Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Hauptseminar - Thesis Outline

Rutkay Küpelikilinc

Ambisonics

Fakultät Technik und Informatik
Studiendepartment Informatik

Faculty of Engineering and Computer Science
Department of Computer Science

Rutkay Küpelikilinc

Ambisonics

Hauptseminar - Thesis Outline eingereicht im Rahmen der Ausarbeitung zum Hauptseminar

im Studiengang Master Informatik
am Department Informatik
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

Eingereicht am: 31.08.2018

Inhaltsverzeichnis

1 Einleitung **1**
- 1.1 Ziel der Arbeit . . . 1
- 1.2 Motivation . . . 2
- 1.3 Wellenfeldsynthese Anlage . . . 2
  - 1.3.1 Ambisonics in der WFS-Anlage . . . 3

2 Grundlagen Ambisonics **4**
- 2.1 Aufbau . . . 4
- 2.2 Physikalisches Prinzip . . . 5
- 2.3 Vorteile von Ambisonics . . . 5
- 2.4 Kritik . . . 7

3 Aktuelle Forschung **7**

4 Leitfrage, Methodik, Risiken **9**
- 4.1 Methodik . . . 10
- 4.2 Risiken . . . 11

5 Fazit **12**

1 Einleitung

In diesem Kapitel soll ein erster Einblick in die Welt von Ambisonics und dem Vorhaben im Rahmen der Masterthesis ermöglicht werden. Um dies zu erreichen werden die wichtigsten Grundlagen kurz dargestellt sowie eine Übersicht der Anwendungsgebiete gegeben.

Das Themengebiet von Ambisonics liegt im Bereich der für räumliche Wiedergabe geeigneten Audioformate. Ziel dieses Forschungsbereiches ist die Optimierung der erzeugten Schallfelder und das Erlangen von Erkenntnissen zur auditiven Wahrnehmung des Menschen. Bei Schallfeldern handelt es sich um aus mehreren Tonquellen (Lautsprechern) zusammengesetzten Schallwellen. Ambisonics selbst ist eine Audioformat welches die zu Wiedergebenden Töne als mathematische Beschreibung der entsprechenden Druckgradienten hinterlegt. Die Berechnung basiert dabei auf der Kugelflächenfunktion. Des Weiteren können mit Hilfe höherer Ordnungen dieser Funktion, Higher Order Ambisonics (HOA) berechnet werden. Mit steigender Ordnung werden die Gradienten in feinere Bereiche aufgeteilt. Dies ermöglicht eine genauere Darstellung der Schallausbreitung, benötigt aber zur Wiedergabe eine größere Anzahl an Lautsprechern und eine höhere Rechenleistung. [Mal03, Dan01]

1.1 Ziel der Arbeit

Ziel der Arbeit ist es einen Überblick über den aktuellen Forschungsstand und über die geplanten Untersuchungen im Rahmen der Masterthesis zum Thema Ambisonics zu bieten. Dafür soll nach der aktuellen Relevanz des Themengebiets auf die Grundlagen von Ambisonics so wie der für die Nutzung vorgesehenen Wellenfeldsynthese-Anlage eingegangen werden. Dabei werden auch die Vorzüge und kritische Stimmen in Bezug auf Ambisonics beleuchtet. Folgend wird ein Überblick über aktuelle Forschungsbereiche und Ergebnisse um das Thema Ambisonics gegeben. Im Letzten Kapitel wird auf die geplante Umsetzung und die Ziele der Untersuchungen eingegangen. Des Weiteren wird auf die Probleme eingegangen, welche im Laufe der Arbeit auftreten können.

1.2 Motivation

Ambisonics wurde schon 1973 von Michael Gerzon vorgestellt [Ger73] hat aber wenig kommerziellen Erfolg verbuchen können, da dass Thema seiner Zeit wohl etwas voraus war. Nun wo virtuelle Realität und 3D oder auch 360-Grad Filme, Spiele und Videos aktueller Forschungsgegenstand sind und immer weiter in den Endverbrauchermarkt vordringen, werden auch passenden Audioformate benötigt, welche ein räumliches Klangbild erzeugen und der visuellen Bewegung anpassen können. Durch die Kopplung von auditiver und visueller räumlicher Wahrnehmung kann so eine starke Immersion in virtuellen Welten erzeugt werden. Bisher wird dieser Bereich größtenteils für die Unterhaltungsindustrie genutzt. Aber durch die kontinuierliche Verbesserung der Hardwarekomponenten in Bezug auf die Qualität der Darstellung werden auch weitere Anwendungsgebiete erschlossen. Zu diesen Anwendungsgebieten dienen etwa Präsentationen von 3D-Modellen. So können etwa Häuser oder Autos vor dem Vertragsabschluss virtuell in der gewählten Ausstattung besichtigt werden um einen Eindruck des ganzen zu bekommen und mögliche Änderungen vorzeitig einzuleiten. Auch die Nutzung der virtuellen Realität im medizinischen Bereich ist ein aktuelles Forschungsgebiet. So können etwa Operationen simuliert und gefahrlos einstudiert werden [GRC+05]. Es sind etliche Trainingssituationen denkbar, wie etwa auch kosteneffektives Flugtraining für Piloten. [BC03]

Auch auf der Weltweit größten Videoplattform Youtube kann Ambisonics als Audioformat in 360-Grad Videos genutzt werden.[1]

