



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Bachelorarbeit

Rutkay Küpelikilinc

**Untersuchung der Höhenwahrnehmung an einer um
Einzellautsprecher erweiterten WFS-Anlage**

*Fakultät Technik und Informatik
Studiendepartment Informatik*

*Faculty of Engineering and Computer Science
Department of Computer Science*

Rutkay Küpelikilinc

**Untersuchung der Höhenwahrnehmung an einer um
Einzellautsprecher erweiterten WFS-Anlage**

Bachelorarbeit eingereicht im Rahmen der Bachelorprüfung

im Studiengang Bachelor of Science Technische Informatik
am Department Informatik
der Fakultät Technik und Informatik
der Hochschule für Angewandte Wissenschaften Hamburg

Betreuender Prüfer: Prof. Dr. Wolfgang Fohl
Zweitgutachter: Prof. Dr.-Ing. Andreas Meisel

Eingereicht am: 12. Mai 2016

Rutkay Küpelikilinc

Thema der Arbeit

Untersuchung der Höhenwahrnehmung an einer um Einzellautsprecher erweiterten WFS-Anlage

Stichworte

Wellenfeldsynthese, 3D-Audiowahrnehmung, 3D-Audio, Akustische Lokalisierung, Schallquelle, Akustische Höhenwahrnehmung

Kurzzusammenfassung

In dieser Arbeit wird die zweidimensionale Wellenfeldsynthese-Anlage der HAW Hamburg mittels zwei weiterer Lautsprecher-Arrays mit je acht Lautsprechern um die dritte Dimension erweitert, um Schallquellen der Wellenfeldsynthese-Anlage auch in der Höhe frei gezielt platzieren zu können. Um das Ergebnis zu analysieren, wurde ein Hörtest, welcher die Präzision der Wahrnehmung von virtuellen Schallquellen an verschiedenen Positionen ermittelt sowie ein vorgeschalteter Referenz-Test zum Vergleich mit der optischen Wahrnehmung konzipiert und implementiert.

Rutkay Küpelikilinc

Title of the paper

Examination of height perception in a WFS-system extended by individual speakers

Keywords

Wave Field Synthesis, 3D-Audio perception, 3D-Audio, Acoustic localization, Soundsource, Acoustic height perception

Abstract

This thesis describes an extension of the wave field synthesis system of the HAW Hamburg by two additional loudspeakerarrays, consisting of eight loudspeakers each, in order to achieve a third dimension. A hearing test is developed to analyze the precision of the perception achieved from virtual sound sources on different positions, together with a visual test for comparison purposes.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Ziel der Arbeit	1
1.2	Themenabgrenzung	1
1.3	Aufbau	2
2	Akustisch-räumliche Wahrnehmung des Menschen	3
2.1	Akustischer Sweet Spot	3
2.2	Phantomschallquelle	4
2.3	Akustische Lokalisation	4
2.4	Wahrnehmung verschiedener Frequenzbereiche	5
3	Audiowiedergabeverfahren	7
3.1	Monophonie	7
3.2	Stereophonie	7
3.3	3D-fähige Wiedergabeverfahren/Stand der Technik	8
3.3.1	Wellenfeldsynthese	8
3.3.2	VBAP	8
3.3.3	Binaurales 3D-Audio	9
3.3.4	Ambisonics	9
3.3.5	Zukunft der dreidimensionalen Audiowiedergabeverfahren	9
4	Wellenfeldsynthese	11
4.1	Funktionsweise der Wellenfeldsynthese	12
4.2	Aufbau einer Wellenfeldsynthese-Anlage	12
4.3	Technische Umsetzung	13
4.4	Darstellungsformen von Audioquellen	14
4.4.1	Punktquellen	14
4.4.2	Linearquellen	14
4.4.3	Fokussierte Quellen	15
4.5	Vor und Nachteile der Wellenfeldsynthese	16
4.5.1	Vorteile	16
4.5.2	Nachteile	16
5	Hardware	17
5.1	Die Wellenfeldsynthese-Anlage der HAW Hamburg	17
5.1.1	Die dritte Dimension der Wellenfeldsynthese-Anlage	18
5.2	Das Tracking-System der Wellenfeldsynthese-Anlage	18

