der Umwelt und des sozialen Gefüges befinden [37]. Informationen über den Raum beinhalten meist Eindrücke über innen und außen, aber auch über die Art und Größe des Raumes, wie eine Kirche oder ein Zimmer. Die Umwelteindrücke können meist Geräusche wie Wind-

rauschen, Vögel oder technische Apparaturen beinhalten, die Aufschluss über die Art der näheren Umwelt liefern. Somit ist dem Spieler klar, dass er sich bei Windrauschen und Vogelzwitschern in der Natur befindet, während er sich beim Surren von Computern auch in einem Büro wiederfinden kann. Das soziale Gefüge verrät dem Spieler, ob es Menschen in der Umgebung gibt oder nicht. Befindet er sich in der Stadt, so werden meist Schritte und Stimmen von Menschen zu hören sein. Ist er jedoch alleine in einem Zimmer werden keine menschlichen Geräusche vorhanden sein. Die Atmo kann dem Zuhörer auch eine Art Tiefe vermitteln, bei dem einzelne Ambience Sounds unterschiedlich weit vom Spieler wahrgenommen werden. All diese Informationen über den Eindruck einer Umgebung können somit durch die Atmo vermittelt werden.

### 3.3 Technische Aspekte

Der wichtigste technische Punkt, der vor allem zur Immersion und Präsenz beiträgt, ist die Echtzeitfähigkeit von VR-Anwendungen. Während der Echtzeitverarbeitung sollen Inhalte aus der virtuellen Umgebung, sowie Bewegungen des Nutzers möglichst verzögerungsfrei dargestellt werden. Würde dies nicht geschehen und die vorhandene Latenz zu groß sein, würde dies den Spieler aus der Illusion reißen und sogar zu Unwohlsein und Übelkeit führen. S. Serafin und G. Serafin beschreiben in [18] mehrere Aspekte, die bei der Verwendung von Audio für naturgetreue Abbildungen einer Umgebung und für das Gefühl eines Ortes wichtig sind. Darunter fallen neben der Art der Soundübertragung, auch die Spielerbewegung, Interaktivität, Vielfalt, Übertreibung und Menge an Schallquellen. Durch die Verbesserung der Geräusche selbst, wie unter anderem einer Aufwertung von tiefen Frequenzen, kommt es zu einer Verbesserung der Präsenz. Würde man nur die Anzahl der Audiokanäle erhöhen, hätte dies keinen Einfluss auf die Präsenz selbst [18, S. 2]. Auch die Interaktivität des Nutzers mit der Umgebung spielt eine große Rolle. Durch diese interaktive Wechselwirkung entsteht erneut ein reales Gefühl des Ortes, da der User auch in Wirklichkeit mit Gegenständen in seiner Umgebung interagieren kann. Sind diese Interaktionen natürlich gestaltet, tragen auch sie zur Verbesserung der Präsenz bei.

### 3.4 Gestalterische Aspekte

Um eine Umgebung lebendiger zu gestalten, wird von den S. Serafin und G. Serafin empfohlen, bewegliche Soundquellen einzusetzen. Diese ermöglichen es, ein dynamisches und reales Bild einer Umgebung darzustellen. Werden in einer Umgebung sich wiederholende Geräusche verwendet, sollte darauf geachtet werden, dass auch diese die Wahrnehmung und Wirkung der virtuellen Welt stören können und als unnatürlich empfunden werden [18, S. 2]. Deswegen sollten Sounds nur dann wiederholt werden, wenn sie dem Nutzer nicht auffallen oder gar gänzlich vermieden werden. Ideal wären sich wiederholende Sounds wie zum Beispiel Ambience Sounds, die sich mit anderen Audiodateien abwechseln und erst dann erneut eine Wiederholung der bereits bekannten Sounds starten. Ebenfalls ist es bei vielen Programmen, wie der Unreal Engine von Epic Games [23] bereits möglich eine kleine Varianz in einen Loop<sup>1</sup> einzubauen. Dies geschieht meist

---

<sup>1</sup>Ein Loop ist eine unveränderte, sich stets wiederholende Sequenz.

durch die zufällige Veränderung von Frequenzen, Lautstärke und Geschwindigkeit der Schallquelle, wodurch ein Durchgang nie exakt wie der zuvor klingt. Ebenso können Sounds mit Signalwirkung gewählt werden, wodurch die Aufmerksamkeit des Nutzers in die gewünschte Richtung gelenkt wird. Auf die unterschiedlichen Klangobjekte und ihre Wirkungen wird in Abschnitt 7.1 genauer eingegangen. Da Audio für die Wahrnehmung und das Gefühl eines Ortes wichtig ist, sollte auch auf die Menge der verwendeten Sounds geachtet werden. Neben der Atmo sollten alle sichtbaren, beweglichen und interaktiven Objekte vertont werden. Objekte, die nicht sichtbar in der Umgebung vorhanden sind, sollten weggelassen werden, da sie den Nutzer eher verwirren könnten und dadurch zu einer verminderten Präsenz führen würden [18, S. 2]. Infolgedessen sollte immer auf die Menge der verwendeten Schallquellen geachtet werden, da viele Geräusche oft gar nicht gebraucht werden und eher schaden als helfen.

### 3.5 Head-Mounted Display

Als Ausgabegerät des VR-Systems dient ein Head-Mounted Display (HMD), das grob als „am Kopf befestigter Bildschirm“ übersetzt werden kann. Diese existieren in Form von Helmen oder Brillen. Grimm et al. beschreiben in [9, S. 147–153] die verschiedenen Arten von HMDs. Das Direktsicht-HMD ist die am häufigsten verwendete und günstigste Variante. Der Nutzer blickt mit Hilfe einer Vergrößerungsoptik auf das verbaute Display und ist mit dieser Variante komplett von der realen Außenwelt abgeschottet. Das Video-HMD bietet eine Frontkamera als Erweiterung zur Direktsicht-Variante. Hierbei werde Bilder der realen Welt durch die Frontkamera aufgenommen und an den angeschlossenen Computer geschickt. Durch diese Variante können auch Gegenstände in der realen Welt erfasst und umgewandelt werden, damit der User sie auch in der virtuellen Welt sieht. Bei der Variante des See-Through-HMDs sieht der Nutzer die reale Welt mit seinen eigenen Augen und virtuelle Elemente werden durch eine präzise Frontkamera auf dem Display augenblicklich verarbeitet und eingeblendet. Mit Hilfe der interaktiven HMD können Benutzer mit den virtuellen Informationen auf dem Display interagieren. Hierbei wird auf eine Eye-Tracking-Kamera gesetzt, die Steueranweisungen durch Augenbewegungen aufnimmt, interpretiert und verarbeitet an das HMD weiterleitet.

### 3.6 Besonderheit Head Tracking

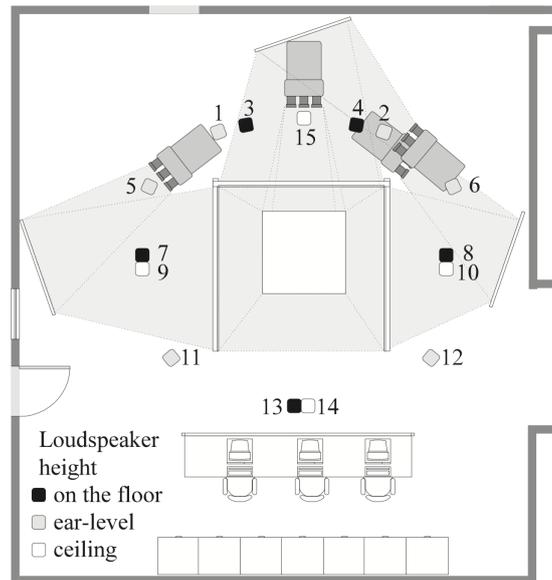
Binaurales Audio in virtuellen Welten funktioniert erst wirklich im Zusammenspiel mit Head Tracking. Beim Head Tracking werden die Kopfbewegungen des Nutzers überwacht und an das VR-System weitergegeben. Dieses berechnet die Bewegungen und überträgt diese an das Wiedergabegerät, damit die Kopfbewegungen in der realen und virtuellen Welt übereinstimmen und sich die Perspektive für den Benutzer korrekt verändert. Durch diese zusätzlichen Informationen der Kopfbewegung, können die binauralen Hinweise aus der virtuellen Umgebung verändert werden, um deren Veränderungen realistischer wirken zu lassen. Die Verwendung von Head Tracking löst dabei auch noch ein Problem von binauralem Sound: Die Identifizierung von vorne und hinten [16, S. 72f.]. Trägt eine Testperson Kopfhörer und werden ihm binaurale Aufnahmen vorgespielt, fällt es ihm schwer zu sagen, ob sich die Soundquelle hinter oder vor ihm befindet. Dieser

Nachteil hängt mit der Art der Aufnahme zusammen, denn durch die Platzierung der Mikrofone in den Gehörgängen, werden Signale von vorne und hinten durch den Kopf gedämpft. Erst durch eine minimale Versetzung der Soundquellen nach links oder rechts bzw. durch eine Kopfbewegung, ist es uns möglich zu sagen, ob das Geräusch von vorne oder hinten kommt. In virtuellen Welten wäre somit eine solche Unterscheidung ohne Head Tracking eher schwieriger bis unmöglich, da erst mit der Kopfbewegung die Audio Hinweise aus der Umgebung angepasst werden. Des Weiteren spielt hier die Erwartung hinzu, da wir es aus dem alltäglichen Leben gewohnt sind, diese Art der Wahrnehmungsreaktion auf eine Schallquelle zu erleben. Eine fehlerhafte Änderung des Sounds würde deswegen zu einem Präsenzbruch und einer verminderten Immersion führen. Die meisten HMDs verwenden daher Trackingverfahren, um die Kopfbewegungen der User zu überwachen.

### 3.7 Virtual Room

Bei einem Virtual (Reality) Room handelt es sich um ein voll immersives VR-System für mehrere Personen, in dem man sich frei bewegen und sogar mit anderen Nutzern kommunizieren kann [33]. Hierbei gibt es verschiedene Varianten des Aufbaus. Eine häufig verwendete Variante wäre zum Beispiel ein Raum, der mit großen Monitoren und einem Lautsprechersystem ausgestattet ist. Während der Bewegung im Raum, werden die User getrackt und die Perspektive auf den Monitoren ändert sich entsprechend. Meist wird jedoch nur die Perspektive einer Person getrackt, die für die anderen entsprechend angepasst wird. Eine weitere Aufbaumöglichkeit wäre ein Raum, in dem mehrere Personen mit einem VR-Headset wie der HTC Vive von HTC und Valve Corporation [31] ausgestattet werden. Dabei werden entweder die bereits integrierten On-Ear Kopfhörer des Headsets oder externe Kopfhörer, die von den Mitarbeitern zur Verfügung gestellt werden verwendet. Um in solch einer Umgebung nicht über Kabel zu stolpern, wird bei dieser Art meist ein tragbarer Computer verwendet.

Ein Beispiel für die erste Aufbauvariante, also mit Monitor- und Lautsprechersystem, wäre der Virtual Room der *Helsinki University of Technology*. Dessen Aufbau und Funktionsweise werden von Hiipakka et al. in [10] erläutert, während der Aufbau schematisch in Abbildung 3.1 dargestellt wird. Die Universität verwendet große Leinwände, deren Bilder von hinten auf die Leinwand projiziert werden. Des Weiteren werden Tracking-Geräte verwendet, um die Kopfbewegungen und Orientierung eines Nutzers zu verfolgen, wodurch die Perspektive auf den Leinwänden entsprechend geändert wird. Befinden sich mehr als eine Person in dem Raum, wird trotzdem nur ein Proband getrackt und die Ansicht für alle Nutzer entsprechend verändert. Dadurch sollen zwar leichte Deformationen für die anderen User entstehen, allerdings ist es ihnen so auch möglich miteinander zu kommunizieren. Für das Audiosystem werden Lautsprecher hinter den Leinwänden verwendet, die auf eine Mehrkanal-Reproduktion zurückgreifen da laut Hiipakka et al. binaurales Audio nur einen sehr kleinen Sweet Spot besitzt. Durch die Mehrkanal-Reproduktionstechnik ist es möglich, Audiosignale in einer natürlichen Art und Weise zu verwenden. Um das beste Audio-Erlebnis bieten zu können, sollte der Raum von sich aus möglichst ruhig sein und mit absorbierendem Material ausgestattet werden, um Hall und andere verändernde Eigenschaften auszumerzen. Der Virtual Room der *Helsinki University of Technology* wird später in Kapitel 5 aufgrund der dort



**Abbildung 3.1:** Aufbau des Virtual Rooms der *Helsinki University of Technology*, der vor allem mit Projektionen auf Leinwänden und einem Lautsprechersystem arbeitet [10].



**Abbildung 3.2:** Exit VR in Wien verwendet eine HTC Vive, Kopfhörer, sowie die HTC Vive Controller und einen tragbaren Computerrucksack von Zotac zum Spielen in einer offenen Umgebung mit  $25\text{m}^2$  [36].

durchgeführten Experimenten noch einmal erwähnt.

Ein Virtual Room kann jedoch auch für spielerische Zwecke verwendet werden. Die deutsche Firma EXIT Adventures [36] bietet eine VR-Version eines Escape the Room Spiels, bei der HTC Vive Headsets verwendet werden. Einen Standort gibt es auch in Wien, auf den im Folgenden genauer eingegangen wird. Die Firma bietet ein Spiel namens *HUXLEY* an, bei dem zwei bis sechs Spieler teilnehmen können und mit einem „speziellen Computerrucksack, einer VR-Brille und zwei Controllern ausgestattet werden, um auf einer  $4\text{x}6\text{m}$  großen Fläche zusammen eine Mission zu erfüllen“ [36]. Man kann sich frei bewegen und auch Bewegungen wie Drehen, Bücken und Springen problemlos ausführen. Bei dem Rucksack handelt es sich um den Zotac VR GO Backpack der

Firma Zotac [39], der einen tragbaren Computer für die Verwendung mit VR-Headsets darstellt. Durch seine Verwendung besitzt jeder Spieler einen eigenen Computer, der auf seine Bewegungen im Raum trackt. Der Raum selbst ist leer und bietet im Grunde nur über die VR-Headsets die Spielerfahrung und Umgebung. Durch Mikrofone können die Spieler während des etwa 45-minütigen Spiels miteinander kommunizieren. Durch Over-Ear Kopfhörer wird die Geräuschkulisse der virtuellen Welt passend zu den Kopf- und Körperbewegungen wiedergegeben. Ebenso ist es durch die Mikrofone und Kopfhörer möglich, bei etwaigen Problemen das Spielleiter-Team zu kontaktieren. Zur Interaktion mit der virtuellen Umwelt werden die Controller der HTC Vive verwendet. Ein solcher Aufbau kann exemplarisch in Abbildung 3.2 beobachtet werden.

## Kapitel 4

# Lautsprecher und Kopfhörer

Bei der Verwendung von binauralem Audio, vor allem für virtuelle Welten, ist besonders die Wahl des Wiedergabemediums wichtig. 3D Audio-Reproduktionen erfordern eine möglichst genaue Kontrolle über das Audiosignal und wie es beim Spieler ankommt. Dies ist vor allem durch den Aufbau von Kopfhörern, aber auch mit Hilfe eines Lautsprechersystems möglich. Im folgenden Kapitel werden die Besonderheiten von Lautsprechern und Kopfhörern im Zusammenhang mit binauralem Audio behandelt.

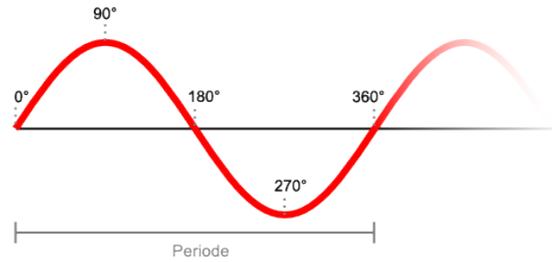
### 4.1 Lautsprecher

Während Lautsprechersignale unverändert auch einigermaßen über Kopfhörer funktionieren, ist es umgekehrt nicht so einfach möglich. Die binaurale Audiowiedergabe über Lautsprecher birgt einige Schwierigkeiten, mit denen sich unter anderem Durand Begault in [1] auseinandersetzt.

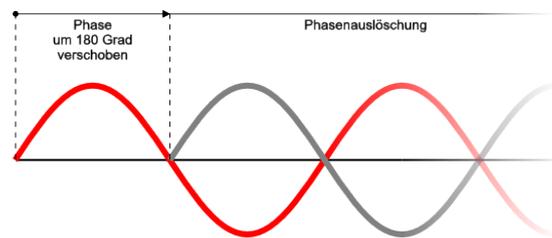
Dabei gibt es laut Begault vor allem drei Probleme, die im Gegensatz zu Kopfhörern bei Lautsprechern auftreten [1, S. 174–177]. Da Lautsprecher das Signal nicht direkt an das Ohr des Zuhörers abliefern können, fließen besonders die Informationen der Umgebung in die Wahrnehmung des Audios mit ein. Die im Raum vorhandenen Reflexionen führen zu einer veränderten Wahrnehmung. Selbst wenn, wie auch in Abschnitt 3.7 kurz erwähnt, absorbierende Materialien zur Schallisolierung des Raumes verwendet werden oder auch zeitgleich Modifikationen am Schallsignal selbst stattfinden, ist es nie möglich diese Veränderungen durch den Raum vollständig zu entfernen.

Ein weiteres Problem beinhaltet den Sweet Spot und die Position von einem oder mehreren Nutzern. Insbesondere die Abbildung einer räumlichen 3D Umgebung über Lautsprecher für mehrere Zuhörer stellt eine schwierige Herausforderung dar. In den meisten Fällen wird nur eine der Personen im Bereich getrackt und aufgrund dieser Position werden Berechnungen für das akustische Raumbild angestellt. Dadurch ergibt sich nur ein kleiner Sweet Spot, in dem das Hörerlebnis ideal funktioniert. Für jegliche Nutzer außerhalb dieses Sweet Spots wird eine immer ungenauer werdende Geräuschkulisse geboten.

Als letztes großes Problem erwähnt Begault noch, dass jedes Lautsprechersignal auch von beiden Ohren gehört wird, wodurch es zum sogenannten Cross-Talk Effekt kommt, der die interaurale Zeitdifferenz, welche den Zeitunterschied vom Eintreffen des



**Abbildung 4.1:** Die Positionen eines Schwingungsdurchganges von  $0^\circ$  bis  $360^\circ$ , wobei zur besseren Verständlichkeit der Grafik Sinustöne in ihrer Reinform verwendet wurden. Ein Durchgang kann auch als Periode bezeichnet werden [26].



**Abbildung 4.2:** Überlagerung zweier Wellen um  $180^\circ$ , wodurch es zu einer Phasenauslöschung kommt, die in Stille resultiert [26].