Des Weiteren soll das Thema auf der Bachelorthesis mit dem Thema „Untersuchung der Höhenwahrnehmung an einer um Einzellautsprecher erweiterten WFS-Anlage“aufbauen und es sollen Untersuchungen angestellt werden, welche das Potenzial von Ambisonics Rendering gegen das Rendering mit Hilfe von Wellenfeldsynthese ermitteln. Dies ist vor allem relevant um auch die Wellenfeldsynthese-Anlage der HAW Hamburg um Ambisonics als weiteres Renderingverfahren zu erweitern und so eine Alternative zur Wellenfeldsynthese zu implementieren, da diese sich als weniger massentauglich erwiesen hat. Auch ist Ambisonics mittlerweile im MPEG-H und ATSC 3.0 Standard inkludiert.

1.3 Wellenfeldsynthese Anlage

Die Implementierung von Ambisonics soll innerhalb der Wellenfeldsynthese Anlage der HAW Hamburg realisiert werden. Die Wellenfeldsynthese stellt dabei ein alternatives Renderingver-

[1] https://support.google.com/youtube/answer/6395969 (27.07.2018)

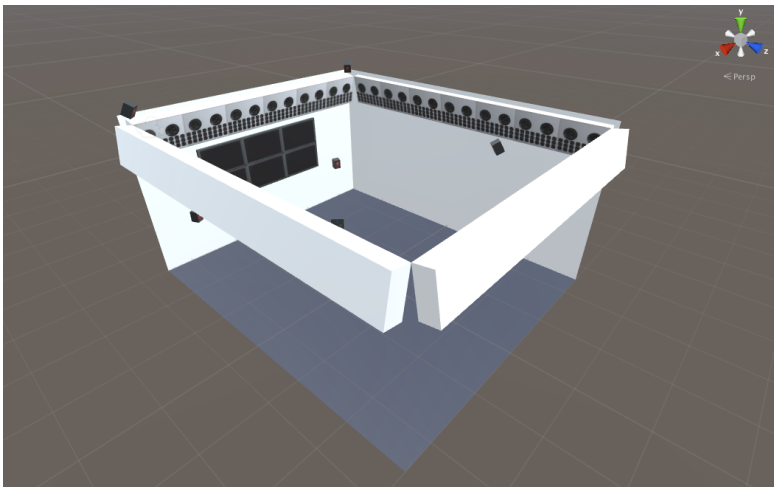


Abbildung 1.1: Die WFS-Anlage der HAW Hamburg als Modell

fahren für die Wiedergabe von räumlichen Klangkulissen dar. Das Grundprinzip der Wellenfeldsynthese beruht dabei auf der Überlagerung der sogenannten Elementarwellen. Durch die zeitlich verzögerte Wiedergabe der unterschiedlich positionierten Elementarwellen welche von den Lautsprecherarrays erzeugt werden überlappen diese sich, abhängig vom Abstand der Lautsprecher zueinander, nach einer bestimmten Distanz zu einer Wellenfront, dies ergibt die synthetisierte Schallwelle. Ein großer Vorteil von Ambisonics gegenüber der Wellenfeldsynthese ist die einfache Erweiterbarkeit vor allem in der Vertikale um eine zusätzliche Dimension der Wahrnehmung zu schaffen. Die Wellenfeldsynthese-Anlage in ihrer ursprünglichen Form ist auf die horizontale Ebene beschränkt. In der Abbildung 1.1 ist ein Modell der Anlage von schräg oben zu sehen. Dort erkennbar sind die 26 auf einer Höhe von zwei Metern aneinandergereihten Lautsprechermodule. Die Module bestehen jeweils aus 24 Breitbandlautsprechern und 2 Tieftonchassis. Weiterhin sind die 6 Infrarotkameras des Trackingsystems sowie die Powerwall zu erkennen. Das Trackingsystem ist für die geplanten Arbeiten relevant, da mit Hilfe dessen genaue Postionen der Probanden so wie Auswertungen auf Basis von berechneten Punkten erstellt werden können. [Foh13]

1.3.1 Ambisonics in der WFS-Anlage

Die Anlage ist speziell auf die Wellenfeldsynthese zugeschnitten und bietet somit eine Vielzahl kleiner Lautsprecher. Ambisonics soll ein Sphärisches Wellenfeld erzeugen, so dass im optimalen Fall ein korrektes Schallfeld und somit die Ortung von Positionen in jede Richtung möglich ist. Dafür werden zusätzliche Lautsprecher an weiteren Positionen im Raum benötigt.

Die Anzahl der Lautsprecher ist dabei abhängig von der gewählten Ordnung mit welcher das Ambisonics Rendering genutzt wird. Für eine korrekte Wiedergabe werden $(N+1)^2$ Lautsprecher benötigt, wobei N der Higher Order Ambisonics Ordnung entspricht [Fra14]. Um die Lautsprecher der WFS-Anlage mit Ambisonics zu nutzen, können die einzelnen Lautsprecherkanäle angesprochen werden, alternativ dazu können per Software virtuelle Schallquellen erzeugt werden, welche in diesem Fall als virtuelle Lautsprecher fungieren. Die Virtuellen Lautsprecher bringen den Vorteil mit sich, dass sie auf der Horizontalebene einfach in die gewünschte Position gesetzt werden können.