5.3	Weitere genutzte Hardware	19
6	Software	21
6.1	WONDER	21
6.1.1	Bestandteile von WONDER	21
6.1.1.1	tWonder	21
6.1.1.2	cWonder	22
6.1.1.3	xWonder	22
6.1.2	Konfiguration	23
6.1.2.1	Hinzufügen der Einzellautsprecher	24
6.1.3	Kommunikation und Steuerung	24
6.2	MoWeC	25
6.2.1	Funktion	25
7	Entwicklung des Hörtests zur Lokalisationsgenauigkeit	26
7.1	Zielsetzung	26
7.2	Berechnung der Abweichung vom Ziel	26
7.3	GUI	29
7.4	Erfassen der Eingaben	30
7.5	Ändern der Quellenposition	31
7.6	Messdaten	31
8	Versuch zur Lokalisierungsgenauigkeit im dreidimensionalen Raum	33
8.1	Erwartete Ergebnisse	33
8.1.1	Beeinflussende Faktoren	33
8.2	Erweiterung der Wellenfeldsynthese-Anlage	34
8.2.1	Anpassung der Lautstärke und Klangfarbe	35
8.2.2	Auswahl des Testtones und der Positionierung	37
8.2.3	Versuchsablauf	38
8.3	Ergebnisse	38
8.3.1	Ergebnisse des Versuches	38
8.3.2	Diskussion der Ergebnisse	46
9	Schluss	48
9.1	Zusammenfassung	48
9.2	Fazit	48
9.3	Ausblick	49

1 Einleitung

Die Wellenfeldsynthese stellt ein Audiowiedergabeverfahren dar, welches ein optimales Hörerlebnis auf einer größeren Fläche als bei herkömmlichen Systemen ermöglicht. Durch Nachbildung eines Schallfeldes kann auf diese Weise eine reale Schallquelle simuliert werden. Dies erlaubt eine realitätsnahe örtliche Wahrnehmung von Schallquellen unabhängig vom Standort des Hörers innerhalb der Anlage. Meist sind solche Systeme aus Kosten- und Platzgründen nur zweidimensional realisiert. In dieser Arbeit soll untersucht werden, wie präzise virtuelle Schallquellen in der dritten Dimension einer simpel erweiterten Wellenfeldsynthese-Anlage lokalisiert werden.

1.1 Ziel der Arbeit

Ziel dieser Arbeit ist die Erweiterung der Wellenfeldsynthese-Anlage der HAW Hamburg um eine Höhenkomponente, indem 16 zusätzliche Einzellautsprecher in diese integriert werden. Aufbauend auf der Arbeit von Marvin Geitner [Gei15] soll geprüft werden, ob durch weitere Lautsprecher ein präzises Ergebnis in der akustischen Lokalisation von virtuellen Schallquellen in der Höhe erzielt werden kann. Dazu wird die von Marvin Geitner modifizierte Version der Wellenfeldsynthese-Software WONDER¹ genutzt, welche 3D-Quellenpositionierung ermöglicht. Des Weiteren soll ein automatisierter Hörtest Aufschluss über die Präzision der Darstellung virtueller Schallquellen geben. Die Ergebnisse sollen mit denen der Arbeit von Matthew Montag [Mon11] verglichen werden, der bei einem ähnlichen Versuchsaufbau eine Kombination aus Wellenfeldsynthese und Vector Base Amplitude Panning (VBAP) genutzt hat.

1.2 Themenabgrenzung

Erweiterungen durch Kombination der Wellenfeldsynthese mit anderen Verfahren zur Verbesserung der Höhenwahrnehmung werden nicht genutzt. Es wurden an WONDER keine Änderungen zum anpassen des Ausbreitungsverhaltens von Schallwellen in der dritten Di-

¹<http://www.sourceforge.net/projects/swonder>[18.12.2015]

mension vorgenommen, lediglich die Konfiguration wurde für den Versuch angepasst. In der Theorie sind die Abstände zwischen den Lautsprecherreihen so groß, dass die korrekte Wiedergabe einer Wellenfront auf Grund von räumlichem Aliasing nicht möglich ist. Trotz dieser Gegebenheit soll untersucht werden wie präzise der Höreindruck bei Vernachlässigung dieses Effektes ist.

1.3 Aufbau

Im ersten Teil der Arbeit soll ein Grundverständnis für die akustische Wahrnehmung des Menschen vermittelt werden. Es werden einige Phänomene aus der Psychoakustik und andere, die akustische Wahrnehmung beeinflussende Faktoren erklärt. Durch diese Grundkenntnisse soll ein Gefühl für die Komplexität der menschlichen Wahrnehmung sowie den Aufwand, welcher von Nöten ist, um die Sinne des Menschen zu überlisten, vermittelt werden. Des Weiteren sind diese Kenntnisse von Vorteil, um Auffälligkeiten aus den Versuchsergebnissen zu erklären. Darauf folgen Kapitel zu Wiedergabeverfahren sowie ein grundlegender Einblick in die Funktionsweise der Wellenfeldsynthese.

Im zweiten Teil befinden sich Kapitel zur genutzten Hard- und Software sowie deren Erweiterung im Zuge der Entwicklung des automatisierten Hörtests. Um die Erweiterungen nachvollziehen zu können, ist eine Grundkenntnis der genutzten Hard- und Software von Vorteil.

Im letzten Teil findet sich der abschließende Versuch zur Lokalisationsgenauigkeit von virtuellen Schallquellen sowie ein Überblick der gewonnenen Erkenntnisse, das Fazit und ein Ausblick über mögliche weitere Arbeiten zu diesem Thema.