Schallsignale in beiden Ohren untersucht, und das Gleichgewicht der Frequenzen beeinflusst. Somit werden nun auch unerwünschte Signale von anderen Schallquellen gehört, welche die Orientierung und Lokalisierung der gewünschten Schallquellen schwierig bis unmöglich macht. Zur Kompensation wird hierfür die sogenannte Cross-Talk Cancellation eingesetzt [1, S. 175]. Diese Methode wurde von Atal und Schroeder 1963 entwickelt um die Wiedergabe von binauralen Aufnahmen über Lautsprecher zu ermöglichen [1, S. 175]. Bei der Cross-Talk Cancellation wird etwas vom linken Signal dem rechten Signal beigefügt und mit der Verzögerung versehen, die es bis zum Eintreffen am anderen Ohr brauchen würde. Das gleiche Prinzip wird auch für das linke Signal angewandt. Dadurch kommt es zu einer Phasenverschiebung – welche durch ihre Verschiebung auch als „aus der Phase“ bezeichnet wird –, bei der das Signal um  $180^\circ$  verschoben und das unerwünschte Signal vollständig ausgelöscht wird [26]. Dieser Vorgang wird in Abbildung 4.2 dargestellt. Bei den  $180^\circ$  handelt es sich um eine Stelle innerhalb des Schwingungsdurchganges von  $0^\circ$  bis  $360^\circ$ , der in Abbildung 4.1 illustriert wird.

Nichtsdestotrotz hängt diese Filterung stark von der Position des Kopfes ab. Durch diese Methode ist die binaurale Wahrnehmung zwar im Sweet Spot möglich, je weiter man sich jedoch von diesem entfernt, desto mehr unerwünschte Signale treten auf und die Filterung wird unwirksam. Um eine effiziente Cross-Talk Cancellation verwenden zu können, braucht es genaue Details zur Hörumgebung, sowie zu den Positionen der Lautsprecher und des Zuhörers. Su und Hsieh haben in [19] herausgefunden, dass durch die Verwendung von mehr als zwei Lautsprechern die Cross-Talk Cancellation weniger empfindlich gegenüber Abweichungen vom Sweet Spot ist und auch das Head Tracking besser eingesetzt werden kann, da mehr Ansprechpunkte in der Hörumgebung

des Nutzers existieren. In ihren Tests haben Su und Hsieh herausgefunden, dass drei bis vier Lautsprecher in Zusammenarbeit mit zwei Cross-Talk Cancellation Filtern bereits robuster gegenüber Abweichungen und den dadurch entstehenden Cross-Talk Signalen sind.

All die genannten Faktoren erschweren zwar die Verwendung von Lautsprechern im Zusammenhang mit binauralem Audio, allerdings ist die Verwendung durch die erwähnten Vorgehensweisen und Methoden unter gewissen Umständen auch ebenso möglich. Laut Begault wäre der Idealzustand für die Verwendung von Lautsprechern die Existenz eines leichten, kabellosen Head Tracking Gerätes, das die Berechnungen zu der relativen Kopfposition direkt an die Lautsprechergruppe sendet und eine Echtzeitverarbeitung des Audiosignals, sowie der Cross-Talk Cancellation ermöglicht [1, S. 177].

## 4.2 Kopfhörer

Durch die bereits in Kapitel 2 erklärte Funktionsweise von binauralem Audio sind besonders Kopfhörer für die Wiedergabe geeignet, da sie nahe am Aufnahmepunkt der Mikrophone sitzen und zusätzliche Veränderungen des Audios, durch beispielsweise Hall oder andere Gegebenheiten eines Raumes abgeschirmt werden. Hierbei spielt unter anderem die Head-Related Transfer Function (HRTF) eine große Rolle, da diese die „[...] komplexe Filterwirkung von Kopf, Außenohr (Pinnae) und Rumpf [bezeichnet]“ [38]. Sie ermöglicht es uns unter anderem auch akustische Signale zu orten. Hierbei werden meist standardisierte Filter entwickelt, die auf Durchschnittsgrößen von menschlichen Köpfen basieren. Individuelle Filter sind meist sehr zeitaufwendig und teuer, was sie vor allem für die Massenanwendung unbrauchbar macht.

Das Problem des Cross-Talk bei Lautsprechern ergibt sich bei Kopfhörern nicht, da die HRTF diese Signale filtert und jedes Ohr auch nur ein einziges, für sie bestimmtes Signal erhält. Das linke und rechte Signal unterscheiden sich durch die interaurale Zeitdifferenz, die interaurale Pegeldifferenz, sowie der Filterwirkung des Kopfes durch die zuvor berechnete HRTF. Ein weiterer Vorteil von Kopfhörern ist, dass der Abstand zwischen den Kopfhörern und dem Gehörgang des Users relativ gleich bleibt. Dadurch kann man davon ausgehen, dass das Verhältnis von Nutzer und Kopfhörer beim Tracken der Bewegungen relativ gleich bleibt. Dies führt zu einem verminderten Rechenaufwand im Gegensatz zu Lautsprechersystemen.

Des Weiteren spielt auch der Kostenfaktor bei der Wahl des Wiedergabemediums eine Rolle. Es ist sehr unwahrscheinlich, dass jemand zwei Versionen für eine VR-Anwendung erstellt, auch wenn die Kosten für die Reproduktion von Lautsprechersystemen über die Zeit vernünftigeren Ausmaße angenommen haben. Da meist Privatpersonen Ziel einer VR-Experience oder eines VR-Spiels sind, wird von fast allen Publishern der Kopfhörer gewählt und das Audio für dieses Wiedergabemedium optimiert. Kopfhörer werden bevorzugt von Privatpersonen verwendet, da man für ihre Verwendung kaum Platz benötigt und viele Modelle im Vergleich zu guten Lautsprechersystemen oft kostengünstiger sind. Ein Problem bei der Reproduktion von raum- und kopfbezogenen Audio ist die Im-Kopf-Lokalisation (IKL), welche Thomas Görne folgendermaßen beschreibt [8, S. 131]:

Weil bei der Kopfhörerwiedergabe der Einfluss des Außenohrs auf das Schallfeld fehlt, werden die Phantomschallquellen im Kopf auf einer Linie zwischen den Ohren geortet. Die IKL ist bei langem Hören ermüdend und führt zu dem typischen „Druck auf den Ohren“.

Dieses Problem sei laut Görne nur mit der Filterung durch eine HRTF lösbar. Diese nehme den Druck von den Ohren. Würde man diese gefilterte Version nun über Lautsprecher abspielen, würde „die HRTF als hochfrequente Klangfärbung wahrgenommen [werden]“ [8, S. 131]. Yuan et al. definieren in [20, S. 420] neben der HRTF noch weitere Faktoren, die benötigt werden um das natürliche Hören imitieren zu können. Darunter befinden sich frühe Reflexionen, später Nachhall und ebenso dynamische Faktoren, wie Kopfbewegung des Zuhörers, sowie Informationen zum Raum und zur Schallquelle. Durch Miteinbeziehung all dieser Faktoren kann ein akkurates Bild der virtuellen Umgebung erschaffen werden, dass dem natürlichen Hören des Menschen ähnelt. Im Idealfall fällt einer Person somit das Tragen des Kopfhörers nach einiger Zeit gar nicht mehr auf und er nimmt das Gehörte als normal hin.

## Kapitel 5

# Ausgewählte Theorien zur Orientierung

### 5.1 Sound Charakteristiken nach McMullen

In [13] beschreibt Kyla McMullen die vielen Vorteile, die Audio für virtuelle Welten birgt, sowie Überlegungen, die zuerst angestellt werden müssen, bevor man sich der Erstellung von räumlichen Audio für virtuelle Welten widmet. Dabei befasst sie sich mit folgenden klanglichen Charakteristiken [13, S. 31–32]: Sprachgebundene und nicht-sprachgebundene Sounds, unregelmäßige und kontinuierliche Sounds, sowie Lautsprecher und Kopfhörer.

Nicht-sprachgebundene Sounds sind leichter zu lokalisieren als sprachgebundene Sounds, und werden meist als *Beacons* und *Auditory Icons* bezeichnet [13, S. 32]. *Beacons* sind normalerweise Geräusche, die akustisch ähnlich zu alltäglichen Lauten sind, wie beispielsweise „Boings“ oder das Quietschen von Türen. Diese Sounds werden vorzugsweise für Alarm- oder Warnsysteme verwendet. Hierbei ist jedoch zu beachten, dass die einzelnen Bedeutungen der Geräusche erst gelernt werden müssen, um sie verstehen zu können. *Auditory Icons* sind „natürliche“ Geräusche, deren Bedeutung bereits aus der realen Welt bekannt ist, wie etwa das Rauschen von Wasser oder das Knistern von Flammen.

Unregelmäßige Sounds lassen sich während einer Bewegung schlechter lokalisieren, da der Spieler sie nur dann wahrnehmen kann, wenn sie zu hören sind [13, S. 32]. Ein Beispiel hierfür wären Schritte, da diese nur dann zu hören sind, wenn sich eine Spielfigur fortbewegt. In den Pausen, in denen die Spielfigur stehen bleibt sind die Schritte nicht zu hören und der Spieler kann die Position der Figur aufgrund der Schritte nicht mehr bestimmen. Kontinuierliche Sounds, wie das Rauschen eines Wasserfalls, können hingegen durch ihr stetiges Vorhandensein selbst bei der Fortbewegung des Spielers lokalisiert werden. Ebenso kann durch kontinuierliche Sounds immer der Abstand zum gehörten Objekt beobachtet und eingeschätzt werden.

Laut McMullen wird dem Kopfhörer im Gegensatz zum Lautsprecher vor allem ein großer Vorteil zugesprochen: Die relative Position der Kopfhörer zum Zuhörer bleibt immer gleich [13, S. 32]. Bei Lautsprechern muss zuerst immer eine Position gefunden werden, damit der räumliche Effekt des Audios mit der in der virtuellen Welt übereinstimmt. Bewegt der Spieler seinen Kopf mit den Kopfhörern, braucht das System allerdings eine sehr gute Echtzeit-Signalverarbeitung, da die auditive Umgebung mit

der Bewegung des Nutzers übereinstimmen muss, um die Illusion dieser Welt aufrecht-erhalten zu können. Bei Lautsprechern ist ein derartiger Rechenaufwand nicht nötig, da die Signalquellen aus den Lautsprechern mit denen aus der virtuellen Welt übereinstimmen und sich durch die Bewegung einer Person im Lautsprecherfeld die Wahrnehmung der Schallereignisse von selbst ändern [13, S. 32].

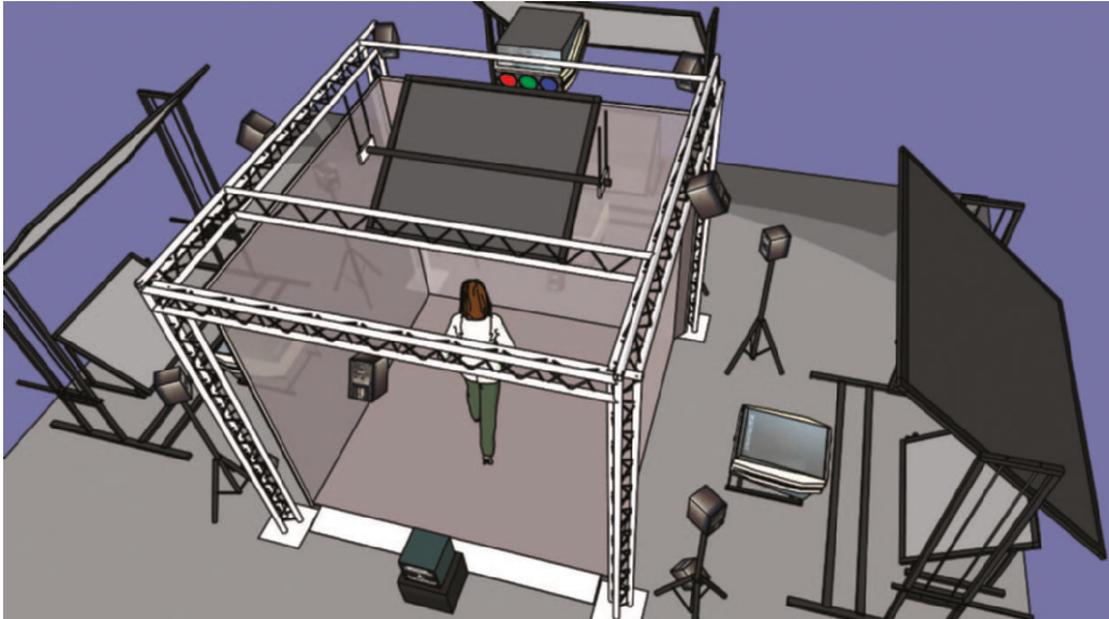
Des Weiteren berücksichtigt sie die Performance von Systemen, da schon die kleinste Verzögerung den Nutzern beeinflussen und ihn aus der Welt reißen könnte, da Audio und Bild nicht mehr zusammenpassen. Ebenso ist hier die Gefahr von „Motion Sickness“ zu beachten, dessen Symptome von leichten Kopfschmerzen bis Übelkeit reichen [13, S. 32]. Ein zweiter wichtiger Punkt bei der Verwendung von 3D Audio ist das Hinzufügen von visuellen Landmarks bzw. Orientierungspunkten, da diese bei der Erinnerung an bestimmte Orte behilflich sein können [13, S. 32]. Kurz wird ebenfalls darauf eingegangen, dass sich Nutzer in einer dichten auditiven Umgebung leichter erinnern und ihre Suchfähigkeit schneller verbessern konnten, als Nutzer in einer größeren Umgebung, in der Soundquellen weiter verteilt wurden [13, S. 32].

## 5.2 Navigation durch auditive Hinweise nach Lokki und Gröhn

Tapio Lokki und Matti Gröhn beschäftigen sich in [12] mit der Frage, wie gut die Navigation in einer virtuellen Umgebung nur anhand von auditiven Hinweisen funktioniert. Hierbei wurde von den beiden Autoren eine Studie im Virtual Room der *Helsinki University of Technology* durchgeführt, dessen Aufbau in Abbildung 5.1 schematisch dargestellt wird. Gearbeitet haben die beiden mit einem Audio-Reproduktionssystem für Lautsprecher, anstelle von Kopfhörern, und mit sogenannten *Auditory Beacons* [12, S. 80]. Diese *Auditory Beacons* ermöglichen eine Navigation durch Audiohinweisen zu Orten, die nicht sichtbar sind, da sie von anderen Objekten, wie z.B. Gebäuden, verdeckt werden. Ebenso können sie einen interessanten Teil eines 3D-Objektes hervorheben. Für die Studie wurden zwei Tests durchgeführt: Im ersten Test wurden visuelle, auditive und audiovisuelle Hinweise verglichen. Beim zweiten Test wurden verschiedene auditive Hinweise ohne jegliche visuelle Information getestet [12, S. 80].

Die Aufgabe für die Testpersonen bestand darin, sich auf einem vordefinierten Weg durch Tore zu navigieren. Die Tore wurden durch einen weißen Ball dargestellt und ebenso wurde eine Punktschallquelle zur akustischen Wahrnehmung der Tore verwendet. Die weißen Bälle befanden sich in einem komplexen 3D Gerüst. Der Proband konnte sich mithilfe eines selbsterstellten Stabes, der aus einer Funkmaus und einem magnetischen 3D Tracker bestand, fortbewegen [12, S. 80]. Die Testperson kann somit die Geschwindigkeit und Bewegungsrichtung mit dem Stab lenken. Das Tor galt als erreicht, sobald man sich in einem bestimmten Radius um den weißen Ball befand. Als auditiven Hinweis gab es eine Art Glockenläuten, um das Erreichen des Tores zu signalisieren. Vor den eigentlichen Tests wurde so die Navigation erklärt, damit die Testpersonen später keine Probleme mehr mit der Benutzung des Stabes haben und sich auf ihre Aufgabe konzentrieren können.

Im ersten Test wurden audiovisuelle, visuelle oder auditive Hinweise in der virtuellen Umgebung verwendet. Bei den visuellen Hinweisen handelte es sich, wie bereits oben erwähnt, um einen weißen Ball, während die klanglichen Hinweise aus Pink Noise mit einem weitem Spektrum und einer zeitlich begrenzten Varianz bestanden. Der Test



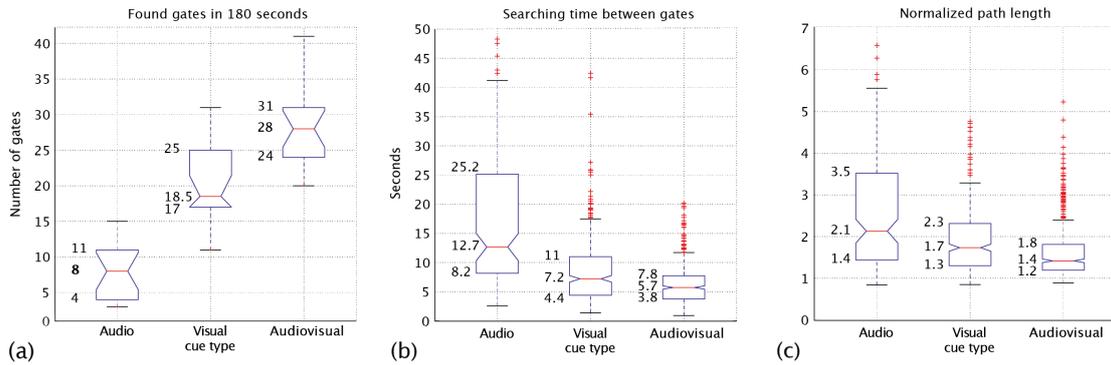
**Abbildung 5.1:** Eine schematische Darstellung des virtuellen Raumes an der *Helsinki University of Technology*, die von Lokki und Gröhn als Umgebung für ihre Navigationsstudie ausgewählt und verwendet wurde [12].

dauert 3 Minuten. Die Testergebnisse zeigen, dass die Navigation mit audiovisuellen Hinweisen bei weitem einfacher ist, als mit rein visuellen oder auditiven Hinweisen [12, S. 82], wie die Ergebnisse in Abbildung 5.2 zeigen. Es kann somit gesagt werden, dass die grobe Ortung zwar aufgrund der auditiven Hinweise stattfindet, jedoch die finale Annäherung an das Ziel nur durch die Bestätigung eines visuellen Reizes vorgenommen wird.