2 Grundlagen Ambisonics

In diesem Abschnitt sollen die Grundlagen zum Thema Ambisonics behandelt werden. Bei Ambisonics handelt es sich um ein Verfahren welches zur Aufnahme und Wiedergabe von Schallfeldern entwickelt wurde. Mit Hilfe von Ambisonics können Schallfelder realitätsnah aufgenommen und wiedergegeben werden. Die Besonderheit dieses Verfahrens ist die Möglichkeit der dreidimensionalen Aufnahme und Wiedergabe des Schallfeldes. So werden Klänge mit allen Richtungsinformationen aufgenommen und wiedergegeben. Außer der Aufnahme und Wiedergabe mit Ambisonics können auch Mono Signale, unabhängig von Aufnahmen im Ambisonics Format, an einem beliebigen Punkt in einem dreidimensionalen Schallfeld wiedergegeben werden.

2.1 Aufbau

Ambisonics funktioniert auf beliebig sphärisch orientierten Lautsprecheraufbauten. Die Mindestanzahl der Lautsprecher orientiert sich dabei an der genutzten Ordnung von Higher Order Ambisonics welche Wiedergegeben werden soll. Wie schon in Abschnitt 1.3.1 beschrieben werden auf Basis der Ordnung $(N+1)^2$ Lautsprecher für eine korrekte Wiedergabe benötigt. Jede Position und Ausrichtung muss im System bekanntgegeben werden. Die genaue Position ist essenziell um akustische Fehler im Schallfeld zu vermeiden, da erst eine gezielte Überlappung der verschiedenen aus den Lautsprechern erzeugten Schallwellen die Funktionsweise von Ambisonics ermöglicht. Fehler können etwa durch falsche Ausrichtung, Bsp.: Schallfronten aus verschiedenen Lautsprechern überlappen sich im falschen Winkel oder ungenauer Abstände

entstehen in diesem Fall treffen die Schallfronten mit zu großer Verzögerung ein und werden als zwei eigenständige Töne wahrgenommen. [Ger73, Fra14, Mal03, WA01]

2.2 Physikalisches Prinzip

Das physikalische Prinzip von Ambisonics basiert auf der Kugelflächenfunktion, mit Hilfe derer die Gradienten des Schalldrucks sowie der Schallschnelle für die einzelnen im System registrierten Lautsprecher errechnet werden. Die Anzahl der Basisfunktionen der Kugelflächenfunktion erhöht sich mit steigender Ordnung. Aus einer höheren Ordnung resultiert auch eine erhöhte Auflösung und damit eine bessere Wiedergabequalität im Sinne der Tonqualität sowie der Darstellung des immersiven Raumklanges. Der Aufbau der verschiedenen Ordnungen und die jeweiligen Basisfunktionen ist in Abbildung 2.1 zu sehen. Die dunklen Flächen stellen dabei die Inverse Polarität der Gradienten dar. Die verbesserte Wiedergabequalität mit steigender Ordnung ist auf die größere Diversifizierung der Basisfunktionen und somit feineren Gradienten zurückzuführen. Durch die Verfeinerung der Gradienten sind diese exakter, da sie für jeweils eine kleinere Fläche zuständig sind und somit präzisere Informationen für die Druckverteilung im Raum erlangt werden. Der Idealfall wäre die größtmögliche Ordnung welche Basisfunktionen enthält welche die gesamte Kugeloberfläche in unendlich kleinen Flächen erfasst. So hätte man im Idealfall den Schalldruck und die Schallschnelle aus exakt jedem Punkt des Umfeldes aufgenommen bzw. errechnet und könnte dies nun Wiedergeben. Das erhöhen der Ordnung ist durch den Aufbau (Anzahl benötigter Lautsprecher) und vermehrt benötigter Rechenleistung (steigende Komplexität) beschränkt. [Ger73, Fra14]

In Abbildung 2.2 ist zu erkennen wie sich die Ausbreitung und die Qualität der resultierenden Wellenfronten verschiedener Ordnungen unterscheiden. Der schwarze Punkt stellt die Hörerposition dar. Bei der höheren Ordnung ist eine deutlich glattere und geschlossene Wellenfront erkennbar.

2.3 Vorteile von Ambisonics

Die Vorteile von Ambisonics gegenüber herkömmlichen kanalbasierten Wiedergabeverfahren zeigen sich vor allem in der flexiblen Positionierung von Soundquellen. Mit einem Ambisonics Encoder können Mono Signale in jeder beliebigen Position im Raum platziert werden. Aber auch schon im Ambisonics Format aufgenommene Audiodateien können einfach verarbeitet und umplatziert werden. Durch die mathematische Basis von Ambisonics sind diese Operationen einfach umsetzbar. Vor allem die Flexibilität beim verdrehen gesamter Audioszenerien (z.B. in