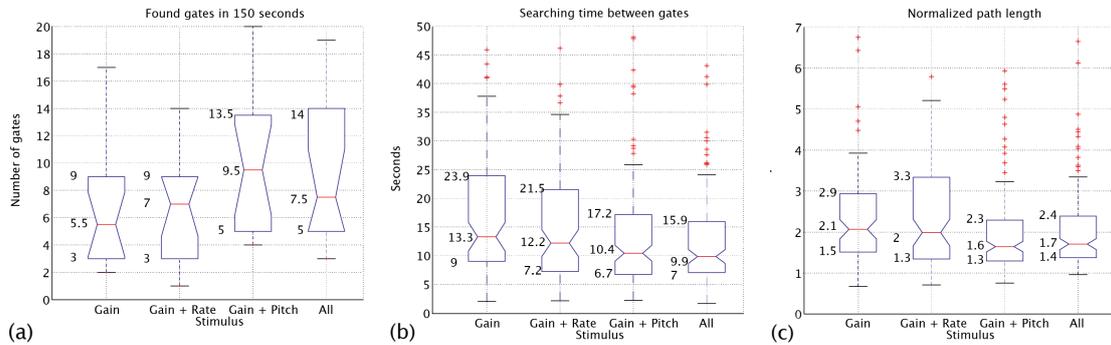
Durch die positiven Resultate der reinen klanglichen Hinweise, wurde ein zweiter Test durchgeführt, der verschiedene Parameter der Audioebene beachtet und diese verändert. Die beiden führten das Hauptproblem der Lokalisation von klanglichen Objekten auf die Ungenauigkeit der Elevation-Wahrnehmung zurück, da sich viele Testpersonen meist über oder unter der Soundquelle befanden [12, S. 83]. Für die auditive Lokalisation spielen die drei Achsen Azimut (Achse von links nach rechts) und Elevation (Achse von oben nach unten), sowie die Entfernung (Distanz einer Schallquelle vom Zuhörer) eine wichtige Rolle [7, S. 286]. Der Test dauert dieses Mal nur 2.5 Minuten und es werden vier verschiedene Stimuli, erneut aus Pink Noise bestehend, getestet: Lautstärke, Lautstärke und Pulsrate, Lautstärke und Tonhöhe, sowie Lautstärke, Pulsrate und Tonhöhe.

Der erste Stimuli (Lautstärke) ist gleich mit dem auditiven Hinweis aus dem ersten Test und besteht somit einfach nur aus Pink Noise mit Entfernungsabfall. Der zweite Stimuli (Lautstärke und Pulsrate) zeigt ebenfalls die Entfernung zum nächsten Tor mithilfe einer Art Pulsrate an – je näher ein Tor ist, desto schneller bzw. dichter wurde die Pulsrate. Der dritte Stimuli (Lautstärke und Tonhöhe) enthält ein schmales Bandpass<sup>1</sup>, durch welches die Höhe ungefähr eingeschätzt werden kann, und Abstufungen

<sup>1</sup>Ein schmales Bandpass ist ein Filter, der nur bestimmte Frequenzbereiche hindurch lässt.



**Abbildung 5.2:** Die Ergebnisse des ersten Tests. In (a) kann man die Anzahl der gefundenen Tore erkennen. In (b) ist die Zeit angegeben, die für das Finden von einem Tor zum anderen gebraucht wurde. In (c) werden die normierten Pfadlängen von einem Tor zum nächsten aufgezeigt [12].



**Abbildung 5.3:** Die Ergebnisse des zweiten Tests. In (a) kann man die Anzahl der gefundenen Tore erkennen. In (b) ist die Zeit angegeben, die für das Finden von einem Tor zum anderen gebraucht wurde. In (c) werden die normierten Pfadlängen von einem Tor zum nächsten aufgezeigt [12].

der Tonhöhe, durch welche die Höhe der Schallquelle exakter angegeben werden kann. Der vierte Stimuli (Lautstärke, Pulsrate und Tonhöhe) vereint alle zuvor genannten Parameter [12, S. 84].

Die Ergebnisse aus dem zweiten Test zeigen, dass der erste Stimuli die gleichen Ergebnisse, wie die auditiven Hinweise aus dem ersten Test liefert, wie in Abbildung 5.3 gezeigt wird. Die zuvor getroffene Annahme der Autoren, dass die ungenaue Information auf der Elevationsebene Schuld an der schlechten Orientierung sei, wurde durch den Test bestätigt. Durch die Verwendung von zusätzlichen Informationen auf der Elevationsebene wurden Tore schneller gefunden und die Reiselänge zwischen zwei Toren konnte reduziert werden [12, S. 85]. Im Grunde kann gesagt werden, dass rein auditive Hinweise gleich gut wie rein visuelle Hinweise funktionieren können, wenn sie sorgsam designt werden, jedoch am allerbesten die Kombination aus beiden Reizen für die Navigation in virtuellen Umgebungen geeignet ist.

### 5.3 Fazit der Theorien

Beide Texte heben gewisse Parameter hervor, die verändert oder einfach beachtet werden sollten um die Orientierung in einer virtuellen Welt zu verbessern.

Lokki und Gröhn haben durch ihre Studien und der Manipulation von Lautstärke, Tonhöhe und Pulsrate geschafft, die Orientierung durch reine auditive Hinweise zu optimieren. Ihre Erkenntnis war, dass die Elevationsebene – also die Achse, die für die Höheninformation zuständig ist – von großer Bedeutung für die Orientierung in komplexen und abstrakten Umgebungen ist. Werden genug Informationen über sie gesendet, ist es dem Spieler möglich sich besser in der Welt zurechtzufinden.

McMullen hebt die Unterschiede verschiedener Bereiche hervor, die bei der Arbeit mit Audio in Virtual Reality beachtet werden sollten. Hierbei spielen für sie auch die Art der Sounds eine wichtige Rolle, die zur verbesserten Orientierung beitragen sollen. Durch kontinuierliche Sounds und auditive Landmarks ist der Spieler in der Lage sich eine gedankliche Karte der Umgebung abzuspeichern. Auditive Landmarks bieten hierbei auch den Vorteil, dass sie nicht gesehen werden müssen um registriert zu werden, sondern auch von weiter weg gehört und gefunden werden können. Ebenfalls empfehlenswert ist hierbei der Einsatz von mehreren kleineren Landmarks, die in der Umgebung verteilt sind. Je dichter eine Umgebung an Audio-Hinweisen ist, desto besser und schneller erinnert man sich später auch an diese [13, S. 32]. Des Weiteren sollte auf die Bedeutung der Sounds geachtet werden, da die Bedeutung von *Beacons* meist erst im Laufe des Spiels definiert werden, während *Auditory Icons* bereits aus der realen Welt bekannt sind und ihre Bedeutung dadurch klar ist. Hierbei sollte jedoch auf die kulturelle Bedeutung dieser Sounds geachtet werden.

Als weiteren Eckpunkt könnte man den Vergleich von Kopfhörern und Lautsprechern betrachten, der in Kapitel 4 noch genauer betrachtet wird. Hierbei spielen auch die Echtzeitverarbeitung der Kopfhörer und die passende Aufstellung der Lautsprecher eine wichtige Rolle. Bei Kopfhörern wird dank der Echtzeitverarbeitung eine gute Hardware gebraucht, während bei den Lautsprechern im Vorhinein eine gute Position für diese gesucht werden muss um den virtuellen Raum ideal abbilden zu können.

## Kapitel 6

# Orientierung in virtuellen auditiven Umgebungen

Beim Aufbau einer virtuellen auditiven Umgebung (VAE) ist neben dem Einsatzgebiet und dem Genre auch auf andere Aspekte zu achten. Im folgenden Kapitel werden der Aufbau einer VAE, Methoden zur Wegfindung, die auditive Lokalisierung in VAE, sowie das Erinnerungsvermögen und Interaktion im Bezug auf explorative, storybasierte VR-Spiele behandelt. Die gefundenen Erkenntnisse sollen bei der Konzeption einer VAE eine Hilfestellung sein.

### 6.1 Aufbau einer VAE

Der Aufbau einer explorativen, storybasierten VR-Umgebung kann sich aus nur einer oder auch mehrerer Räumlichkeiten zusammensetzen und bietet dem Spieler eine offene Umgebung, die er neben der Führung durch die Geschichte erkunden kann. Hierbei sollte sich der Entwickler bewusst sein, dass eine komplex aufgebaute Welt Orientierungspunkte braucht. Allerdings sollten diese Orientierungspunkte nicht nur visuell repräsentiert werden, sondern auch akustisch. Bei einem rein visuellen Orientierungspunkt ist es oft schwierig diesen hinter anderen Objekten, wie beispielsweise Wänden oder gar Gebäuden, wahrzunehmen. Wird dieser Bezugspunkt jedoch akustisch unterstützt, ist dieser auch durch andere Objekte und ohne ihn direkt betrachten zu müssen wahrnehmbar. Diesbezüglich kann es auch hilfreich sein kleinere Orientierungspunkte in die Umgebung einzubauen. Dies könnte ein kleiner Wasserfall, ein großes Lagerfeuer oder auch ein laut aufgedrehter Radio sein. All diese akustischen Objekte können als Referenzpunkte angesehen werden, wodurch der Nutzer seine Position und Entfernung zu bestimmten Punkten oder Orten in der Welt bestimmen kann.

Des Weiteren sollte eine Art der Führung bzw. Hilfestellung zum Vorankommen im Spiel vorhanden sein. Weiß der Nutzer beispielsweise nicht, wo er nach dem Erkunden eines Ortes weitersuchen soll, kann dies durch Hinweise in Form eines kurzen auditiven Signals am Zielort oder der Stimme der Spielfigur stattfinden. Oft werden hörbare Signale, die sich vom Rest der VAE unterscheiden untersucht. Stimmen von Spielfiguren können darüber hinaus weitere wichtige Informationen zur Spielwelt liefern, die visuell vielleicht gar nicht in dieser Form übermittelt werden können. Nebencharaktere können

sich mit dem Protagonisten bzw. dem Spieler unterhalten oder in Form von Tonaufnahmen Informationen zum gerade erkundeten Ort liefern. Durch das Nutzen der Stimme können zusätzliche Informationen der Geschichte über die Audioebene transportiert werden.

Das Interesse einer VAE besteht darin, eine natürliche Soundkulisse zu erschaffen und dem Nutzer eine problemlose und natürliche Interaktion und Wahrnehmung zu ermöglichen. Wie in Abschnitt 5.1 bereits durch McMullen erklärt, können sogenannte *Auditory Icons* eingebaut werden, deren Bedeutung durch die reale Welt bereits bekannt sind. Infolgedessen sollten vor allem interaktive Gegenstände vertont werden. In diese Kategorie würden Schubladen und Schränke, aber auch Türen fallen. Würde man eine Tasse fallen lassen, ohne dass diese ein Geräusch beim Zerschlagen von sich gibt, würde dies unserem Verständnis und unserer Erfahrung aus der realen Welt widersprechen und ein ungutes Gefühl hinterlassen, welches sogar in einer Störung der Immersion resultieren könnte.

Ebenso wichtig für den Aufbau einer VAE sind die atmosphärischen Hintergrundgeräusche, die Atmos. Diese stellen einen Bezug zur Lage des Ortes her. Stellt die Szene zum Beispiel einen Wald dar, werden unter anderem das Rauschen von Blättern, das Summen von Insekten, sowie vielleicht das Plätschern eines Flusses zu hören sein. All dies trägt zum Gefühl der Umgebung bei und es kann auch ohne die visuelle Repräsentation eine Annahme über die Art des Ortes getroffen werden. Eine Stadt besitzt somit eine andere Atmo und kann auch mit geschlossenen Augen von einem Wald unterschieden werden. Dadurch ist es häufig hilfreich, sich Orte aus der realen Welt anzusehen und deren Atmo genau zu untersuchen. Selbst abstrakte Orte können Referenzen aus der realen Welt zum Aufbau einer Atmo heranziehen und gewohnte Sounds verfremden oder sogar modifizieren.

All diese Gesichtspunkte sollten bei der Konzeption einer explorativen VAE im Vorhinein berücksichtigt werden, um ein möglichst immersives und natürliches Spielerlebnis liefern zu können. Bei der Konzeption all dieser Sounds kann es unter anderem hilfreich sein, selbst eine Karte der virtuellen Welt samt allen wichtigen, vorkommenden akustischen Bezugspunkten und Führungshilfen anzufertigen. Dadurch kann man sich einen Überblick über alle auditiven Hinweise auf dem Weg des Spielers beim Erkunden der Geschichte verschaffen. Durch dieses Hilfsmittel während der Konzeption ist es eventuell möglich, Schwachstellen im eigenen Audiokonzept zu erkennen und diese frühzeitig zu beheben.

## 6.2 Methoden der Wegfindung

Bei der Wegfindung bzw. Orientierung in einer virtuellen Umgebung gibt es mehrere Aspekte, die diese beeinflussen können. Um Entscheidungen über eine Umgebung treffen zu können muss die Umgebung analysiert werden, wodurch unbewusst eine kognitive Karte der Umgebung entsteht. Dörner et al. definieren drei Arten von Wissen über die Umgebung [5, S. 169f.]: Landmarkenwissen, Routenwissen und Übersichtswissen.

Bei Landmarkenwissen handelt es sich um das Wissen von herausstechenden, oft einmaligen Bezugspunkten im Raum. Diese Bezugspunkte werden Landmarks genannt und bieten durch ihre Einzigartigkeit und weite Sichtbarkeit in der Umgebung Bezugspunkte zur Orientierung. Ein Beispiel hierfür wäre ein Leuchtturm, der auch von weiter

Weg noch sichtbar ist. Von großem Vorteil wäre es, wenn diese Landmarks auch vertont sind und von Weitem, oft hinter anderen Objekten verdeckt noch wahrgenommen werden können und so die Orientierung in der Umgebung erhalten bleibt. Durch ihre normalerweise statische Position ist es möglich die eigene Position, aber auch die von anderen Schallquellen, sowie die Entfernung zur Landmark kontinuierlich abzuschätzen.

Beim Routenwissen, das auch prozedurales Wissen genannt wird spielen die Aktionen, die zur Routenfolge wichtig sind eine Rolle. Dabei merkt sich der Nutzer die Anhaltspunkte in der Szene und die dafür benötigten Interaktionen, um der Route weiter folgen zu können. Häufig „[...] können Hilfsmittel wie ein digitaler Kompass, Hinweisschilder oder Wegmarkierungen den Erwerb von Routenwissen unterstützen“ [5, S. 169]. Wird auch hier auf die Hilfe von auditiven Hinweisen gesetzt, kann sich der Spieler im Nachhinein an die verwendeten Punkte im Spiel erinnern.

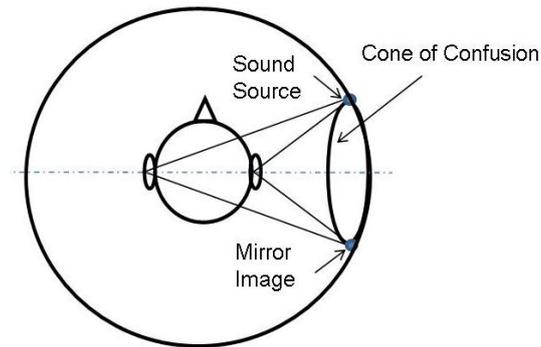
Beim Übersichtswissen wird auf die topologischen Eigenschaften der Umgebung eingegangen. Der Erwerb dieser Fähigkeit dauert laut Dörner et al. am längsten und es wird dabei auf das Landmarken- und Routenwissen zurückgegriffen. Visuell wird dieses Wissen häufig mit Miniaturdarstellungen der Umgebung oder interaktiven Landkarten unterstützt [5, S. 169]. Beim Übersichtswissen werden alle Sinne des Nutzers angesprochen, somit auch der Hörsinn. Wie zuvor bei McMullen in Abschnitt 5.1 erklärt, ist es Nutzern möglich sich an eine dichte auditive Umgebung besser zu erinnern als eine dünn besiedelte. Durch die Erinnerung an die verschiedenen akustischen Reize ist es somit möglich, sich verschiedene Bereiche einer Umgebung besser zu merken. Ebenfalls eine Möglichkeit sind kleine Objekte oder auditive Hinweise, die die Aufmerksamkeit des Spielers so auf sich lenken, dass er diese genauer erkunden möchte und so schrittweise den vordefinierten Pfad des Entwicklers folgt. Hierbei kommt das Prinzip der Salienz zum Einsatz. Dörner und Steinicke beschreiben dieses Prinzip folgendermaßen [3, S. 57]:

Die menschliche Wahrnehmung hat nicht die Kapazität, um alle Umweltreize gleichermaßen ausführlich zu verarbeiten. Es werden Schwerpunkte gesetzt, der Mensch kann Aufmerksamkeit auf bestimmte Aspekte richten.

Durch diese gefilterte Aufmerksamkeit von einzelnen Schallquellen an wichtigen Stellen der Geschichte oder des Weges, ist es dem Spieleentwickler möglich, die Aufmerksamkeit des Spielers an gewissen Punkten zu bündeln und ihn so langsam auf den gewollten Weg zu führen.

### 6.3 Auditive Lokalisierung in VAE

T. Letowski und S. Letowski beschäftigen sich in [11] mit verschiedenen Studien und Erkenntnissen zum Thema der auditiven Lokalisierung in virtuellen Welten und erklären, dass sich die auditive Lokalisierung aus Beurteilungen der Positionen von Schallquellen auf der horizontalen und vertikalen Ebene zusammensetzt. Meist ist die Lokalisierung auf der horizontalen Ebene durch binaurale Signale einfacher als auf der vertikalen Ebene. Wie bereits von Lokki und Gröhn in Abschnitt 5.2 festgestellt wurde, resultieren diese Lokalisationsfehler auf der vertikalen Ebene aus fehlenden Informationen auf der Elevationsebene. Häufig wird bei Schallquellen, die auf der vertikalen Ebene platziert werden nicht auf die Besonderheit der Höheninformation geachtet. Die von Lokki und



**Abbildung 6.1:** Darstellung des „Cone of Confusion“, bei dem sich ein imaginärer Kegel von beiden Ohren bis zur Schallquelle erstreckt und die gleiche Laufzeitdifferenz eine Lokalisierung unmöglich macht [11].

Gröhn empfohlenen Veränderungen an den Parametern Lautstärke, Tonhöhe und Pulsrate schufen bessere Resultate zu Lokalisierung von Sounds in der vertikalen Ebene. Befindet sich der Spieler nun in einer Umgebung, in der er herausfinden soll, dass eine Schallquelle sich beispielsweise auf einem bestimmten Berg in der virtuellen Umgebung befindet, würde ihm die stetig ansteigende Lautstärke und die veränderte Tonhöhe darauf hinweisen, auf welchem der Berge sich dieser Sound befindet.

Binaurale Hinweise sind auf die horizontale Ebene limitiert, da die interaurale Zeitdifferenz (ITD) und die interaurale Pegeldifferenz (IID) nur Informationen auf der Azimuth-Ebene bieten. Während die ITD durch tiefe Frequenzen unter etwa 800 Hz für die links-rechts Lokalisierung verantwortlich ist, beinhaltet die IID hohen Frequenzen über 1600 Hz für die links-rechts Lokalisierung [11, S. 12]. Höhere Frequenzen werden zwar leichter lokalisiert, dennoch bieten nur beide zusammen eine effektive Art der Lokalisierung. Würden jedoch mehrere Töne mit der gleichen Laufzeitdifferenz beim Nutzer ankommen, könnte ein „Cone of Confusion“ entstehen [11, S. 13], dessen Prinzip in Abbildung 6.1 schematisch abgebildet wird. Durch das gleiche Verhalten der Schallquellen ist es dem Zuhörer nicht möglich die Position der Schallquelle auszumachen, wodurch es ihm auch nicht möglich ist zu sagen, ob das Geräusch von hinten oder von vorne kommt. Erst durch das Drehen des Kopfes kann eine Differenzierung zwischen den Positionen der beiden Schallquellen getroffen werden.

Nicht nur durch die Kopfbewegung, sondern auch durch dynamische Schallquellen ist es möglich diese Differenzierung zwischen vorne und hinten zu treffen. Dabei spielt vor allem die Dauer der dargebotenen Sounds eine große Rolle. T. Letowski und S. Letowski sprechen in [11, S. 17] von einer Mindestdauer von 600–800 ms, damit der Effekt der Kopfbewegung überhaupt zur Lokalisierung beitragen kann. Je länger ein Geräusch anhält, desto leichter fällt es dessen Position zu bestimmen. Des Weiteren spielt die Vertrautheit mit einer Schallquelle eine wichtige Rolle, da vertraute Geräusche leichter lokalisiert werden können, als fremd klingende Sounds. Je länger eine Schallquelle erklingt, desto länger hat der Spieler Zeit sich an den wiedergegebenen Sound zu erinnern.

T. Letowski und S. Letowski erwähnen ebenfalls den *Präzedenz-Effekt*, bei dem der Hörer ein Urteil über die Position in einem geschlossenen Raum anhand des zuerst ein-

treffenden Schallsignals trifft und dabei alle gleich klingenden Schallquellen aus anderen Richtungen ignoriert [11, S. 18]. Besitzt das zweite eintreffende Signal eine Zeitverzögerung von unter 1 ms, so befindet sich die wahrgenommene Schallquelle zwischen diesen beiden existierenden Schallquellen. Würde das zweite eintreffende Signal eine Verzögerung von etwa 20 ms besitzen, wäre es als Echo zu hören [11, S. 18f.]. Dies zeigt, dass Schallquellen auch in reflektiven Umgebung lokalisiert werden können. Allerdings ist die Position einer impulsartigen Schallquelle leichter zu lokalisieren, als die einer kontinuierlich klingenden oder langsam steigenden Schallquelle. Dies liegt daran, dass Reflexionen von einmaligen, kurzen Signalen leichter zugeordnet werden können, als sich ständig erneuernden Reflexionen von kontinuierlichen Signalen. In einem VR-Spiel könnte dies zum Beispiel ein Schuss sein, der kurz in einem reflektiven Raum zu hören ist. Die Lokalisierung dieses Schusses wäre zwar schwierig, aber der Spieler könnte dennoch die Richtung und ungefähre Position bestimmen.

Ein weiterer Punkt ist das Zusammenspiel zwischen Seh- und Hörsinn. Hierbei würde immer der Sehsinn gewinnen, da Menschen sich mehr auf ihre Augen als ihre Ohren verlassen. Dieser Vorzug des Sehsinns kann auch die Wahrnehmung des Audios eines VR-Spiels beeinflussen. Hierfür wichtig ist der *Bauchredner-Effekt*, bei dem Menschen ein Objekt als die Quelle eines Sounds ansehen, selbst wenn es gar nicht die wahre Quelle dieses Sounds ist [11, S. 20]. Neben dem Bauchredner-Effekt sind auch die Erwartung und Vertrautheit eines Sounds erneut von Bedeutung. Diese führen dazu, dass akustischen Signale im Sinne der Erwartung oder Wiedererkennung eines Sounds übergangen werden. Ein Beispiel hierzu wäre laut T. Letowski und S. Letowski der Schrei eines Adlers, bei dem die meisten Menschen instinktiv nach oben schauen würden und das, obwohl der Schrei in Bodenhöhe abgespielt wurde [11, S. 21].

Beim Einsatz von auditiven Hinweisen ist ebenso auf den Umkehrlokalisierungsfehler zu achten, bei dem die geschätzte Richtung der Schallquelle genau gegenüber liegt [11, S. 65]. Dieser Fehler kann auftreten, wenn die binaurale Information gleich gut zu einer anderen räumlichen Position passt, ähnlich dem Cone of Confusion. Sie kann durch Störungen des umgebenden Raumes, wie beispielsweise Reflexionen, durch das verwendete Headset oder auch die Erwartung des Spielers ausgelöst werden. Meist tritt dieses Phänomen bei kurzen Signalen oder Schmalband<sup>1</sup>-Signalen auf [11, S. 65].

Bei der Verwendung von mehreren Schallquellen sollte auf die Wahl der Signale geachtet werden. Existieren ähnlich klingende Sounds gleichzeitig an unterschiedlichen Orten, tritt eine Gruppierung dieser Sounds auf und es wird nur eine Position wahrgenommen – die Position der Schallquelle mit der tiefsten Frequenz [11, S. 76]. Deswegen sollten Sounds, wenn sie als nicht zusammengehörig angesehen werden sollen, gut differenzierbar und in der Umgebung verteilt sein, damit sie der Spieler wie intendiert wahrnimmt. Menschen können zwischen der Bewegungsrichtung, der geschätzten zurückgelegten Entfernung und der Geschwindigkeit einer beweglichen Schallquelle unterscheiden. All dies ist unter anderem durch die Winkel- und die Radialgeschwindigkeit, sowie der Intensität, der interauralen Zeitdifferenz und einer Frequenzverschiebung möglich. Die Winkelgeschwindigkeit gibt die Geschwindigkeit an, mit der sich der Sound um den Nutzer bewegt, während die Radialgeschwindigkeit die Geschwindigkeit angibt, mit der sich der Sound auf den Spieler zu oder von ihm weg bewegt [11, S. 78]. Wenn sich ei-

---

<sup>1</sup>Schmalband beschreibt einen Filter, der nur einen bestimmten Frequenzbereich hindurch lässt.

ne Schallquelle nun zum Spieler hin bewegt ändert sich die Intensität jener Schallquelle, was zu veränderten Frequenzen führt. Diese Frequenzänderung wird als *Doppler Effekt* wahrgenommen, bei der sich die Tonhöhe wahrnehmbar ändert [11, S. 78f.]. Nähert sich die Schallquelle dem Nutzer, so wird das Schallsignal höher, bewegt sich diese erneut weiter weg, wird die Tonhöhe erneut tiefer. Würde sich der Spieler in der virtuellen Umgebung an einer Straße wiederfinden und ein Krankenwagen mit eingeschaltetem Martinshorn vorbeifahren, so würde sich das Signal je nach Position zum Spieler verändern. Ist der Krankenwagen noch weiter entfernt, klingt das Signal heller, ist dieser bereits vorbeigefahren, klingt er danach etwas tiefer.

## 6.4 Erinnerungsvermögen, Diegese und Interaktivität

Zur besseren Orientierung in einer VAE ist es wichtig das Erinnerungsvermögen des Spielers zu steigern. Die Interaktion mit der Audioebene ist eine gute Möglichkeit das Erinnerungsvermögen der virtuellen Umgebung zu steigern.

### 6.4.1 Erinnerungsvermögen

„Die Wahrnehmung ist dort feiner strukturiert, wo ein reichhaltiger Erfahrungshintergrund besteht“ [6, S. 109]. Durch den Lernprozess, den ein Mensch im Laufe seiner Lebens durchmacht reichen oft nur Bruchstücke aus, die Dank unserer Erfahrung und Erinnerung zur Wiedererkennung führen. McMullen und Wakefield haben eine Studie [14] zum Thema Erinnerungsvermögen der Bezeichnung und Position von gleichzeitig abgespielten Sounds durchgeführt. Dabei mussten die Probanden die Positionen und die Bezeichnung der gleichzeitig abgespielten Sounds wiedergeben. In der Testphase wurden drei verschiedene Varianten der Benennung getestet. Zuerst wurden die Positionen der Schallquellen vorgespielt, zeitgleich mit dem Hören nacheinander markiert und nach dem Test benannt. Beim zweiten Durchgang wurden zuerst alle vorhandenen Sounds abgespielt und im Nachhinein deren Position und Bezeichnung wiedergegeben. Die letzte Variante sah vor, dass die Sounds zuerst alle gehört wurden, dann zwei dieser Sounds erneut vorgespielt wurden und der Proband diese in der richtigen zuvor gehörten Reihenfolge markieren und benennen sollte. Die Test hierzu haben ergeben, dass Menschen sehr wohl dazu in der Lage sind sich an die Bezeichnungen und Positionen der gleichzeitig abgespielten Sounds zu erinnern. Des Weiteren wird empfohlen, dass eine VAE kreiert wird, welche die Dauer zwischen dem Wahrnehmen der Sounds und deren erinnern minimiert, um eine möglichst genaue Erinnerung dieser Schallquellen und deren Positionen im Kopf zu behalten [14, S. 102].

### 6.4.2 Diegese

Der Begriff der Diegese behandelt grob die Tatsache, ob sich etwas innerhalb oder außerhalb der erzählten Welt befindet [28]. Dieser Begriff wird oft im Filmbereich dafür verwendet, um die Anordnung der unterschiedlichen Ebenen einer Erzählung und ihre Beziehungen zueinander zu beschreiben. Im Zusammenhang mit VR ist dies allerdings etwas anders, da der Spieler und die Hauptfigur des Spiels fast immer ein und dieselbe Person sind. Salamoni et al. befassen sich in [17] mit dem Thema der diegetischen Be-

nutzeroberflächen. In VR seien alle Charaktere, Objekte, Geräusche und Umgebungen innerhalb der virtuellen Welt, die der Nutzer instinktiv als Teil der Narrative wahrnehmen und direkt selbst erfahren kann, diegetisch [17, S. 174]. Unter nicht-diegetisch kann man daher alle Elemente, unter anderem auch akustische und visuelle Elemente, verstehen, die zwar kontextuell gestaltet sein können, jedoch externe Informationen darstellen [17, S. 174]. Somit wären beispielsweise Overlay-Menüs<sup>2</sup> samt ihren Interaktionsgeräuschen nicht-diegetisch.

### 6.4.3 Interaktivität

Um die Dauer des Erinnerungsvermögens dennoch zu steigern, kann die Interaktivität von großer Bedeutung sein. Collins wandelt in [2] den aus der Filmanalyse bekannten Begriff der Diegese für die Kategorisierung von Videospiel-Sound ab. Videospiel-Sounds können grob in diegetisch und nicht-diegetisch, aber auch in nicht-dynamische und dynamische Sounds eingeteilt werden. Danach erfolgt eine weitere Einteilung in die Typen der dynamischen Aktivität, die mit der Diegese und dem Spieler zusammenhängen. Allerdings ist zu beachten, dass der Grad der dynamischen Aktivität in Videospielen meist eher fließend ist, was eine Klassifizierung eher schwierig macht [2, S. 125]. Da diese Kategorien für klassische Videospiele ausgelegt sind, wurde versucht die bestehenden Kategorien für VR-Spiele abzuändern. Die Bezeichnungen der Kategorien aus dem Buch wurden, da sie nicht mit den Gegebenheiten von VR übereinstimmen, weggelassen und nur die Erklärungen selbst wurden adaptiert.

Die erste Kategorie beschreibt Cinematics oder Introfilme, bei der keine Interaktion mit dem Audio selbst stattfindet [2, S. 125f.]. Der Spieler ist nur Beobachter in solchen Szenen, kann die Perspektive der Kamera im Regelfall jedoch beliebig durch Kopf- und Körperbewegungen ändern. Somit ist es ihm zwar möglich die „Kameraeinstellungen“ zu ändern, er kann jedoch nicht mit dem Audio oder gar der Szene selbst interagieren. Ein Beispiel hierfür wäre *Batman: Arkham VR* [25], bei dem der Spieler die Geschichte vom Tod der beiden Eltern Batmans in Form eines Cinematics miterlebt, wobei ein kurzer Ausschnitt der Szene in Abbildung 6.2 zu sehen ist. Da die Spielfiguren sich in der gleichen virtuellen Welt wie der Spieler befinden, handelt es sich bei den Stimmen und Geräuschen um diegetische Sounds.

Die zweite Kategorie befasst sich mit adaptiven Sounds, die durch das Gameplay selbst, aber nicht auf die direkte Spielerbewegung reagieren [2, S. 126]. Hierbei sind die Sounds erneut diegetisch, auch wenn der Spieler sie nicht direkt beeinflussen kann. Die Events werden durch Ereignisse der Spielmechanik selbst ausgelöst, zum Beispiel der Wechsel von Tag und Nacht, der das Audio beeinflusst. Beispielsweise könnte nur nachts das Zirpen von Grillen zu hören sein, während dieser Sound tagsüber nicht in der virtuellen Welt existiert.

Durch die Interaktion des Spielers in der dritten Kategorie, ist es möglich, dass Audio durch direkte Interaktionen des Spielers selbst zu verändern [2, S. 126]. Würde sich der Spieler zum Beispiel einem gegnerischen Charakter nähern, der durch Musik begleitet wird, könnte er die Lautstärke der Musik oder die Lautstärke der Geräusche

---

<sup>2</sup>Das Overlay-Menü ist eine Form des Menüs, das über der Spielwelt eingeblendet wird und diese überlagert. In VR können diese Art von Menüs sogar blickbasiert sein und sitzen somit auch bei Kopfbewegungen immer im Sichtfeld des Spielers.



**Abbildung 6.2:** Ausschnitt der Szene in der Batmans Eltern ermordet werden. Der Spieler kann nur die Kameraperspektive ändern, jedoch mit keinem Objekt in der Szene interagieren [25].

des Gegners anhand seiner Position zu diesem beeinflussen. Je näher er ihm käme, desto lauter wird die Musik und je weiter weg der Spieler sich vom Gegner bewegt, desto leiser wird sie. Das Audio ist somit durch den Spieler direkt beeinflussbar.

Die nächste Kategorie beinhaltet nicht-dynamische, adaptive Sounds, die jedoch nicht interaktiv sind [2, S. 126]. Hierbei handelt es sich um diegetisches Audio, das in der Welt der Spielfigur erscheint, allerdings kann erneut keine Beeinflussung des Audios stattfinden. In *The Gallery: Episode 1 – Call of the Starseed* [22] ist die Stimme des Professors an bestimmten Stellen der Geschichte immer zu hören, was der Ortung seiner Position und zum Verständnis der nächsten Aufgabe dient. Der Ausschnitt einer Szene mit dem Professor ist in Abbildung 6.3 zu sehen. Seine Stimme lässt sich nicht vom Spieler ändern, befindet sich aber in der gleichen virtuellen Welt.

Des Weiteren gibt es adaptive, diegetische Sounds, die meist für Ambience Sounds, wie Vogelzwitschern oder Grillenzirpen beim Tag-Nacht-Wechsel verwendet werden [2, S. 126f.]. Diese sind somit adaptiv an der Tageszeit orientiert und diegetisch für den Hauptcharakter des Spiels zu hören. Interaktive, diegetische Sounds befinden sich erneut in der Welt der Spielfigur und können direkt von dieser beeinflusst werden. Meist handelt es sich bei ihnen um Soundeffekte, wie Schritte oder auch Waffengeräusche des Hauptcharakters im Kampf.

Der letzte wichtige Punkt laut Collins sind die sogenannten kinetischen Gesteninteraktionen, die bei diegetischen und nicht-diegetischen Sounds vorkommen, bei denen der Spieler selbst körperlich mit den Sounds in der Spielwelt interagiert [2, S. 127]. Einige VR-Anwendungen können auch mit Gamepads gespielt werden, allerdings sollte, falls vorhanden, immer eine Bewegungssteuerung wie die Oculus Touch<sup>3</sup> verwendet werden. Nicht-diegetische Sounds könnten beispielsweise die Soundeffekte bei der Auswahl

<sup>3</sup>Bei den Oculus Touch handelt es sich um die Controller der Oculus Rift für Gesteninteraktionen.



**Abbildung 6.3:** Ausschnitt einer Szene, in der die Stimme des Professors zuerst zu hören ist bevor er sich schließlich zeigt [22].

im Overlay-Menü betreffen, während diegetische Sounds bei Interaktionen mit Gegenständen auftauchen. Beispiele von Sounds hierfür wären beim Öffnen bzw. Schließen von Schubladen oder Türen. Durch die vertonten Interaktionen lässt sich die Immersion der VAE steigern. Ebenso kann durch Interaktionen mit vertonten Objekten auch Feedback erhalten werden. Erst durch die Interaktion mit dem Gegenstand wird der passende Sound ausgelöst und bietet dem Spieler somit Feedback zur Situation dieses Gegenstandes.

Interaktionen können zur Optimierung der Orientierung ebenfalls genutzt werden. Es gibt in VR auch Interaktionen, wie interaktive, diegetische Sounds, die dem Spieler später bei der Orientierung – durch Erinnerung der Aktionen und der damit verbundenen Positionen – und dem „Zurechtfinden“ im virtuellen Raum helfen können. Durch das direkte Interagieren mit einem Gegenstand und dadurch der Audioebene, vergrößert sich die kognitiv angelegte Karte erneut. Je dichter eine Umgebung an gut durchdachten und gesetzten Schallquellen ist, desto besser lässt sich diese später vom Spieler rekonstruieren und in Erinnerung rufen.

## Kapitel 7

# Aufbereitung von auditiven Hinweisen

Anknüpfend an das Wissen der vorhergehenden Kapitel, wird nun auf die Kategorisierung von Klangobjekten, sowie die Aufbereitung von Audiodateien in der exemplarisch ausgewählten Unreal Engine eingegangen.

### 7.1 Kategorisierung von Klangobjekten

Bei der Verwendung von Klangobjekten sollte auf ihre gesellschaftliche Bedeutung geachtet werden. Gewisse Geräusche oder Signale wurden durch die Gesellschaft definiert und sind den meisten Menschen durch die Begegnung im Alltag bekannt. Flückiger beschreibt und kategorisiert diese Klangobjekte [6, S. 158–189].

*Signale* sind „Klangobjekte, die einen gesellschaftlich definierten kommunikativen Gehalt haben“ [6, S. 159]. Meist sind Signale eine Aufforderung zu einer bestimmten Handlung und sind mit Warnhinweisen verknüpft. Sie sind normalerweise einfach gestaltet – bestehen nur aus einem oder zwei Sinustönen – und sollen durch ihre warnende Funktion auffallen, wodurch die Grundfrequenz sich daher zwischen 1 und 2 kHz befindet, worauf der Hörsinn äußerst sensible reagiert [6, S. 159]. Beispiele für Signale sind Glocken, Hupen, Sirenen oder auch Telefonklingeln.

Bei der Kategorie der *Symbole* gibt es keine wirklich festgelegte Bedeutung, sondern sie benötigen häufig eine Interpretation zum Verständnis [6, S. 164]. Ein Merkmal des Symbols ist es, dass es sprachlich nicht restlos aufgeklärt werden kann und bietet somit eher ein abstraktes Konzept. Das Symbol kann je nach Situation für ein anderes Thema stehen. Eine Glocke beispielsweise kann laut Flückiger je nach Situation unterschiedliche Bedeutungen haben. Einerseits kann sie eine reine Signalfunktion, andererseits eine Orientierungsfunktion besitzen oder aber auch für den Tod stehen [6, S. 166ff.]. Natürlich kann sie aber auch eine eigens definierte Symbolik haben, die nur innerhalb der selbst erstellten Welt gilt.