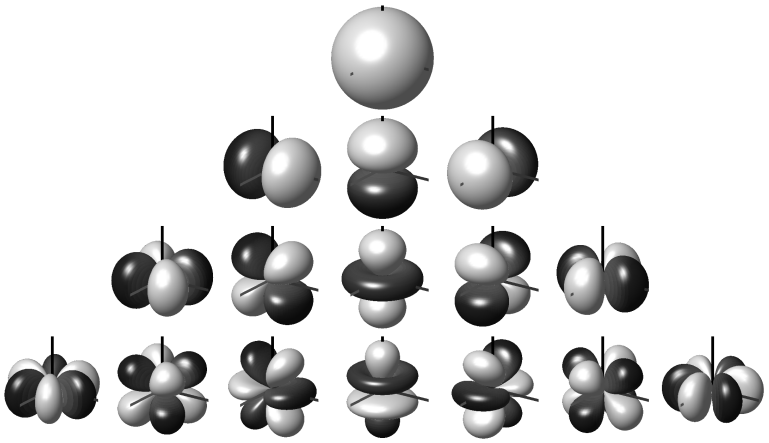
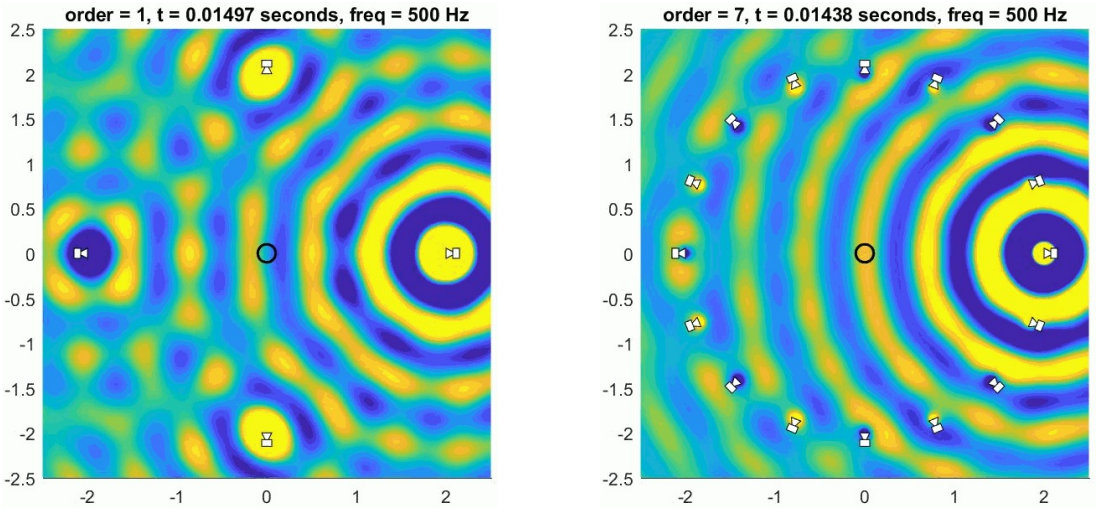


Abbildung 2.1: Ambisonics Ordnungen von 0 bis 3 (v.o.n.u) grafisch dargestellt [1]



(a) HOA Ordnung 1

(b) HOA Ordnung 7

Abbildung 2.2: Ausbreitung von mit Higher Order Ambisonics berechneten Schallwellen [2]

Kombination mit der Bewegung eines Head Mounted Displays) macht Ambisonics interessant für den VR-Bereich.

Bei dem Aufbau erweist sich die Flexibilität in der Positionierung der Lautsprecher als sehr Komfortabel, da so besser auf die räumlichen Gegebenheiten eingegangen werden kann. Auch kann mit Ambisonics schon mit nur vier Wiedergabekanälen ein dreidimensionales Schallfeld wiedergegeben werden. Andere räumliche Wiedergabeverfahren benötigen für die korrekte Darstellung eines räumlichen Schallfeldes eine größere Anzahl an Lautsprechern.

Ein weiterer Vorteil ist die Kompatibilität bzw. Flexibilität zu herkömmlichen Formaten, wie etwa Stereo oder 5.1. So können im Ambisonics Format gespeicherte Daten einfach in herkömmliche Formate umgewandelt werden. Eine weitere Methode ist es Ambisonics binaural zu dekodieren und über herkömmliche Kopfhörer zu nutzen, was hardwaretechnisch betrachtet die kostengünstigste Variante darstellt. [NSMH03]

2.4 Kritik

Im Laufe der Jahre gab es immer wieder kritische Stimmen, welche behaupten, dass Ambisonics nicht funktionieren könne. Unter anderem wurde behauptet, dass der Klang sich innerhalb des Kopfes abspielen würde [Too08]. Um dies zu widerlegen haben sich E. Benjamin, A. Heller und R. Lee dem Thema angenommen und sich mit den einzelnen Kritikpunkten beschäftigt [BHL10]. Im Zuge ihrer Untersuchungen konnten sie die entsprechenden Behauptungen widerlegen. Der Klang innerhalb des Kopfes kann laut der Ergebnisse in jedem Lautsprechersystem mit mehr als einem Lautsprecher entstehen, ließe sich aber im Falle von Ambisonics aber mit Nutzung einer höheren Ordnung für Aufnahme und Wiedergabe beheben.