Unter *Key Sounds* sind all jene Klangobjekte versammelt, die durch ihre deutliche Häufung in Kernszenen mit einer spezifischen Bedeutung geladen werden [6, S. 175f.]. Durch diese Regelmäßigkeit wird vom Spieler angenommen, dass hinter der Verwendung dieses Sounds noch eine tiefere Bedeutung steckt. Ein Key Sound kann leicht verändert werden und behält trotzdem seine Bedeutung. Ein Beispiel hierfür wäre ein Sound der bereits zu Beginn des Spiels hörbar ist, sich über die Zeit hinweg stetig steigert und

schließlich mit einer Person oder einem Gegenstand in Verbindung gebracht wird.

*Stereotypen* werden erst durch die häufige Wiederholung mit Bedeutung versehen und werden so als Konstante im Langzeitgedächtnis gespeichert [6, S. 176f.]. Die Verwendung von Stereotypen wird meist nur während einer bestimmten kulturellen Zeit und in bestimmten Kulturkreisen verstanden, da sie auch dem Wandel der Zeit unterliegen und nur durch ständige Wiederholungen in unterschiedlichen Werken verstanden werden können. Eine Glocke, die während einer angedeuteten Todesszene spielt, kann durch ihre häufige Verwendung in anderen Werken als das Stereotyp des Todes angesehen werden. Spieler, die andere Werke mit ähnlichen Sequenzen und dem Einsatz der Glocke für den Tod gespielt haben, verstehen somit die Andeutungen.

Die letzte Kategorie betrifft den Begriff *Leitmotiv*, der von Neubauer folgendermaßen definiert wird [15, S. 178]:

[...] eine Tonfolge, welche innerhalb eines Musikstückes eine oder mehrere Personen, Ereignisse, Situationen, Gegenstände charakterisieren soll. Leit-motive sind während eines musikalischen Ablaufs variierbar. Die jeweiligen Veränderungen eines Leitmotivs symbolisieren die Entwicklungen und Wandlungen einer Bezugsgröße.

Erneut spielt die Wiederholung eine bedeutende Rolle, da durch den wiederholten Einsatz die Bedeutung des Leitmotivs erst gelernt werden kann. Flückiger erklärt, dass auch ein Klangobjekt zu einem Leitmotiv werden kann [6, S. 187]. Hierfür muss es wiederholt eingesetzt und dynamisch verändert werden, damit es mit einer abstrakten Symbolik versehen werden kann. Ein Sound ohne irgendeine Veränderung kann jedoch nicht als Leitmotiv angesehen werden, da jeglicher dynamischer Aufbau bzw. Veränderung fehlt, der für das musikalische Leitmotiv typisch ist.

## 7.2 Technische Voraussetzungen der Audiodateien für VR-Anwendungen

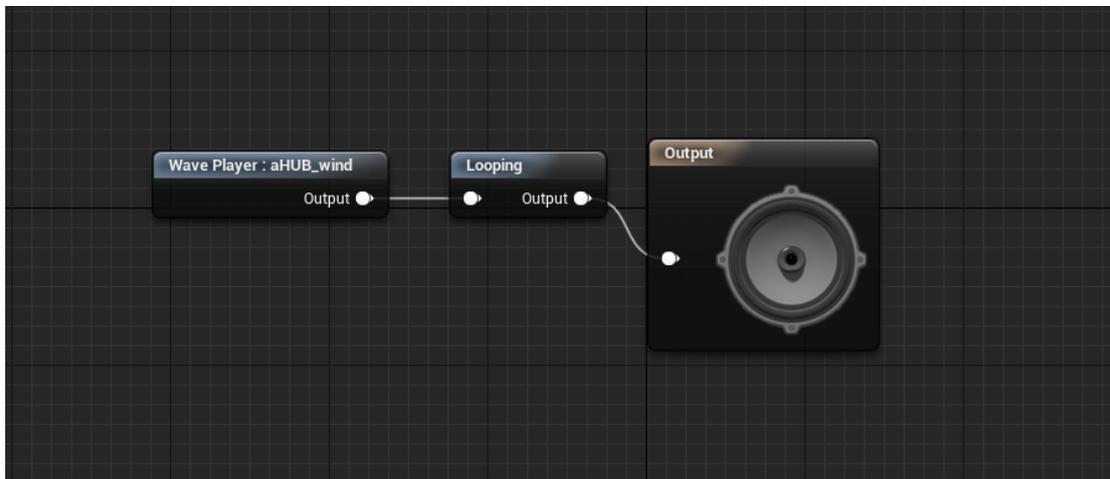
Um den Effekt des 3D Audios auch in der VR-Anwendung berechnen zu können, sind einige bestimmte Voraussetzungen der Audiodateien nötig. Als Beispiel für eine Game Engine und ihre Funktionsweisen wird die Unreal Engine in der Version 4.19 genutzt, zu der die wichtigsten Audioeinstellungen erklärt werden.

### 7.2.1 Vorbereitung

Um eine funktionierende Audiodatei für die Unreal Engine aufzubereiten, müssen einige Faktoren beachtet werden. Audiodateien müssen zuvor als Mono mit 44100 Hz und 16 Bit im WAV-Dateiformat exportiert werden [30]. Andernfalls ist es dem Programm nicht möglich, die Audiodatei mit den passenden 3D Audio-Effekten zu versehen oder diese je nach verwendetem Dateiformat gar zu öffnen. Die Audiodatei muss deshalb Mono sein, damit der Sound auch nur von einer bestimmten Quelle zu kommen scheint und die Berechnungen der Position des Sounds in der virtuellen Umgebung nicht mit Informationen von zusätzlichen Kanälen und somit Richtungen erschwert werden. Danach können diese ins gewünschte Projekt importiert werden.

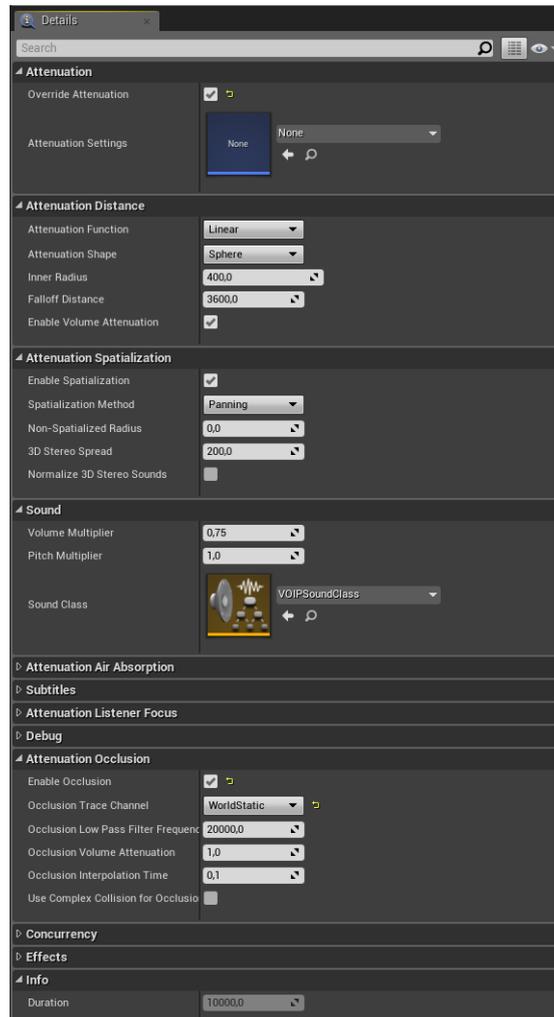
### 7.2.2 Audioeinstellungen der Unreal Engine

In der Unreal Engine gibt es einige Möglichkeiten, um den binauralen und räumlichen Effekt imitieren und anwenden zu können. Möchte man diese Effekte auf die Audiodatei einwenden, so sollte im Idealfall zuvor ein *Sound Cue* erstellt werden. Dieser ermöglicht eine Bearbeitung durch das Node-System der Unreal Engine, das in Abbildung 7.1 abgebildet wird. Allerdings sollte man auf die *Auto Activate* Funktion achten, da diese immer aktiviert wird, sobald ein Sound Cue ins Level gesetzt wird. Diese bewirkt, dass der Sound, sobald das Level im Spiel gestartet wird, automatisch beginnt zu spielen. Die Option *Spatialize* ermöglicht es, dass Audio von nur einer bestimmten Position zu hören ist und ahmt somit den binauralen Effekt des Hörens nach. Ohne diese Funktion würde die Lokalisierung bei einer Kopfbewegung nicht möglich sein.



**Abbildung 7.1:** Der Audio Node Graph zeigt die Audioquelle und die damit verbundenen Nodes an, die es verändern [23].

Eine weitere wichtige Funktion stellt die *Attenuation* dar, die den minimalen und maximalen Radius des Lautstärkenabfalls regelt [30]. Befindet man sich innerhalb des minimalen Radius, so ist die Lautstärke der Schallquelle auf 100%, bewegt man sich jedoch von dem minimalen in den maximalen Radius, nimmt die Lautstärke der Schallquelle je nach Berechnungsart ab. Verlässt man den maximalen Radius, so sollte die Schallquelle gar nicht mehr zu hören sein. Es gibt unterschiedliche Arten der Attenuation, unter anderem linear, logarithmisch und natürlich, welche die Stärke des Lautstärkenabfalls und deren Form beeinflussen. Gleich im Zuge dessen sollte überlegt werden, ob die Option *Virtualize when Silent* gebraucht wird. Diese Option gibt an, ob die Audiodatei stumm weiterspielen soll, selbst wenn der Spieler den maximalen Radius der Attenuation verlassen hat. Durch diese Option ist es möglich, dass der Spieler beim erneuten betreten des Attenuation-Bereichs die gleiche Schallquelle weiterspielen hört. Andernfalls würde sie einfach stoppen und beim erneuten Betreten nicht mehr zu hören sein. Eine weitere nützliche Option ist *Occlusion*, durch die Sound Cues durch zuvor definierte Objekte der virtuellen Umgebung abgedämpft werden. So wäre zum Beispiel ein Radio in einem Zimmer durch die Wand des Nebenzimmers gedämpft. Die Stärke der Dämpfung kann



**Abbildung 7.2:** Das Details Panel zeigt die verschiedenen Einstellungsoptionen bezüglich Attenuation, Spatialize, Virtualize when Silent und Occlusion an [23].

in der Option selbst eingestellt werden. Diese und alle zuvor erwähnten Einstellungen des *Details Panels* können in Abbildung 7.2 betrachtet werden.

Des Weiteren besteht die Möglichkeit des *Volume Ducking*, bei der eine bestimmte Sound Class<sup>1</sup>, die zuvor erstellt und jedem Sound zugewiesen werden sollte, hervorgehoben wird, indem alle anderen Sound Classes in ihrer Lautstärke reduziert werden [30]. Normalerweise wird diese Funktion für Dialoge angewandt. Im Allgemeinen sollte man auf die Lautstärken der eingesetzten Sound Cues achten. Werden viele Audiodateien auf einmal verwendet summiert sich die Lautstärke all jener, bis einige Sounds nicht mehr hörbar sind oder gar übersteuern.

Es gibt noch viele weitere Möglichkeiten Audio mit Hilfe der Unreal Engine zu bearbeiten oder auch Erweiterungen wie *FMOD Studio* [29], welche die Möglichkeiten

<sup>1</sup>Eine Sound Class ist eine Gruppe, der bestimmte Typen von Sounds wie Hintergrundgeräusche oder Dialoge zugewiesen werden, um alle gleichzeitig bearbeiten zu können.

zur Audiotbearbeitung noch erweitern. Diese Optionen können jedoch als Grundlagen angesehen werden und bieten bereits die Möglichkeit 3D Audio für eine VR-Anwendung zu nutzen.

# Kapitel 8

## Leitlinien

Anhand der Analysen der beiden Studien aus Kapitel 5 und durch die Beschäftigung mit der Funktionsweise des binauralen Audios samt Wahrnehmungsprozessen in dieser Arbeit, ist es nun möglich die wichtigsten angesprochenen Punkte zu Leitlinien zusammenzufassen. Diese Leitlinien sollen eine Hilfestellung bei der Planung und Umsetzung eines Audiokonzepts für VR-Spiele darstellen. Dabei werden die wichtigsten Aspekte in Kategorien gesammelt und deren Bedeutung kurz erläutert.

### 8.1 Wahl des Ausgabemediums

Zu Beginn des Projektes sollte überlegt werden, welches Ausgabegerät zum Verwendungszweck passt. Handelt es sich um ein VR-Spiel, das Privatpersonen zu Hause spielen sollen, ist hierfür der Kopfhörer die beste Wahl. Führt man jedoch eine kleine Studie mit seinem VR-Spiel durch, kann auch die Verwendung eines Virtual Rooms mit Lautsprechern erfolgen. Verwendet man eine Game Engine ist diese normalerweise für die Verwendung von Kopfhörern optimiert.

### 8.2 Technische Voraussetzungen einer Game Engine

Bei der Verwendung einer Game Engine sollte immer darauf geachtet werden, welche technischen Voraussetzungen an die Audiodateien gestellt werden. Die Unreal Engine beispielsweise verlangt die Aufbereitung in Mono mit 44100 Hz und 16 Bit. Andernfalls würden die zur 3D Wahrnehmung beitragenden Effekte gar nicht funktionieren.

### 8.3 Vertrautheit akustischer Signale

Das Audio selbst sollte möglichst natürlich klingen, da bekannte Sounds schneller und besser lokalisiert werden als fremd klingende Sounds. Ebenso sollte auf die Länge der Schallquellen geachtet werden, da sie bei zu kurzer Dauer durch die Kopfbewegung nicht verarbeitet und lokalisiert werden können.

## 8.4 Ambience Sounds

Um eine bestimmte Stimmung zu vermitteln, sollten auch Ambience Sounds verwendet werden. Hierbei sollte beachtet werden, da es sich häufig um einen Loop handelt, bei dem die Wiederholungen im Audio so gesetzt bzw. variiert werden, dass sie dem Spieler nicht auffallen. Wären die Schnittstellen im Audio offensichtlich, könnte dies zu einer Störung der Immersion und Präsenz führen. Ebenfalls sollte darauf geachtet werden, dass Sounds, die ähnlich zu denen der Hintergrundgeräusche sind, mit diesen gruppiert wahrgenommen und nicht als eigenständige Sounds betrachtet werden. Ambience Sounds können einen Bezug zu einem realen, bekannten Ort herstellen, wie einem Wald oder einer Großstadt. Werden in der virtuellen Welt verschiedene Arten von Umgebung gewählt, so können Atmos Auskunft darüber geben, in welchem dieser Abschnitte man sich gerade befindet.

## 8.5 Vertonung sichtbarer und interaktiver Objekte

Um die Geräuschkulisse möglichst instinktiv und natürlich wirken zu lassen, sollten alle sichtbaren und interaktiven Elemente der Spielwelt vertont werden. Hierfür sollten aus der realen Welt bekannte Geräusche gewählt werden. Würden beispielsweise nur einige der Gegenstände vertont werden während andere keine akustische Repräsentation erhalten, so würde dies ein ungutes Gefühl beim Spieler hinterlassen, da die gelernte Erfahrung aus der realen Welt nicht mit der Erfahrung aus der virtuellen Welt übereinstimmt. Die *Suspension of Disbelief* könnte somit gestört werden und der Spieler könnte an der „Wirklichkeit“ dieser Welt zweifeln.

## 8.6 Akustische Orientierungspunkte/Landmarks

Da jeder Spieler unbewusst eine kognitive Karte der virtuellen Welt anlegt, sollte die Erinnerung an bestimmte Orte mit dem Setzen von Orientierungspunkten bzw. Landmarks unterstützt werden. Landmarks, die akustisch unterstützt werden bieten darüber hinaus auch eine Orientierungspunkt, falls deren Sichtbarkeit durch andere Objekte verhindert wird. Es ist auch ratsam kleinere Landmarks in der virtuellen Umgebung zu verteilen, damit immer die eigene Position und die der wichtigsten Orte ermittelt werden kann. Bewegt sich der Spieler von einem Landmark zum anderen, so kann er auch die Beziehung unter diesen abspeichern und abrufen, falls er diese Information zur Orientierung benötigt.

## 8.7 Führungspunkte setzen

Führungspunkte auf dem Weg des Spielers können ebenfalls zur Orientierung beitragen. Verläuft sich ein Spieler oder weiß dieser nicht mehr wo er noch hingehen soll, könnten ihm diese Führungspunkte weiterhelfen. Dadurch ist es dem Spieler möglich der Geschichte und dem Weg frustfrei zu folgen. Hierfür kann das Prinzip der Salienz verwendet werden, um die Aufmerksamkeit des Nutzers zu filtern und ihn nur auf bestimmte Reize aus der Umgebung aufmerksam zu machen.

## 8.8 Zusätzliche Höheninformationen

Des Weiteren sollte auf die Lokalisierung von Sound auf der vertikalen Ebene besonders geachtet werden. Da die Höheninformationen von unbearbeiteten Audiodateien normalerweise nicht ausreicht um diese zu lokalisieren, sollten zusätzliche Informationen eingebaut werden, um diese Ortung zu erleichtern. Durch das einarbeiten der Lautstärke, Tonhöhe und/oder Pulsrate ist es möglich, eine genauere Identifikation der Schallquellen-Position zu erhalten. Hochfrequente Schallquellen sind darüber hinaus leichter zu lokalisieren, als tiefe Frequenzen. Dennoch werden beide Frequenzbereiche gebraucht, um eine effektive Ortung zu gewährleisten.

## 8.9 Reflektive Umgebungen

In reflektiven Umgebungen sollte darauf geachtet werden, dass das erste, eintreffende Signal immer zur Lokalisierung verwendet wird und das zweite, eintreffende Signal der Raumbestimmung gilt, da sie entweder als Teil des ersten Signals oder als Echo angesehen werden. Die Lokalisierung in solchen Umgebungen ist schwierig, falls es sich bei den verwendeten Schallquellen um kontinuierliche Sounds handelt. Soll es möglich sein eine Schallquelle in einer reflektiven Umgebung zu lokalisieren, so sollten kurze, impulsartige Geräusche verwendet werden.

## 8.10 Interaktivität und Wiederholungen

Wie bereits erwähnt, spielt das Erinnerungsvermögen einer Umgebung eine wichtige Rolle für die Orientierung. Durch Interaktionen und Wiederholungen kann dieses trainiert und die kognitive Karte des Spieler ausgebaut werden. Nur durch eine gute kognitive Karte ist dem Spieler auch eine gute Orientierung gewährleistet. Da die Erinnerung mit der Zeit abnimmt, sollte der Aufbau der Umgebung nicht allzu kompliziert sein oder der Spieler immer wieder die Möglichkeit haben, sein Wissen über die Umgebung wiederholen und auffrischen zu können. Interaktionen tragen ebenfalls dazu bei sich an gewisse Orte der virtuellen Welt zu erinnern.

## 8.11 Verständlichkeit

Es sollte darauf geachtet werden, dass die VAE der Spielwelt nicht mit unnötigen Sounds gefüllt ist, die den Spieler eventuell sogar verwirren anstatt ihm bei der Orientierung zu helfen. Deswegen sollten vor allem alle sichtbaren und interaktiven Elemente der Spielwelt vertont werden. Hierfür sollten aus der realen Welt bekannte Geräusche gewählt werden. Geräusche, die harmonisch den Geräuschen der Atmo gleichen sollten ausgetauscht oder bearbeitet werden, da sie ansonsten als Teil der Atmo angesehen werden. Das gleiche Prinzip der Gruppierung gilt genauso für Schallquellen, die sich örtlich nahe nebeneinander befinden und ebenfalls harmonisch gleich strukturiert sind. Diese sollte man sinnvoll in der Umgebung verteilen oder unterscheidbar machen. Ansonsten kann es sein, dass sie als ein einzelner Sound mit einer einzigen Position wahrgenommen werde.

## 8.12 Statische und dynamische Schallquellen

Neben statischen Schallquellen sollten genauso dynamische Schallquellen existieren, die es dem Spieler erlauben auch ohne Kopfbewegung die Richtung eines Sounds zu lokalisieren. Des Weiteren tragen bewegliche, dynamische Schallquellen auch zu einer etwas lebhafteren VAE bei.

## 8.13 Bedeutung von Klangobjekte beachten

Werden Klangobjekte genutzt, sollte auf deren kulturelle Bedeutung geachtet werden. Je nach Situation oder Kultur kann ein und dasselbe Signal etwas anderes bedeuten. Soll Gefahr symbolisiert werden, können für die meisten westlichen Kulturen Sirenen oder beim Thema Tod Glocken verwendet werden. Leitmotive können gewissen Charakteren oder Situationen zugewiesen werden, die durch ihre stetige Wiederholung vom Spieler verstanden werden können. Selbst wenn der Charakter nicht in einer Situation bzw. Szene anwesend ist, so verbindet man das Leitmotiv dennoch mit ihm.

## 8.14 Zusätzliche Informationen über die Audioebene

Die Audioebene kann darüber hinaus genutzt werden, um zusätzliche Informationen über die Geschichte oder die Umgebung zu liefern. Gängige Methoden finden in Form von Gesprächen zwischen Charakteren oder auch Tonaufnahmen statt. Diese können in der Spielwelt verteilt sein und als optionale Hintergrundinformationen angesehen werden.

## 8.15 Wechselwirkung der visuellen und auditiven Ebene

Des Weiteren sollte auf die Verbindung zwischen visueller und akustischer Ebene geachtet werden. Da sich Menschen normalerweise mehr auf den Sehsinn als den Hörsinn verlassen, dominiert dieser und kann zu einer veränderten Wahrnehmung des Audios führen. Wie beim Bauchredner-Effekt zuvor in Abschnitt 6.3 erwähnt, sieht der Spieler ein Objekt als die Quelle eines Sounds an, selbst wenn es gar nicht die wahre Quelle dieses Sounds ist. Dieser Effekt kann entweder für die eigenen Zwecke genutzt oder zumindest bei der Konzeption beachtet werden.

# Kapitel 9

## Beispielanalyse

Im Folgenden wird versucht drei Beispiele anhand der soeben aufgestellten Leitlinien zu untersuchen. Nach einer Kurzbeschreibung des Spiels wird das Audio analysiert und im jeweiligen Fazit beschrieben, was gut und schlecht funktioniert hat. Bei den Beispielen handelt es sich um explorative, storybasierte VR-Spiele, die mit der HTC Vive und den dazugehörigen Bewegungscontrollern getestet wurden.

### 9.1 *The Gallery – Episode 1: Call of the Starseed*

#### 9.1.1 Kurzbeschreibung

*The Gallery – Episode 1: Call of the Starseed* [22] ist ein exploratives VR-Spiel, dessen Geschichte der Spieler durch die Perspektive von Alex erkunden kann. Alex versucht seine verschwundene Zwillingschwester Elsie wiederzufinden. Dabei muss der Spieler den Nachrichten von Elsie in Form von Tonbändern folgen. Der Spieler trifft auf den merkwürdigen Professor mit dem Elsie zusammengearbeitet hat und der Alex nach dem Verschwinden seiner Schwester nun für seine eigenen Zwecke einsetzen möchte. Die Suche führt Alex zu einer geheimnisvollen Maschine, die mit dem Verschwinden von Elsie im Zusammenhang steht.

#### 9.1.2 Analyse

Das Spiel ist für die HTC Vive, Oculus Rift, sowie Windows Mixed Reality verfügbar und wurde für die Verwendung von Bewegungscontrollern optimiert [35]. Es ist möglich das Spiel raumfüllend oder auch stehend zu spielen und ist für Einzelspieler gedacht. Bewegen kann der Spieler sich dank Teleportation und falls Room Scale verwendet wird, auch durch Gehen. Das Inventar besteht aus einem Rucksack, der verwendet werden kann, indem der Spieler hinter den Rücken greifen und ihn vor sich in der Luft loslässt. Darin befinden sich alle gefundenen Kassetten und Gegenstände. Das VR-Spiel besteht aus mehreren Abschnitten, die durch Ladebildschirme voneinander getrennt sind.

Zu Beginn des Spiels befindet sich der Spieler in einer Höhle nachts am Strand. Dort sind bereits zu Beginn das Rauschen von Wellen und das Zirpen von Grillen zu hören. Durch die verwendeten Ambience Geräusche, lässt sich selbst ohne das Meer gesehen zu haben und während der Spieler sich noch in der Höhle befindet sagen, dass es Nacht



**Abbildung 9.1:** Der Spieler findet sich zu Beginn des Levels am Strand wieder. Dort fungiert die Glocke der Boje im Meer als Landmark und später auch als Teil eines Rätsels [22].

ist und er sich in der Nähe eines Meeres befindet. Des Weiteren ist das Krächzen von Möwen zu hören. Sobald der Spieler die Höhle verlässt, setzt etwas Musik ein. Diese unterstreicht die mysteriöse Situation, in der sich der Spieler befindet und den Blick auf das Schiffswrack, sowie den Leuchtturm. Der Strand-Abschnitt verwendet das Rauschen des Meeres, das sich nur auf einer Seite des Levels befindet und das Glockenläuten der Boje als Landmarks. Die Glocke der Boje wird später sogar Teil einer Aufgabe und ist in Abbildung 9.1 zu sehen. Diese Landmarks sind auch noch in der bewohnten Höhle etwas weiter weg gut zu hören. Somit weiß der Spieler, wo sich der zuvor untersuchte Strand befindet. Neben diesen großen Orientierungspunkten, wird auch noch ein Radio verwendet, das in einem kleineren Bereich ebenso zu hören ist. Auch er dient in einem bestimmten Radius als Orientierungspunkt. Hat der Spieler das Rätsel mit den Glocken geschafft, setzt Regen samt Gewitter ein, das ebenfalls vertont wird.

Im zweiten Abschnitt befindet sich der Spieler in einem Abwasserkanal, wie in Abbildung 9.2 zu sehen ist. Dort sind zu Beginn immer noch die Geräusche des Meeres und der Möwen zu hören, allerdings gedämpfter als zuvor. Je weiter der Spieler in den Abwasserkanal vordringt, desto leiser werden die Ambience Sounds des vorherigen Abschnitts. Die Atmo des neuen Abschnitts besteht aus Wasserplätschern der Metallrohre und dem Windrauschen, das durch einige Löcher in der Decke zu hören ist. Ebenso zu hören ist ein tiefes Brummen, das vermutlich das Rauschen des alten Rohrsystems betonen soll. Die ganze Zeit über begleitet Musik das Erkunden des Abwasserkanals. Kurz vor Ende dieses Abschnitts ist das Summen von Fliegen zu hören, die über Müll, das sich im Wasser neben dem Spieler befindet, kreisen. Im Abschnitt des Abwasserkanals gibt es keinen großen Orientierungspunkt, da der Spieler seinen Weg durch das Labyrinth anhand der gefundenen Karte und der Pfeile an den Wänden finden soll. Dennoch werden auch hier kleinere Landmarks gesetzt, darunter brennende Tonnen und tropfende Metallrohre. Diese ermöglichen es dem Spieler zu erkennen, ob er sich an dieser Stelle im Labyrinth bereits befunden hat.



**Abbildung 9.2:** Innerhalb des Abwasserkanals muss der Spieler anhand einer Karte und den Wandmarkierungen den Ausweg aus diesem Labyrinth finden. Kleinere Landmarks geben ihm Hinweise, ob er sich an dieser Stelle bereits befunden hat [22].



**Abbildung 9.3:** Im Versteck des Professors müssen einige Rätsel gelöst werden während seine Stimme als Landmark und Hilfestellung zugleich funktioniert [22].

Der nächste Abschnitt zeigt das Versteck des Professors, das in Abbildung 9.3 dargestellt ist. Hier sind piepende und metallene Geräusche, wie bei technischen Apparaturen zu hören. Ebenso wird dieses Level von Musik untermalt. Da der Spieler sich immer noch in dem Abwasserkanal befindet, bleibt das Wasserplätschern aus dem vorherigen Abschnitt bestehen. Das Abwasser spielt für sein Versteck eine wichtige Rolle, da der Professor verschiedene Mechanismen entwickelt hat, die mit Wasserkraft angetrieben werden. Im Versteck des Professors, agiert vor allem er selbst bzw. seine Stimme als Landmark. Hat der Spieler die Sicherung eingesetzt oder das Versteck betreten, ändert sich die Position des Professors immer etwas. Allerdings bleibt er für die Dauer der Aufgabe immer an der gleichen Stelle und ändert erst mit Beginn der nächsten Aufgabe



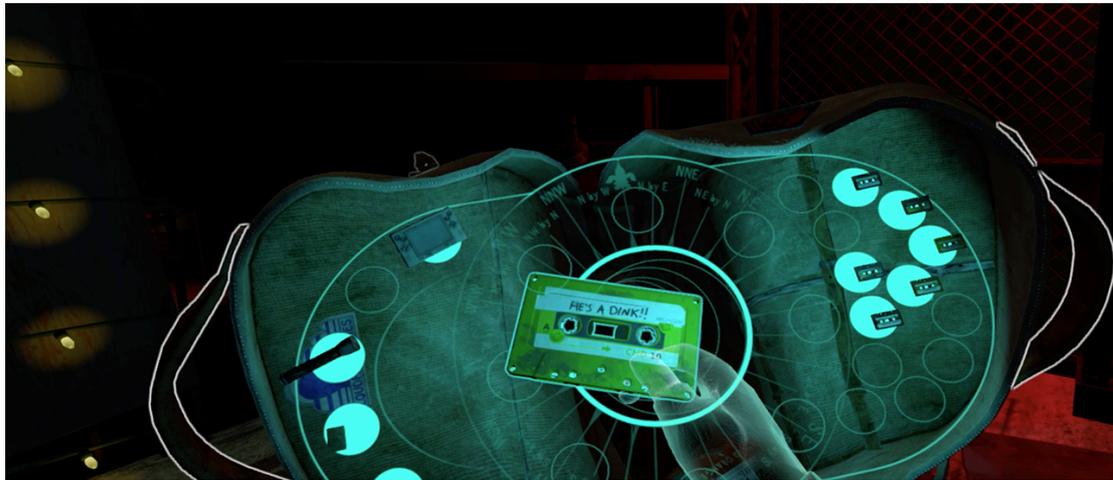
**Abbildung 9.4:** Der Labor-Abschnitt wird stetig von klassischer Musik und elektromagnetischem Summen begleitet [22].

seine Position. So sind der Bereich der nächsten Aufgabe und sein Aufenthaltsort klar.

Im darauffolgenden Abschnitt wird man durch den Lift des Professors in ein Labor mit elektromagnetischem Schild geschickt, welches in Abbildung 9.4 zu sehen ist. Die Ambient Sound in diesem Level setzen sich aus einem leichten elektromagnetischen Summen des Schildes und dem Klappern der umherfliegenden Gegenstände zusammen. Des Weiteren läuft Johann Strauss' Walzer „An der schönen, blauen Donau“ die ganze Zeit über die Lautsprecher im Hintergrund. Sobald der Spieler den Kasten mit dem Kern des Elektromagneten geöffnet hat, endet die Musik, indem sie langsamer und tiefer wird, so als wäre sie defekt, während das elektromagnetische Summen weiter zu hören ist. Der elektromagnetische Bereich mit dem versperrten Artefakt besitzt keinen wirklichen Orientierungspunkt. Dieser ist jedoch in diesem Abschnitt auch nicht nötig, da der zu untersuchende Spielbereich kleiner als in den anderen Leveln ist. Der Spieler steht nur vor einer Maschine und versucht die verschlossene Kammer mit dem Gerät zu öffnen.

Daraufhin befindet sich der Spieler im letzten Abschnitt in einer großen Höhle und einer außerirdisch wirkenden Maschine. Hierbei ist ein dumpfes Dröhnen zu hören, dessen Hall durch die große Höhle zu hören ist. Immer wieder sind künstlich wirkende Geräusche zu hören, die das Alienhafte der Umgebung unterstreichen. Bewegt sich die Plattform langsam auf die Apparatur zu, ist hauptsächlich Musik samt Chor zu hören. In der Maschine selbst enden Atmo und Musik. Erst als sich der Spieler in einer runden Kammer mit merkwürdigen Symbolen befindet setzt erneut ein elektromagnetisches Dröhnen und etwas Musik ein. In diesen letzten beiden Abschnitten kann sich der Spieler nicht mehr von der fixen Position der Plattform entfernen oder teleportieren. Dadurch existieren keine Orientierungspunkte mehr, da nur noch die Geschichte aufgelöst wird und der Spieler selbst nicht mehr navigieren kann.

Es gab eine Vielzahl an Interaktivitäten innerhalb des Spiels. Da jeder Bereich in sich geschlossen war, musste sich der Spieler die Positionen der einzelnen Hinweispunkte nur für die Dauer des Levels merken. Die meisten Umgebungen beinhalteten nur eine



**Abbildung 9.5:** Informationen zur Hintergrundgeschichte oder zu bevorstehenden Aufgaben wurden über Kassetten erzählt. Ebenso im Bild zu sehen ist das vertonte Inventar in Form eines Rucksacks [22].

oder zwei Aufgaben, dadurch fand kaum eine Wiederholung einer Interaktion innerhalb eines Levels statt. Dadurch stachen vor allem einmalige Besonderheiten, wie das Aufploppen lassen von Popcorn über dem Feuer hervor und blieben dem Spieler so besser in Erinnerung. Der Spieler wusste danach immer, dass er im Feuer am Strand neben dem Schiffswrack Popcorn machen konnte. Soweit beurteilbar, wurden alle sichtbaren und interaktiven Objekte vertont. Dies führte dazu, dass sobald man mit einem Gegenstand interagiert, dieser auch ein passendes Geräusch von sich gab. Diese wurden auch mit vertrauten Sounds aus der realen Welt vertont und trugen somit zur Immersion der virtuellen Welt bei. Des Weiteren wurde auch das Inventar, das in Form eines Rucksacks vorhanden ist vertont. Nimmt man ihn nach vorne, hört es sich so an als würde man den Rucksack mit einem Reißverschluss öffnen. Packt man Gegenstände hinein oder bewegt man ihn wieder hinter den Rücken, so klingt es, als würde der Reißverschluss erneut zugemacht werden. Dies signalisiert dem Spieler, ob ein Objekt verstaut wurde und ob sich der Rucksack wieder am Rücken befindet.

Informationen über die Hintergrundgeschichte wurden beinahe ausschließlich über die Audioebene gehandhabt. Hierbei werden Kassetten verwendet, in denen hauptsächlich Elsie spricht. In Abbildung 9.5 ist solch eine Kassette zu sehen, auf der sie ihre Reise dokumentiert und für ihren Bruder Hinweise hinterlässt, die ihm beim Weiterkommen in den Leveln helfen sollen. Unter anderem kann man auch Kassetten finden, auf denen der Professor etwas zur Entdeckung des Artefakts erzählt. Die Stimme des Professors besitzt die gleiche Funktion wie die Kassetten. Auch er gibt dem Spieler Anweisungen bzw. Hinweise, was als nächstes zu tun ist.

Die Teleportation, hier „Blink“ genannt, wurde ebenfalls vertont. Hierfür werden während eines Blinks Schritte abgespielt. Je nach Untergrund und Level kann sich diese Vertonung auch ändern. Während am Strand Schritte auf festem Untergrund zu hören waren, wurde beispielsweise für den Abwasserkanal Wassergeräusche bei der Fortbewegung gewählt. Diese Anpassung des Audios an die Art der Umgebung führt zu einer

besseren Präsenz und Immersion, da Schritte auf festem Untergrund kaum zum Durchwaten einer hüfthohen Wasserumgebung gepasst hätte.

### 9.1.3 Fazit

Besonders hilfreich sind die akustischen Orientierungspunkte und Kassetten-Hinweise, da sie der Wegfindung am besten dienen. Durch die Stimme des Professors lässt sich etwas über die Handlung, aber auch über die momentanen Aufgaben herausfinden. Die sich verändernde Atmo und Teleportationsgeräusche verbessern die Präsenz durch die Anpassung an die unterschiedlichen Umgebungen. Es existieren keine wirklichen zusätzlichen Führungspunkte im Spiel, da dem Spieler solange Zeit gegeben wird bis er den nächsten Hinweis von selbst gefunden hat. Dies kann manchmal frustrierend sein. Vor allem, wenn man das Rätsel mit den drei Glocken nicht sofort versteht und lange am Strand nach einem weiteren Hinweis sucht. An manchen Stellen, wie dem Rätsel mit den drei Glocken wäre es sicher angebracht, akustische Führungspunkte einzusetzen. Diese akustischen Hilfestellungen könnten nach einer gewissen Zeit alle drei Glocken erklingen lassen, damit der Spieler versteht, dass mit „3B“ die drei Glocken gemeint sind. Auch im Labyrinth wäre vielleicht ein akustischer Orientierungspunkt am Ziel nicht verkehrt gewesen. Der Spieler bekommt zwar eine visuelle Karte, dennoch ist diese nutzlos, da mehrere Personen darauf Notizen hinterlassen haben und die Qualität der Textur nicht ausreicht, um sie vollends entziffern zu können. Auch das Rätsel am Ende, mit der das elektromagnetische Feld ausgeschaltet werden konnte war sehr frustrierend, da man nur zum ersten Wert einen Hinweis bekam. Vielleicht hätte man auch zu den anderen drei Werten Hinweise in der Umgebung verstecken können, die nach einer gewissen Zeit akustische hervorgehoben werden. An anderen Stellen jedoch hat die Wegfindung auch ohne zusätzliche Führungspunkte gut funktioniert. Da es viele verschiedene Umgebungen mit unterschiedlichem Aufbau gab, waren die getrennten Abschnitte hilfreich. Dadurch musste der Aufbau eines Levels nur für die Dauer der Aufgaben in diesem Bereich gemerkt werden und konnten danach wieder verworfen und für den neuen Ort angelegt werden.

## 9.2 *Elena*

### 9.2.1 Kurzbeschreibung

*Elena* [21] ist ein exploratives VR-Spiel, das die Geschichte eines Paares aus der Sicht von Anna zeigt. Anna muss im Laufe des Spiels herausfinden, was mit ihrem Ehemann Lennard passiert ist, der plötzlich verschwunden ist. Durch einen unbekanntem Anrufer und dem Durchsuchen des Hauses versucht sie eine Erklärung für das Verschwinden ihres Mannes und die Identität von Elena zu finden.