Auch der sweet spot, der Bereich innerhalb des Lautsprecheraufbaus an dem das Schallfeld optimal reproduziert wird, wird oft kritisch gesehen. Der sweet spot bei Ambisonics Anwendungen ist Berechnungen zufolge nur wenige Zentimeter groß [WA01] und somit in der Theorie nicht einmal ausreichend groß um einen menschlichen Kopf zu umfassen. Diese Problematik besteht aufgrund der menschlichen Wahrnehmung vor allem in höheren Frequenzbereichen. Für dieses Problem gibt es mehrere Lösungsansätze, einerseits ist der vom Hörer wahrgenommene sweet spot umfangreicher als der Theoretisch errechnete [FZS15] des Weiteren kann auch der theoretisch errechnete sweet spot mit erhöhen der Ambisonics Ordnung erweitert werden.

3 Aktuelle Forschung

Die aktuelle Forschung beschäftigt sich vor allem im Bereich der Optimierung des Schallfeldes/Hörerlebnisses meist in Kombination mit VR-Brillen bzw. Head Mounted Displays. Eine Optimierungsmöglichkeit ist etwa eine um Artefakte reduzierte und somit einheitlichere Wellenfront [WHWS17]. Die Optimierungen zielen oftmals zusätzlich auf eine Reduzierung

der Berechnungszeiten ab um so etwa möglichst schnell höhere Ordnungen von Ambisonics berechnen zu können oder dieses mit wenig Rechenaufwand zu verbessern.

Ein Teilbereich der Optimierung des Hörerlebnisses liegt dabei in der Erforschung des Wiedergabeverhaltens von direktionalem Klang und der Verbesserung durch gezieltes Verstärken bestimmter Signale aus ausgewählten Richtungen z.B. mit Hilfe des Clebsch-Gordan Koeffizienten. So kann die Richtungswahrnehmung des Menschen in der künstlich erzeugten, auditiven Umgebung präzisiert werden. Vor allem Ambisonics niedriger Ordnungen können von diesem Verfahren profitieren und kommen mit weniger Rechenaufwand aus. Laut Bastiaan Kleijn könne so die Wiedergabequalität auf die von höheren Ordnungen hochskaliert werden. [Kle18]

Eine weitere Arbeit in diesem Gebiet untersucht die resultierende Abstrahlcharakteristik bei der Wiedergabe der im Ambisonics Format berechneten Tonspuren. Für die Untersuchungen wurden verschiedene Softwarelösungen miteinander verglichen, da diese verschiedene Methoden nutzen um das Klangfeld zu optimieren. Es wurde dabei vor allem mit modalem Beamforming verglichen, da die Prozesse zum berechnen von Ambisonics diesen sehr ähneln. [MEB18].

Einen Ansatz zur Individualisierung und damit Optimierung der Wiedergabe haben Binelli et al. kürzlich vorgestellt. Dabei wird die Head-Related Transfer Function (HRTF) um so auf die individuelle Wahrnehmung des Menschen eingehen zu können. Die HRTF beschreibt eine auf den Kopf und Gehörgang eines Individuums bezogene Funktion welche die filternde Wirkung des Kopfes sowie den individuellen Einfall und die Reflexionen innerhalb des Gehörganges beschreibt [Bla74]. Mit HRTF können somit hochindividuelle Anpassungen zur Optimierung der auditiven Wahrnehmung vorgenommen werden, wodurch die richtungsbasierte Wahrnehmung gestärkt wird. Ein Problem stellt dabei die genau Analyse der individuellen Charakteristiken dar, gerade wenn es von einem breiterem Publikum genutzt werden soll. Für die Untersuchungen wurde zwecks Vereinfachung und Vergleichbarkeit ein Dummy-Kopf in einem schalltotem Raum genutzt. Die entwickelte Software unterstützt die dynamische Wiedergabe von Ambisonics in Verbindung mit HRTFs, gekoppelt mit der Wiedergabe einer 360° Videoumgebung. [BPNF18]

Auch durch einfach wirkende Eingriffe und Filter kann das Klangergebnis optimiert werden. So hat etwa Shu-Nung Yao einen Artikel veröffentlicht aus welchem hervorgeht, dass mit der Filterung bestimmter Frequenzen, welche sich negativ beeinflussen können, eine verbesserte Richtungswahrnehmung ergäbe. Dies sei vor allem bei tiefen Frequenzen der Fall. Bei denen eine akustische Lokalisierung für den Menschen ohnehin schwieriger ist [WRS13]. Des Weiteren

vereinfache sich durch die Nutzung des Filters auch die Decodierung auf höhere Ambisonics Ordnungen [Yao18]

Viele der Verbesserungen entspringen Erkenntnissen aus der Psychoakustik, dabei geht es um die akustische Wahrnehmung des Menschen. Wie etwa verschiedene Frequenzen wahrgenommen werden, vor allem im Bezug auf die akustische Lokalisierung. So kann die Wiedergabe von symmetrisch versetzten Schallquellen dazu führen, dass es zu einer Unsicherheit kommt ob der Ton vor, hinter oder über einem abgespielt wird. Dies ist bedingt durch die zu kleine interaurale Zeitdifferenz bei dem eintreffen der Töne in die Ohrmuscheln. Diesem Problem wurde in Bezug auf allgemeine 3D-Audiowiedergabe schon 2006 von Cho et al. untersucht [COC06]. Matthias Frank und und Franz Zotter haben sich dem Problem der Front-Back Confusion erneut angenommen und haben einen Lösungsweg für Ambisonics erarbeitet. Um einen von vorne stammenden Ton besser erkennbar zu machen wurde ein Kuhschwanzfilter eingesetzt um Töne ab einer Frequenz von 3 kHz zu verstärken. Umgekehrt werden für die Darstellung eines von hinten stammenden Geräusches die selben Frequenzen in ihrer Intensität verringert. [FZ18]

Aber auch Untersuchungen zu der Wahrnehmung von Ambisonics werden betrieben. So wurde etwa die Wahrnehmung des sweet spots untersucht, dieser ist in der Theorie durch Simulationen als zu klein für eine ordentliche Wahrnehmung befunden worden. In der Realität stellte sich aber heraus, dass die gewünschten räumlichen Effekte von den Probanden trotzdem wie gewünscht wahrgenommen werden [FZ17]. In diesem Bereich wird auch der Fokus der geplanten Arbeit liegen.