### 9.2.2 Analyse

Das Spiel ist für die HTC Vive, Oculus Rift, sowie als normale Desktop-Anwendung verfügbar [34]. Es ist möglich das Spiel sitzend, raumfüllend oder auch stehend zu spielen und ist für Einzelspieler gedacht. Bewegen kann der Spieler sich dank Teleportation und falls Room Scale verwendet wird, auch durch Gehen. Das Spiel gliedert sich in



**Abbildung 9.6:** Die zu untersuchenden Räume erstrahlen mit weißem Licht und werden durch einen Sound unterstützt, um die Aufmerksamkeit des Spielers auf sich zu ziehen [21].

mehrere Abschnitte und Umgebungen. In jedem Abschnitt wird einer der Räume der Wohnung untersucht und es können Hinweiskarten gefunden werden. Diese Hinweiskarten müssen an der Pinnwand im Vorraum angebracht werden und können nach und nach miteinander verbunden werden.

Im Spiel werden für jede Aktion Führungspunkte gesetzt, da der Spieler ansonsten nicht so einfach wüsste, welchen der Räume er nun untersuchen muss. Wie in Abbildung 9.6 zu sehen ist, leuchtet zu Beginn jeder Aufgabe der zu untersuchende Raum weiß und ein lauter Whoosh-Sound samt elektrischem Knistern ist zu hören. Dies soll den Spieler darauf aufmerksam machen, dass ein neuer Raum zur Untersuchung bereit steht. Da das weiße Licht als Hinweis nicht ausreicht, aktiviert der Sound die Aufmerksamkeit des Spielers und er blickt in die gewünschte Richtung. Vor allem das Klingeln des Telefons zieht durch seine Signalwirkung die Aufmerksamkeit des Spielers auf sich. Durch die zusätzlich gelernte Bedeutung des Handlungshinweises eilt der Spieler meist direkt zum Telefon und unterbricht seine aktuelle Handlung. Durch die gelernte Bedeutung des Whoosh-Sounds bei den weiß erleuchteten Räumen besitzt auch dieser eine Signalwirkung, die den Spieler aufhorchen lässt. Die ganze Zeit über gibt es keinerlei Ambience Sounds in der Wohnung, sondern nur Musik. Gleich beim Betreten des Levels setzt Musik ein. Sobald Krach aus dem Bad zu hören ist, ändert sich die Musik und es sind nur noch Klaviertöne zu hören. Ebenso kann man das Tropfen von Wasser wahrnehmen, dass aus dem Bad zu kommen scheint. Die Klaviertöne werden düster und erschreckend als ein Vogel auf den Spieler zufliegt. Sobald man den Telefonhörer in die Hand nimmt, beginnt erneut Musik zu spielen. Dieses Musikstück spielt durchgehend bis zu der Aufnahme mit dem Streitgespräch zwischen Elena und Lennard. Hierbei ist alles still und nur die Stimmen sind zu hören. Während dem Telefonat setzt erneut Musik ein, dieses Mal jedoch mit mysteriöser und unheilvoller Stimmung, allerdings besitzt sie immer noch eine sehr ähnliche Melodie. Dieses Musikstück endet sobald man die Simulation und Kapsel verlässt und sich in der realen Welt befindet. Danach sind weder Ambience



**Abbildung 9.7:** Die Stimmen aller Spielfiguren sind während Unterhaltungen zu hören, wodurch die Handlung der Geschichte und Hinweise erläutert werden. Diese Gespräche kommen in Form von Telefonaten oder Videoanrufen vor [21].

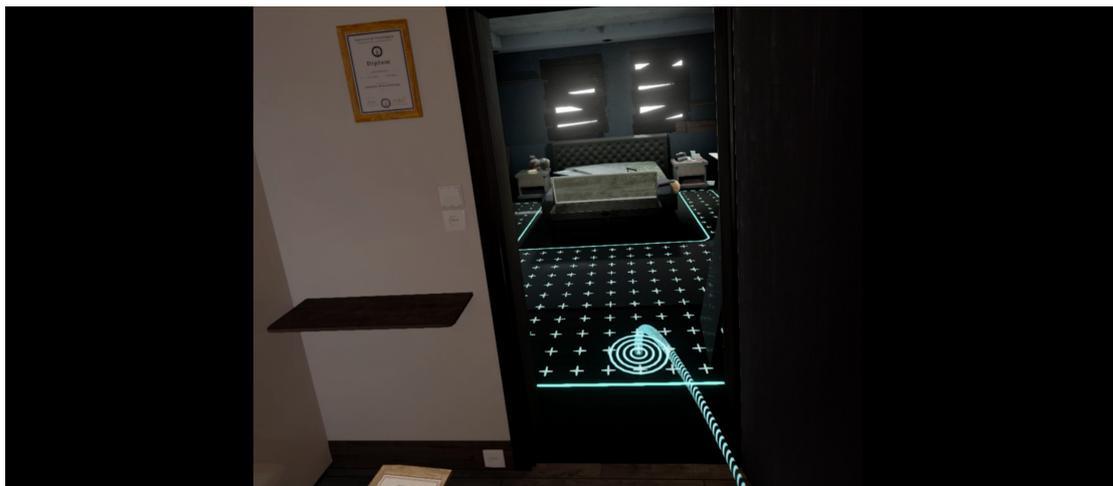
Sounds noch Musik zu hören. Erst ab dem Moment wieder, indem Lennard alles erklärt und seine Taten über einen Videoanruf offenlegt, der in Abbildung 9.7 dargestellt wird.

Es gibt nur vereinzelt besondere Interaktivitäten, wie den Regenstab oder die Spieluhr. Diese Objekte und deren Position behält sich der Spieler im Gedächtnis, auch wenn diese für das Vorankommen der Geschichte nicht relevant sind. Da jeder Raum nur einmal untersucht werden muss, braucht der Spieler sich den Aufbau der Räume auch nur für den Moment der Interaktion selbst zu merken. Dadurch findet auch keinerlei Wiederholung bestimmter Interaktionen in einem Raum statt, da keine genaue kognitive Karte für längere Zeit angelegt werden muss. Vor allem durch Stimmen werden Informationen über die Geschichte, aber auch Hinweise zum Lösen der nächsten Aufgabe geliefert. Annas Stimme kommentiert einige Gegenstände in der Wohnung, wodurch der Spieler mehr über das Leben von Anna und Lennard herausfinden kann. Zusätzlich kommentiert Anna die gefundenen Hinweiskarten und wenn sie an der Pinnwand richtig verbunden werden, wie in Abbildung 9.8 exemplarisch zu sehen. Im Gespräch mit dem unbekanntem Mann werden Hinweise zum nächsten Raum oder auch zur Geschichte selbst geliefert. Am Ende ist es auch ein Gespräch mit Lennard, dass die Auflösung der Geschichte verrät.

Es wurden nicht alle Elemente vertont, was vor allem bei der Interaktion mit einigen Objekten negativ auffällt. Die Geräusche der Schubladen oder Kleiderschranktüren passen beispielsweise nicht zur Interaktion mit diesen, da der Sound auch weiterspielte, obwohl sich die Schublade oder die Schranktür gar nicht mehr bewegten. Des Weiteren wurden interaktive Objekte, wie die Truhe im Schlafzimmer oder auch der Kühlschrank nicht vertont. Dies widerspricht der Wahrnehmung des Spielers und kann zu einer Störung der Präsenz führen. Die Teleportation wurde ebenfalls vertont. Hierfür ist die Teleportationsfläche beim Benutzen des zuständigen Knopfes sichtbar, wie in Abbildung 9.9 gezeigt wird. Befindet sich der Endpunkt des Teleportationsstrahls innerhalb dieser Fläche ist ein künstlicher, zischender Sound zu hören. Trifft der Endpunkt außer-



**Abbildung 9.8:** Anna kommentiert jede gefundene und richtig verbundene Hinweiskarte. Dadurch wird dem Spieler auch der nächste zu suchende Hinweis erklärt, der erneut mit denen auf der Pinnwand verbunden werden soll [21].



**Abbildung 9.9:** Der Sound der Teleportation ändert sich je nachdem, ob der Endpunkt sich innerhalb oder außerhalb der Teleportationsfläche befindet [21].

halb dieser Fläche auf, klingt der Sounds viel langsamer und auch tiefer. Dies signalisiert dem Spieler, ob er sich innerhalb der Teleportationsfläche befindet.

### 9.2.3 Fazit

Da es sich in der Wohnung die ganze Zeit über um eine Simulation gehandelt hat, wurde vermutlich auf eine Atmo verzichtet. Dadurch ist nicht festzustellen, ob sich die Wohnung in einer Stadt oder am Land befindet. Auch keine Geräusche, wie das Surren des Kühlschranks oder andere atmosphärisch Sounds wurden verwendet. Da der Spielbereich der Wohnung überschaubar ist, wurde vermutlich auf den Einsatz von akustischen

Orientierungspunkten verzichtet. Nichtsdestotrotz hätten kleinere Landmarks, wie der surrende Kühlschrank oder das Erklingen der Spieluhr zur schnelleren Orientierung innerhalb der Wohnung beigetragen. Die Signalwirkung des Telefonklingelns hingegen eignet sich hervorragend um die Aufmerksamkeit des Spielers auf sich zu ziehen. Durch die Klangcharakteristik des Klingelns ist dieses von anderen Schallquellen und der Musik unterscheidbar und von überall aus in der Wohnung zu hören. Die Vertonung der Teleportation sticht sehr aus der Soundkulisse heraus und hätte gegebenenfalls dezenter gestaltet werden können. Die Stimmen der Charaktere geben dem Spieler zusätzliche Informationen und leiten ihn bei den Aufgaben an, ohne das es zu gezwungen wirkt. Manchmal ist Annas Stimme etwas anstrengend, da sie beinahe pausenlos Objekte kommentiert oder Gespräche führt. Etwas weniger Kommentare hätten der Geschichte oder dem Spiel sicher nicht geschadet, da durch die Hinweiskarten der Pinnwand bereits sehr viel erklärt wurde, dass sie oft einfach nur wiederholt hat. Dennoch funktioniert die Orientierung im Spiel gut, vor allem durch die auffälligen Sounds während der Weißblenden. Diese verdeutlichen, wo der nächste Hinweis gefunden werden kann. Das einfache Abspielen von Tonspuren ohne Berücksichtigung der Interaktion selbst stört sehr häufig. Ein Beispiel hierfür wären Schubladen, da der Sound oft auch weiterspielt obwohl der Spieler gar nicht mehr mit ihnen interagiert. Um die Immersion zu steigern, sollten einige der interaktiven und sichtbaren Objekte vertont werden.

### 9.3 *Letzte Worte*

#### 9.3.1 Kurzbeschreibung

Bei *Letzte Worte* [24] handelt es sich um das Masterprojekt in Form eines VR-Spiels von Matthias Patscheider, Bianca Zankl und mir, das neben dieser Arbeit entstanden ist. Allerdings befindet sich das Projekt zum jetzigen Zeitpunkt noch in der Produktionsphase. Der Spieler übernimmt die Rolle von Lisa Wagner, die durch Schuldgefühle und den Selbstmord ihres Bruders Leo geplagt wird. Durch diese verwirrenden Gefühle beginnt sie, ihre verdrängten Kindheitserinnerungen zu erforschen und erfährt nach und nach von den dunklen Erinnerungen, die ihren Bruder und sie seither belasten.

#### 9.3.2 Analyse

Das Spiel ist für die HTC Vive und Oculus Rift verfügbar. Es ist möglich das Spiel raumfüllend oder auch stehend zu spielen und ist für einen einzelnen Spieler gedacht. Durch Teleportation und Gehen bei der Verwendung von Room Scale kann der Spieler sich in der virtuellen Welt fortbewegen. Wir haben uns gegen eine Vertonung der Teleportation entschieden, da der Fortbewegung keine zu große Bedeutung beigemessen werden sollte.

Wie in Abbildung 9.10 zu sehen, startet der Spieler in der abstrakten Traumwelt, in der das ganze Level über ein tiefer Basston zu hören ist. Bei der Berührung der Lichtkugel ertönen Klaviertöne. Sobald die Lichtgestalten die Erinnerungen der beiden Kinder in Form von Gesprächen zeigen, werden diese durch ein Musikstück aus Klavier und Cello unterstützt. Die Lichtkugel besitzt einen Glitzer-Sound, der dem Spieler beim Finden der Lichtkugel helfen soll. Dreht der Spieler also der Lichtkugel den Rücken zu



**Abbildung 9.10:** Die Lichtgestalten führen Gespräche aus der Vergangenheit, während die Lichtkugel einen Eigensound zur besseren Orientierung besitzt [24].



**Abbildung 9.11:** Lisas Wohnung besitzt mehrere interaktive Objekte die vertont wurden. Neben einem Schallplattenspieler kann der Spieler auch Tassen zerbrechen und Objekte fallen lassen [24].

oder fliegt sie zu schnell weg, kann er sie aufgrund ihres Eigensounds lokalisieren.

Danach befindet der Spieler sich in Lisas Wohnung, die in Abbildung 9.11 zu sehen ist. In Lisas Wohnung ist Straßenlärm von draußen zu hören. Dieser setzt sich aus vorbeifahrenden Autos und Bussen zusammen. Einmalig ist auch das Martinshorn eines Krankenwagen zu hören. Diese Geräusche sollen zeigen, dass sich die Wohnung in einer belebten Stadt befindet. Da Lisas Wohnung, das Baumhaus und Leos Studentenzimmer im Grunde nur aus einem kleinen übersichtlichen Zimmer bestehen, wurde hier auf einen akustischen Orientierungspunkt verzichtet. Um den Moment vor dem Selbstmord des Bruders festzuhalten, wurde ein Telefonat der Geschwister gewählt. Das Klingeln des Telefons zieht durch seine Signalwirkung die Aufmerksamkeit des Spielers auf sich.

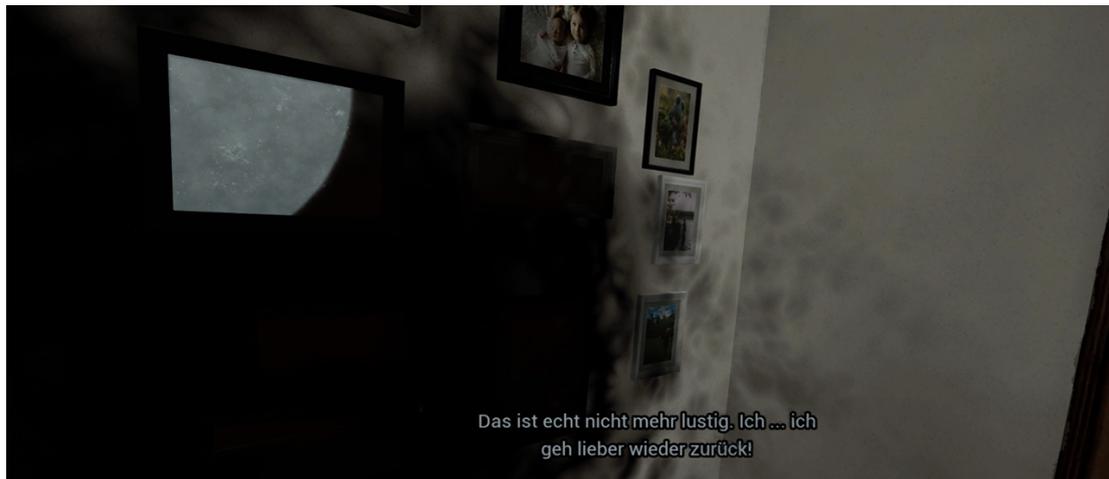


**Abbildung 9.12:** Das Windrauschen der Berginsel dient als Landmark und ist von den beiden anderen Inseln aus ebenfalls zu hören. Je nach Höhe verändert sich der Sound [24].

In Lisas Wohnung befindet sich ein Schallplattenspieler, den der Spieler nach Belieben einschalten kann.

In der Nebelwelt ist vor allem der Wind zu hören. Je höher der Spieler kommt, desto lauter und höher wird das Windrauschen. Um die Höheninformation des Windrauschens am Berg besser einschätzen zu können, wurde die Lautstärke verändert. Ebenso kommt ein zusätzliches Windpfeifen hinzu, sobald der Spieler eine gewisse Höhe erreicht hat. Diese Informationen sollen dem Spieler helfen sich auch auf der vertikalen Ebene zurecht zu finden. Befindet sich der Spieler in einem steinernen Durchgang, so wird der Wind leiser und das Windpfeifen ist lauter zu hören. Die Lautstärke des Windes hängt von der Höhe, aber auch von der Entfernung zum Berg ab. Besichtigt der Spieler gerade die Elternhaus-Insel, so ist der Wind kaum wahrnehmbar. Darüber hinaus dient der Wind als Orientierungspunkt. Je nach Höhe oder Entfernung zur Spitze verhält sich die Schallquelle anders. Befindet man sich vor dem Elternhaus kann man erraten, wo sich der Berg ungefähr befindet.

Das Elternhaus besitzt zufällig auftretendes Holzknarren, was das Alter des Bauernhauses hervorheben soll. Ansonsten sind Geräusche, wie das Surren des Kühlschranks oder das Knistern der Flammen im Bauernofen zu hören. Diese Geräusche werden lauter, je näher der Spieler dem jeweiligen Objekt kommt. Im Elternhaus selbst tragen der Kühlschrank und der Bauernofen nicht nur zur Atmo bei, sondern sie dienen dem Spieler auch als kleinere Landmarks. Um Informationen über die Kindheit liefern zu können, entschieden wir uns Kassetten zu verwenden. Diese kommen im Elternhaus und Baumhaus vor. Hierbei war der Gedanke, dass Leo ein Aufnahmegerät besaß, mit dem er alltägliche Situationen mit seiner Schwester aufnahm. So ist es dem Spieler möglich etwas über die Beziehung und Kindheit der beiden zu erfahren. Auf der Insel des Baumhauses ist bereits von unten das Rauschen der Blätter zu hören. Betritt der Spieler oben das Baumhaus, so werden das Rauschen der Blätter und der Wind von außen abgedämpft. Im Inneren ist dann neben den dumpfen Geräuschen von außen noch das Knarren des Holzbodens zu hören. Beide Inseln besitzen ein Objekt, das vom Spieler



**Abbildung 9.13:** Sobald das Schlüsselement des Levels gefunden und an der richtigen Stelle angebracht wurde, ertönt das Leitmotiv des Onkels und Lisa kommentiert den Abschluss des Bereichs [24].

gefunden oder zusammengebaut werden muss, damit er das Level abgeschlossen hat und weiterkommt. Dies wird durch die visuelle Veränderung, aber auch durch einen abstrakten Sound angezeigt. Dieser Sound ist das Leitmotiv des Onkels und tritt immer wieder an Stellen auf, die mit ihm in Verbindung stehen. Wie in Abbildung 9.13 gezeigt wird, wird dieser Effekt durch Lisas Kommentar begleitet, dass dieses Level beendet ist und der Spieler sich einem neuen Ort zuwenden kann.