4 Leitfrage, Methodik, Risiken

Die Fragestellung im Laufe der Masterthesis dreht sich um die Nutzung von Ambisonics in Kombination mit der Wellenfeldsynthese-Anlage der HAW Hamburg. Im Laufe der Forschungen sollen Fragen zur Umsetzbarkeit einer solchen Implementierung geklärt werden wie z.B.: Ist ein starrer Aufbau wie die Wellenfeldsynthese-Anlage geeignet um diese mit in ein Ambisonics Setup zu integrieren. Dabei sind viele Fragen und mögliche Probleme zu beachten. Wie etwa verschiedene Frequenzgänge und Abstrahlcharakteristik der genutzten Hardware. Der darauf Folgende Schritt ist der Aufbau eines Ambisonics Setups. Dafür sollen auch diverse Optimierungen aus dem Abschnitt 3 umgesetzt werden. Nach erfolgreichem Aufbau des Systems sollen

Hörtests mit Probanden durchgeführt werden sowie Auswertungen des erzeugten Schallfeldes mit Hilfe eines Soundfieldmikrofons[1], welches in der Lage ist ein Schallfeld dreidimensional aufzunehmen.

Mit Untersuchungen in der Wellenfeldsynthese-Anlage soll verglichen werden ob sich Ambisonics in dem gegebenen Umfeld gegen die Wellenfeldsynthese durchsetzen kann. Durch die Analyse der Ergebnisse können die erlangten Erkenntnisse auch weiteres Optimierungspotenzial zum Vorschein bringen. So wird ein Blick auf den aktuellen Stand von Ambisonics erlangt und es können verschiedene Optimierungsmaßnahmen kombiniert und auf ihre Effektivität im Verbund untersucht werden.

Die Leitfrage der Untersuchung lautet dabei:

> Ist eine signifikante Verbesserung der auditiven Lokalisierung mit Ambisonics gegenüber der Wellenfeldsynthese gegeben, insbesondere in Hinblick auf die vertikale Ebene?

4.1 Methodik

In diesem Abschnitt soll skizziert werden auf welchem Weg das Ziel, von aussagekräftigen Erkenntnissen durch Hörtests und Messungen, erreicht werden soll.

Ein wichtiger Schritt dafür ist die Konzeption und Durchführung eines Hörversuchs welcher aussagekräftige Daten liefert und Vergleiche innerhalb der Versuchsdaten sowie zu Ergebnissen ähnlicher Untersuchungen zulässt. Dies ist essenziell Notwendig um korrekte Schlüsse aus den Daten ziehen zu können und diese Ergebnisse auch nach außen mit Erkenntnissen anderer Forscher verifizieren zu können.

Der Kern des Hörversuchs dreht sich um die Ortungspräzision der Ursprungsposition des wiedergegebenen Tons in synthetisierten Schallfeldern. Da sowohl die Wellenfeldsynthese als auch Ambisonics als Wiedergabeverfahren synthetisierte Schallfelder erzeugen ist somit ein Vergleich der Ergebnisse beider Verfahren möglich. Die technische Seite ist dabei gut mit dem Trackingsystem der Wellenfeldsynthese Anlage umsetzbar, welches zur Positionsermittlung des anvisierten Tonursprungs der Probanden genutzt werden kann. Dazu soll der Proband ein von der Anlage erfasstes Objekt auf den vermeintlichen Tonursprung richten. Im System wird dann mit Hilfe von Richtungsvektoren ein anvisierter Punkt auf einer Ebene und die Abweichung zum tatsächlichen Tonursprung berechnet [Küp16]. Dieses System soll im Laufe der Arbeit im Sinne der Nutzerfreundlichkeit optimiert werden. Die so erfassten Ergebnisse werden

[1] https://www.soundfield.com/products/sps200 (20.08.2018)

dann zu Auswertungszwecken in Dateien im XML Format geschrieben, um eine einfache Auswertung mit Statistikprogrammen zu gewährleisten. Auch die subjektive Wahrnehmung des Tons spielt eine Rolle, denn auch wenn die Ortung gut funktionieren sollte ist es wichtig, dass der Ton auch angenehm und natürlich wahrgenommen wird, da für den Erfolg des Systems auch eine gewisse Akzeptanz der Hörer erforderlich ist. Die Ergebnisse sollen mit vorangegangenen Versuchen, die u. a. im Rahmen der Bachelorthesis vorgenommen wurden, verglichen werden [Küp16, Mon11].

Abgesehen von den subjektiven Hörtests sollen Aufnahmen und Auswertungen des Schallfeldes mit Hilfe des Soundfieldmikrofons umgesetzt werden. Durch die Aufnahmen des Soundfieldmikrofons kann das jeweils erzeugte Schallfeld mit Tools wie etwa Matlab visualisiert und untersucht werden. Kriterien die dabei besonders beachtet werden sind die Homogenität der Wellenfront und Artefaktbildung im Schallfeld. Durch dieses Verfahren kann auch der theoretische sweet spot bestimmt werden und untersucht werden in wie weit dieser mit den Hörtests übereinstimmt. Als Vergleich sollen die Signale auch direkt am Ausgang abgegriffen und erfasst werden. So können die theoretisch erzeugten Signale (Abgreifen der Signale) mit den real erzeugten (Aufnahmen des Soundfieldmikrofons) verglichen und auf Unregelmäßigkeiten überprüft werden.

4.2 Risiken

Abgesehen von den in der Einleitung des Kapitels aufgezählten Risiken besteht die Gefahr, dass die Hörversuche mit Probanden keine eindeutigen Ergebnisse liefern oder falsche Schlüsse aus den Ergebnissen gezogen werden. Um dies zu verhindern muss der Hörversuch bedacht geplant werden. Auch die Vergleichsdaten können Fehlschlüsse enthalten, welches sich bei Nutzung der Daten auf die eigenen Ergebnisse auswirken kann. Eine weitere potenzielle Gefahr ist dabei, dass die Probanden die Fehler im Schallfeld kompensieren und somit ein unnötiger Aufwand getrieben wird um dieses zu optimieren.

Bei Hörversuchen mit Probanden ist es wichtig andere Wahrnehmungseinflüsse zu verhindern, da diese das Ergebnis stark beeinflussen können. So können visuelle Einflüsse, wie zum Beispiel sichtbare Lautsprecher dazu führen, dass der Tonursprung auf genau dieser Höhe erwartet wird und sich nicht auf die rein auditive Wahrnehmung verlassen wird. Um dies zu verhindern sind Aufbauten mit schalldurchlässigen Vorhängen oder eine VR-Brille als Sichtschutz denkbar. Die VR-Brille würde auch weitere Untersuchungen zum Zusammenspiel der visuellen und auditiven Wahrnehmungen zulassen wie etwa in Versuchen zum Redirected Walking [NF16]. Dazu können den Probanden verschiedene Szenarien vorgetäuscht werden.

So entsteht eine Möglichkeit die Abhängigkeit der akustischen und visuellen Wahrnehmung zu erforschen.

Auch weitere psychoakustische Effekte sind zu beachten wie etwa die unterschiedliche Wahrnehmung von Frequenzbereichen in Lautstärke, Qualität der Lokalisierung und der empfundenen Ursprungsposition [Bla13]. Auch der korrekte Aufbau und die Abstimmung der Komponenten aufeinander spielt dabei eine wichtige Rolle da sonst die Wahrnehmung, durch etwa falsche Lautstärkeregelung oder durch den Aufbau bedingte Artefakte in der Wellenfront, stark beeinflusst werden kann [SWR+13].

Auch das Zeigen auf einem Punkt mit dem getrackten Objekt ist dabei kritisch zu betrachten, da durch die Verlängerung des Armes eine gewisse Ungenauigkeit entsteht. Sinngemäß der Fragestellung: „Wie genau kann ein Mensch eigentlich zeigen?“ Diese Abweichung könnte weiter untersucht werden. In Kombination mit der VR-Brille ist es vorstellbar die Handstücke zu Nutzen, da so in dem virtuellen Raum der anvisierte Punkt angezeigt werden kann. Eine manuelle Option wäre die Nutzung eines feinen Rasters mit Koordinaten, welche zur Positionsbestimmung genannt werden.

5 Fazit

Das Thema Ambisonics wird durch die starke Ausbreitung von VR-Anwendungen immer interessanter und zeigt durch Unterstützung in beliebten Codecs so wie auf wichtigen Plattformen wie YouTube auch eine längerfristige Relevanz auf. Nach dem Aufzeigen der Grundlagen so wie der Umgebung in der die Versuche ausgeführt werden sollen, wurden aktuelle Forschungen zum Thema Ambisonics vorgestellt. Dabei wurden vor allem auf die Optimierungsmöglichkeiten von Ambisonics eingegangen. Zum Abschluss wurde auf die Umsetzungsmöglichkeiten einer Ambisonics Implementierung, innerhalb der Wellefenldsynthese-Anlage der HAW Hamburg, im Rahmen der Masterthesis eingegangen. Dabei wurden Fragestellungen die im Laufe der Untersuchungen auftreten können beleuchtet sowie die geplante Durchführung der Versuche zum Erlangen von Erkenntnissen über die Wahrnehmung von Ambisonics generierten Schallfeldern. Schlussendlich wurde auf die Forschungsrisiken und auf Alternativen oder Möglichkeiten zur Minimierung dieser eingegangen.

Literaturverzeichnis

[BC03] Grigore C. Burdea and Philippe Coiffet. *Virtual Reality Technology, Second Edition with CD-ROM.* Wiley-IEEE Press, June 2003.

[BHL10] Eric Benjamin, Aaron Heller, and Richard Lee. Why ambisonics does work. In *Audio Engineering Society Convention 129*, Nov 2010.