Erneut in der Traumwelt hat sich die Atmo bereits geändert. Nun sind ein geisterhafter hoher Ton und ab der ersten Berührung der Lichtkugel auch ein Herzschlag zu hören. Der Herzschlag wird mit jeder Berührung schneller und höher, bis der Spieler schließlich die Lichtfigur vor der Tür berührt. Danach ändert sich das Herzklopfen und die Atmo nimmt den tiefen Basston vom ersten Level erneut an. Des Weiteren besitzt die Lichtkugel erneut einen Eigensound, der genau wie im ersten Level bei der Lokalisierung helfen kann.

In Leos Studentenzimmer herrscht erneut eine realere Atmo. Hier sind Straßengeräusche und Vogelzwitschern zu hören. Diese dezenten Straßengeräusche und das Vogelzwitschern sollen einen etwas ruhigeren Ort suggerieren, als an dem Lisa lebt. Wie in Abbildung 9.14 dargestellt, handelt es sich bei der letzten Kassette um den Abschiedsbrief von Leo, den er seiner Schwester Lisa widmet. Darin erklärt er ihr und dem Spieler, dass es nicht ihre Schuld gewesen sei und er immer froh war, sie gehabt zu haben.

Weiß der Spieler nicht mehr, wo er sich noch umschauen soll oder wo er als nächstes suchen soll, gibt es Führungspunkte in Form von Lisas Stimme. Sie sagt dem Spieler nach einer gewissen Zeit oder direkt, ob sie einen bestimmten Ort verlassen möchte oder gibt Hinweise zu den gesuchten Gegenständen. Berührt der Spieler beispielsweise die Haustür des Elternhauses, so gibt sie ihm den Hinweis, dass es einen Schlüssel in der Nähe geben müsste. Findet er diesen nicht und verlässt die Insel, gibt sie ihm beim erneuten Betreten den Hinweis, dass sich der Schlüssel doch irgendwo in den Blumenkästen befunden hat. Kann der Spieler ein Level durch eine Tür, wie zum Beispiel in Lisas Woh-



**Abbildung 9.14:** Leos Abschiedsbrief kann wie die Kindheitserinnerungen auf einer Kassette gefunden werden [24].



**Abbildung 9.15:** Ist der Weg zum nächsten Level freigegeben, erleuchtet die Tür weiß und gibt ein glitzerndes Geräusch von sich, das auch von weiter weg zu hören ist [24].

nung verlassen, ist neben einem weißen Lichtstrahl auch ein helles glitzerndes Geräusch zu hören, der die Aufmerksamkeit des Spielers durch seinen ungewöhnlichen Sound auf sich zieht. Eine solche Tür ist in Abbildung 9.15 zu sehen. Es gibt viele Interaktionen, die der Spieler in den Umgebungen vornehmen kann. Er kann Schubladen, Schränke und Türen öffnen, einige Lampen und Herdplatten einschalten. Außerdem kann er die Schaukel schwingen lassen und einige Objekte zerbrechen lassen. Durch die Wiederholung der Kassetten, weiß der Spieler bereits, dass sich auch welche in einem anderen Level verstecken können. Durch die Interaktionen mit zum Beispiel den Schubladen erweitert sich die kognitive Karte im Kopf des Spielers stetig. Besondere Interaktionen, wie das Abspielen einer Schallplatte oder das Ein- und Ausschalten der Herdplatten bleiben dem Spieler besonders in Erinnerung. Meist wird sich auch die Position dieser

Interaktionen gemerkt, da sie ungewöhnlich und einmalig waren. Darüber hinaus wurde versucht alle sichtbaren und interaktiven Elemente zu vertonen. Besonders viel Augenmerk wurde auf die Vertonung von Öffnen und Schließen der Schubladen, Schränken und Türen gesteckt. Hierfür wird je nach Bewegungsrichtung ein Öffnen oder Schließen Geräusch abgespielt und der Öffnungsgrad und die Geschwindigkeit der Interaktion bestimmen darüber hinaus, wie laut oder hoch sich ein Ton anhört. Bei Schaukeln oder Tischlampen wurde ein Quietschen ab dem Erreichen eines bestimmten Winkels eingefügt. Andernfalls würden die verwendeten Sounds nie zur Geschwindigkeit und Art der Bewegung passen. Werden Objekte fallen gelassen, geben sie je nach Größe ein dumpfes, lautes oder ein helles, leises Geräusch von sich. Einige leichte Objekte, wie Zettel bekommen keinen Eigensound. Tassen oder andere zerbrechliche Objekte zerspringen beim Fallen lassen und geben beim Aufkommen auf den Boden ein passendes Geräusch von sich.

### 9.3.3 Fazit

Das Projekt befindet sich derzeit immer noch in der Produktionsphase. Dennoch wurden bereits alle Audiodateien ins Spiel eingebaut. Bei den Atmos wurde versucht sie möglichst für jeden Ort passend zu gestalten. Ebenfalls wurden alle Elemente, die sichtbar oder interaktiv sind vertont. In den Playtests<sup>1</sup> stellte sich heraus, dass einige der Hinweis-Kommentare von Lisa sehr hilfreich waren und den Spielern das Gefühl gaben, dass sie mit einer Aufgabe fertig waren oder wüssten, was sie nun tun sollten. Einigen wenigen reichte dieser Audiokommentar jedoch nicht und sie hätten sich mehr Feedback bei Beendigung eines Levels gewünscht. Des Weiteren wurde versucht die kommentierende Stimme von Lisa so einzusetzen, dass sie als nicht zu penetrant empfunden wurde. Im Unterschied zu vielen anderen VR-Spielen haben wir sehr auf die Geräusche von Schubladen und Türen geachtet. Es werden die Geschwindigkeit und Position der Tür beachtet und je nachdem der Sound abgespielt. Schlägt der Spieler die Tür schnell zu, ist das Quietschen höher und lauter. Öffnet der Spieler allerdings die Tür langsam, so ist das Quietschen der Tür leise und tiefer. Viele andere VR-Spiele, wie zum Beispiel *Elena* aktivieren ein Geräusch, egal wie schnell der Spieler die Tür auch öffnet oder schließt. In den Playtests wurde der Anruf in Lisas Wohnung stets angenommen, auch wenn die Möglichkeit bestand den Anruf nicht anzunehmen und stattdessen den Anrufbeantworter abzuhören. Dies liegt vermutlich an der hohen Signalfunktion des Telefonklingelns.

---

<sup>1</sup>Ein Playtest bezeichnet die Testung eines Spiels durch Probanden, um Feedback zu gewissen Aspekten des Spiels erhalten zu können.

# Kapitel 10

## Fazit

### 10.1 Zusammenfassung

Durch die Analyse zweier Studien, die sich mit der auditiven Orientierung in Virtual Reality Anwendungen beschäftigen, sowie die im Zuge dieser Arbeit recherchierte Theorien zur menschlichen auditiven Wahrnehmung und der Funktionsweise von binauralem Audio, konnten Leitlinien zur Verbesserung der Orientierung durch Audio in virtuellen Welten aufgestellt werden. Diese Leitlinien stellen meine Antwort auf die zu Beginn dieser Arbeit gestellten Frage, wie man die Audioebene nutzen kann, um die Orientierung in komplexen virtuellen Welten zu steigern, dar. Dabei stellte sich als schwierigster Punkt die Neuheit moderner Virtual Reality Technologien heraus, wodurch es noch keine einheitlichen Richtlinien oder Aspekte gibt, die bei der Arbeit mit Audio in Virtual Reality beachtet werden können. Durch diese fehlenden allgemeinen Informationen wurde selbst versucht, den Zusammenhang von Audio und Orientierung durch bestehende Studien und Theorien herzustellen. Aufgrund dieser fehlenden Informationen konnte nur ein Teilbereich der möglichen zu beachtenden Aspekte aufgestellt werden und umfasst daher vermutlich nicht alle bestehenden Möglichkeiten der Orientierung durch Audio. Dabei wurde versucht durch die Leitlinien Anhaltspunkte bei der Erstellung eines Audiokonzepts zu geben, auf die während der Planung geachtet werden sollte.

### 10.2 Ausblick

Die Verwendung von Audio im Zusammenhang mit Virtual Reality bietet eine Vielzahl an Möglichkeiten, um in verschiedene Richtungen weiter recherchieren zu können. So bietet die zukünftige Entwicklung von Lautsprechern und Kopfhörern für 3D Audio einen wichtigen Ansatz, da sich nicht nur die Eingabegeräte, sondern auch die Wiedergabegeräte weiterentwickeln müssen. Es wird somit interessant sein zu sehen, welche technischen Möglichkeiten Lautsprecher für den Privatgebrauch in Zukunft schaffen können. Ein weiterer interessanter Bereich ist der Einsatz der eigenen Stimme im Zusammenspiel mit virtuellen Elementen und wie die Spielwelt auf diese reagiert. In einigen Jahren ist es eventuell möglich, dass der Spieler mit seiner eigenen Stimme Gespräche mit NPCs führt und diese sogar darauf antworten.

# Anhang A

## Inhalt der CD-ROM/DVD

Format:

### A.1 PDF-Dateien

Pfad: /

Povolny\_Samantha\_2018.pdf Masterarbeit

### A.2 Bildmaterial

Pfad: /images

\*.jpg, \*.png . . . . . Verwendetes Bildmaterial der Masterarbeit

### A.3 Online-Ressourcen

Pfad: /onlinequellen

\*.pdf . . . . . Archivierte Onlinequellen

# Quellenverzeichnis

## Literatur

- [1] Durand R. Begault. *3-D Sound for Virtual Reality and Multimedia*. Kalifornien: The NASA STI Program Office, 2000 (siehe S. 14–16).
- [2] Karen Collins. *Game Sound. An Introduction to the History, Theory, and Practice of Video Game Music and Sound Design*. Cambridge: The MIT Press, 2008 (siehe S. 29, 30).
- [3] Ralf Dörner und Frank Steinicke. „Wahrnehmungsaspekte von VR“. In: *Virtual and Augmented Reality (VR/AR): Grundlagen und Methoden der Virtuellen und Augmentierten Realität*. Hrsg. von Ralf Dörner u. a. Berlin: Springer Verlag, 2013. Kap. 2, S. 33–64 (siehe S. 7, 8, 25).
- [4] Ralf Dörner u. a. „Einleitung“. In: *Virtual and Augmented Reality (VR/AR): Grundlagen und Methoden der Virtuellen und Augmentierten Realität*. Hrsg. von Ralf Dörner u. a. Berlin: Springer Verlag, 2013. Kap. 1, S. 1–32 (siehe S. 7, 8).
- [5] Ralf Dörner u. a. „Interaktionen in Virtuellen Welten“. In: *Virtual and Augmented Reality (VR/AR): Grundlagen und Methoden der Virtuellen und Augmentierten Realität*. Hrsg. von Ralf Dörner u. a. Berlin: Springer Verlag, 2013. Kap. 6, S. 157–194 (siehe S. 24, 25).
- [6] Barbara Flückiger. *Sound Design. Die virtuelle Klangwelt des Films*. 3. Aufl. Marburg: Schüren Verlag, 2007 (siehe S. 28, 32, 33).
- [7] E. Bruce Goldstein. *Wahrnehmungspsychologie. Der Grundkurs*. Hrsg. von Karl R. Gegenfurtner. 9. Aufl. Berlin: Springer Verlag, 2015 (siehe S. 4–6, 20).
- [8] Thomas Görne. *Tontechnik. Hören, Schallwandler, Impulsantwort und Faltung, digitale Signale, Mehrkanaltechnik, tontechnische Praxis*. Hrsg. von Ulrich Schmidt. 4. Aufl. München: Carl Hanser Verlag, 2014 (siehe S. 16, 17).
- [9] Paul Grimm u. a. „VR-Ausgabegeräte“. In: *Virtual and Augmented Reality (VR/AR): Grundlagen und Methoden der Virtuellen und Augmentierten Realität*. Hrsg. von Ralf Dörner u. a. Berlin: Springer Verlag, 2013. Kap. 5, S. 127–156 (siehe S. 10).
- [10] Jarmo Hiipakka u. a. „Sound signal processing for a virtual room“. In: *10th European Signal Processing Conference (Tampere)*. IEEE, Sep. 2000, S. 1–4 (siehe S. 11, 12).

- [11] Tomasz R. Letowski und Szymon T. Letowski. *Auditory Spatial Perception: Auditory Localization*. Techn. Ber. ARL-TR-6016. Aberdeen Proving Ground, MD: U.S. Army Research Laboratory, Mai 2012. URL: [www.dtic.mil/dtic/tr/fulltext/u2/a562292.pdf](http://www.dtic.mil/dtic/tr/fulltext/u2/a562292.pdf) (siehe S. 25–28).
- [12] Tapio Lokki und Matti Gröhn. „Navigation with Auditory Cues in a Virtual Environment“. *IEEE MultiMedia* 12.2 (2005), S. 80–86 (siehe S. 19–21).
- [13] Kyla A. McMullen. „The Potentials for Spatial Audio to Convey Information in Virtual Environments“. In: *IEEE VR Workshop: Sonic Interaction in Virtual Environments (SIVE)* (Minneapolis). IEEE, 2014, S. 31–34 (siehe S. 18, 19, 22).
- [14] Kyla A. McMullen und Gregory H. Wakefield. „3D Sound Memory in Virtual Environments“. In: *IEEE Symposium on 3D User Interfaces* (Minnesota). IEEE, 2014, S. 99–102 (siehe S. 28).
- [15] Günther-Armin Neubauer. *Musik. Lexikon der Grundbegriffe*. Hamburg: Rowohlt, 1994 (siehe S. 33).
- [16] Francis Rumsey. *Spatial Audio*. Oxford: Focal Press, 2001 (siehe S. 3, 6, 10).
- [17] Paola Salomoni u. a. „Diegetic user interfaces for virtual environments with HMDs: a user experience study with oculus rift“. *J. Multimodal User Interfaces* 11.2017 (2005), S. 173–184 (siehe S. 7, 8, 28, 29).
- [18] Stefania Serafin und Giovanni Serafin. „Sound Design to Enhance Presence in Photorealistic Virtual Reality“. In: *Proceedings of ICAD 04-Tenth Meeting of the International Conference on Auditory Display* (Sydney). ICAD, Juni 2004, S. 1–4 (siehe S. 9, 10).
- [19] Da-Jhuang Su und Shih-Fu Hsieh. „Robust Crosstalk Cancellation for 3D Sound using Multiple Loudspeakers“. In: *IEEE International Conference on Acoustics, Speech and Signal Processing* (Honolulu). IEEE, 2007, S. 181–184 (siehe S. 15).
- [20] Yougen Yuan u. a. „Sound image externalization for headphone based real-time 3D audio“. *Frontiers of Computer Science* 11.3 (2017), S. 419–428 (siehe S. 17).

## Software

- [21] Catch A Cat. *Elena*. HTC Vive, Oculus Rift. 2016 (siehe S. 46–49).
- [22] Cloudhead Games. *The Gallery: Episode 1 - Call of the Starseed*. Playstation VR, HTC Vive, Oculus Rift, Windows Mixed Reality. 2016 (siehe S. 30, 31, 41–45).
- [23] Epic Games. *Unreal Engine 4*. Windows, Linux, macOS. 2014 (siehe S. 9, 34, 35).
- [24] Matthias Patscheider, Samantha Povolny und Bianca Zankl. *Letzte Worte*. HTC Vive, Oculus Rift. 2018 (siehe S. 2, 50–54).
- [25] Rocksteady Studios. *Batman: Arkham VR*. Playstation VR, HTC Vive, Oculus Rift. 2016 (siehe S. 29, 30).

## Online-Quellen

- [26] Felix Baarß. *FAQ: Was ist Phasenverschiebung?* Nov. 2012. URL: [www.delamar.de/faq/phasenverschiebung-17344/](http://www.delamar.de/faq/phasenverschiebung-17344/) (besucht am 17.08.2018) (siehe S. 15).
- [27] Duden. *binaural, seltener biaural*. URL: [www.duden.de/rechtschreibung/binaural](http://www.duden.de/rechtschreibung/binaural) (besucht am 26.08.2018) (siehe S. 3).
- [28] educalingo. *Wörterbuch*. URL: [educalingo.com/de/dic-de/diegesi](http://educalingo.com/de/dic-de/diegesi) (besucht am 25.08.2018) (siehe S. 28).
- [29] FMOD. *FMOD Studio*. URL: [www.fmod.com/studio](http://www.fmod.com/studio) (besucht am 25.08.2018) (siehe S. 35).
- [30] Epic Games. *Audio System Overview in Unreal Engine Documentation*. URL: [docs.unrealengine.com/en-US/Engine/Audio/Overview#soundcue](http://docs.unrealengine.com/en-US/Engine/Audio/Overview#soundcue) (besucht am 25.07.2018) (siehe S. 33–35).
- [31] HTC. *HTC Vive*. URL: [www.vive.com/de/product/](http://www.vive.com/de/product/) (besucht am 25.07.2018) (siehe S. 11).
- [32] Neumann. *Kunstkopf KU 100*. URL: [de-de.neumann.com/ku-100](http://de-de.neumann.com/ku-100) (besucht am 05.08.2018) (siehe S. 4).
- [33] Margaret Rouse. *VR room (virtual reality room)*. URL: [whatis.techtarget.com/definition/VR-room-virtual-reality-room](http://whatis.techtarget.com/definition/VR-room-virtual-reality-room) (besucht am 25.07.2018) (siehe S. 11).
- [34] Steam. *Elena*. URL: [store.steampowered.com/app/513860/Elena/](http://store.steampowered.com/app/513860/Elena/) (besucht am 03.09.2018) (siehe S. 46).
- [35] Steam. *The Gallery - Episode 1: Call of the Starseed*. URL: [store.steampowered.com/app/270130/The\\_Gallery\\_\\_Episode\\_1\\_Call\\_of\\_the\\_Starseed/](http://store.steampowered.com/app/270130/The_Gallery__Episode_1_Call_of_the_Starseed/) (besucht am 03.09.2018) (siehe S. 41).
- [36] Exit VR. *Exit VR Wien*. URL: [exit-vr.at/wien/](http://exit-vr.at/wien/) (besucht am 25.07.2018) (siehe S. 12).
- [37] Wikipedia. *Atmo*. Aug. 2018. URL: [de.wikipedia.org/wiki/Atmo](http://de.wikipedia.org/wiki/Atmo) (besucht am 10.09.2018) (siehe S. 8).
- [38] Wikipedia. *Head-Related Transfer Function*. Feb. 2018. URL: [de.wikipedia.org/wiki/Head-Related\\_Transfer\\_Function](http://de.wikipedia.org/wiki/Head-Related_Transfer_Function) (besucht am 17.08.2018) (siehe S. 16).
- [39] Zotac. *VR Go*. URL: [www.zotac.com/us/product/mini\\_pcs/vr-go](http://www.zotac.com/us/product/mini_pcs/vr-go) (besucht am 25.07.2018) (siehe S. 13).