[Bla74] J. Blauert. *Räumliches Hören.* Monographien der Nachrichtentechnik. Hirzel, 1974.

[Bla13] Jens Blauert. *The Technology of Binaural Listening.* Springer Science & Business Media, Berlin Heidelberg, 2013. aufl. edition, 2013.

[BPNF18] Marco Binelli, Daniel Pinardi, Tiziano Nili, and Angelo Farina. Individualized hrtf for playing vr videos with ambisonics spatial audio on hmds. In *Audio Engineering Society Conference: 2018 AES International Conference on Audio for Virtual and Augmented Reality*, Aug 2018.

[COC06] S. J. Cho, A. Ovcharenko, and U. Chong. Front-back confusion resolution in 3d sound localization with hrtf databases. In *2006 International Forum on Strategic Technology*, pages 239–243, Oct 2006.

[Dan01] Jerome Daniel. *Représentation de champs acoustiques, application à la transmission et à la reproduction de scènes sonores complexes dans un contexte multimédia.* PhD thesis, Université Paris 6, 2001.

[Foh13] Wolfgang Fohl. *The Wave Field Synthesis Lab at the HAW Hamburg in Sound - Perception - Performance.* Springer Science & Business Media, Berlin Heidelberg, 2013. aufl. edition, 2013.

[Fra14] Matthias Frank. How to make ambisonics sound good. 2014.

[FZ17] Matthias Frank and Franz Zotter. Exploring the perceptual sweet area in ambisonics. In *Audio Engineering Society Convention 142*, May 2017.

[FZ18] Matthias Frank and Franz Zotter. Simple reduction of front-back confusion in static binaural rendering. *Fortschritte der Akustik, DAGA*, 2018.

[FZS15] Matthias Frank, Franz Zotter, and Alois Sontacchi. Producing 3d audio in ambisonics. In *Audio Engineering Society Conference: 57th International Conference: The Future of Audio Entertainment Technology – Cinema, Television and the Internet*, Mar 2015.

[Ger73] Michael A. Gerzon. Periphony: With-height sound reproduction. *J. Audio Eng. Soc*, 21(1):2–10, 1973.

[GRC+05] Anthony G Gallagher, E Matt Ritter, Howard Champion, Gerald Higgins, Marvin P Fried, Gerald Moses, C Daniel Smith, and Richard M Satava. Virtual reality simulation for the operating room: proficiency-based training as a paradigm shift in surgical skills training. *Annals of surgery*, 241(2):364–372, February 2005.

[Kle18] W. B. Kleijn. Directional emphasis in ambisonics. *IEEE Signal Processing Letters*, 25(7):1079–1083, July 2018.

[Küp16] Rutkay Küpelikilinc. Untersuchung der Höhenwahrnehmung an einer um Einzellautsprecher erweiterten WFS-Anlage, 2016.

[Mal03] Dave Malham. Higher-order Ambisonic systems, 2003.

[MEB18] Jens Meyer, Gary W. Elko, and Steven Backer. Modal beamformer analysis of ambisonics decoders. *The Journal of the Acoustical Society of America*, 143(3):1823–1823, 2018.

[Mon11] Matthew N. Montag. Wave Field Synthesis In Three Dimensions By Multiple Line Arrays. Masterthesis, University of Miami, 2011.

[NF16] M. Nogalski and W. Fohl. Acoustic redirected walking with auditory cues by means of wave field synthesis. In *2016 IEEE Virtual Reality (VR)*, pages 245–246, March 2016.

[NSMH03] Markus Noisternig, Alois Sontacchi, Thomas Musil, and Robert Holdrich. A 3d ambisonic based binaural sound reproduction system. In *Audio Engineering Society Conference: 24th International Conference: Multichannel Audio, The New Reality*, Jun 2003.

[SWR+13] S. Spors, H. Wierstorf, A. Raake, F. Melchior, M. Frank, and F. Zotter. Spatial sound with loudspeakers and its perception: A review of the current state. *Proceedings of the IEEE*, 101(9):1920–1938, Sept 2013.

[Too08] F.E. Toole. *Sound Reproduction: Loudspeakers and Rooms*. Audio Engineering Society Presents. Elsevier, 2008.

[WA01] D. B. Ward and T. D. Abhayapala. Reproduction of a plane-wave sound field using an array of loudspeakers. *IEEE Transactions on Speech and Audio Processing*, 9(6):697–707, Sept 2001.

[WHWS17] Fiete Winter, Nara Hahn, Hagen Wierstorf, and Sascha Spors. Azimuthal localization in 2.5d near-field-compensated higher order ambisonics. *The Journal of the Acoustical Society of America*, 141(5):3934–3934, 2017.

[WRS13] H. Wierstorf, A. Raake, and S. Spors. *The Technology of Binaural Listening*, chapter Binaural Assessment of Multichannel Reproduction, pages 255–278. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.

[Yao18] Shu-Nung Yao. Equalization in ambisonics. *Applied Acoustics*, 139:129 – 139, 2018.

Bildquellen

[1] Dr. Franz Zotter, CC BY-SA 3.0, https://upload.wikimedia.org/wikipedia/commons/c/c1/Spherical_Harmonics_deg3.png (17.08.2018)
[2] https://www.ssa-plugins.com/blog/2017/07/18/what-is-higher-order-ambisonics/ (15.05.2018)