2 Akustisch-räumliche Wahrnehmung des Menschen

Die räumliche Wahrnehmung des Menschen lässt ihn Geräusche schnell und meist präzise einer Richtung zuordnen. Doch schon kleine Änderungen und Unstimmigkeiten können die Wahrnehmung beeinflussen. In diesem Kapitel werden Eigenschaften des Menschen in seiner akustischen Wahrnehmung sowie die Auswirkungen dieser beschrieben. Diese Eigenschaften sind von Bedeutung, da sie Fehlerursachen in der Messung darstellen können und zur Auswahl des verwendeten Tonmaterials beitragen.

2.1 Akustischer Sweet Spot

Damit ein räumlicher Höreindruck entsteht, müssen mindestens zwei Wiedergabekanäle zur Verfügung stehen. Bei vielen Wiedergabeverfahren spielt die Position des Hörers eine große Rolle. So ist bei den üblichen Wiedergabeverfahren, wie der Stereophonie, darauf zu achten, dass sich der Hörer auf dem sogenannten Sweet Spot befindet. Als Sweet Spot wird der Punkt bezeichnet, an dem sich der Hörer im optimalen Schnittpunkt der Schallwellen befindet. Durch Ausrichtung der Lautsprecher in einem gedachten gleichschenkligen Dreieck, wie in [Abbildung 2.1](#) zu sehen, kann dieses erreicht werden. An diesem Schnittpunkt ist es dem Hörer möglich, einen räumlichen Höreindruck zu gewinnen. Befindet sich der Hörer im Sweet Spot, werden die Tonspuren so wahrgenommen wie vom Produzenten beabsichtigt. Auch bei herkömmlichen Mehrkanalverfahren entsteht ein Sweet Spot. Bei Mehrkanalverfahren werden die Lautsprecher üblicherweise an einem gedachten Kreis ausgerichtet, sodass sich der Sweet Spot im Mittelpunkt dieses Kreises befindet. Bei einigen 3D-Audio Verfahren kann der Sweet Spot ausgeweitet werden.

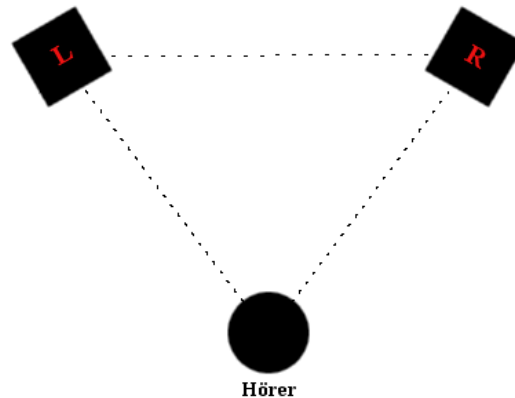


Abbildung 2.1: Der Hörer befindet sich im Sweet Spot der beiden Lautsprecher

2.2 Phantomschallquelle

Bei der Phantomschallquelle handelt es sich um ein psychoakustisches Phänomen, bei welchem der Hörer eine Schallquelle zwischen den Lautsprechern wahrnimmt. Da dort aber bei physischer Betrachtung keine reale Schallquelle vorliegt, wird diese als Phantomschallquelle bezeichnet. Vorausgesetzt der Hörer befindet sich auf dem Sweet Spot eines stereophonen Tonwiedergabesystems und die Lautsprecher geben synchron dasselbe Audiosignal aus, entsteht für den Hörer genau mittig zwischen den beiden Lautsprechern eine Phantomschallquelle [Ahr12]. Durch Variation der Lautstärke auf den verschiedenen Kanälen kann die Position der Phantomschallquelle nach Belieben zwischen den Lautsprechern verändert werden (siehe Abb. 2.2). Dieses gezielte Einsetzen von Phantomschallquellen wird auch als Stereo Panning bezeichnet. Auf diese Weise können einzelne Tonspuren auch zwischen den Lautsprechern platziert werden. [DDHW14]

2.3 Akustische Lokalisation

Menschen können durch akustische Lokalisation Geräusche oft genau einer Richtung zuordnen und die Entfernung zum Ursprung des Geräusches grob einschätzen. Diese Eigenschaften werden als Richtungshören und Entfernungshören bezeichnet. Bei binauralem Hören funktioniert die Lokalisation der Schallquellen durch das zeitversetzte Eintreffen der Schallwellen in dem jeweiligen Gehörgang. Durch die Zeit- und Pegeldifferenzen kann der Mensch lokalisieren, ob die Schallquelle links oder rechts von ihm liegt. Zwischen vorne und hinten kann mit Hilfe dieser Faktoren nicht unterschieden werden. Die Unterscheidung, ob etwas vor oder hinter dem Menschen liegt, wird durch den Einfallswinkel des Schalls in das Ohr bestimmt.

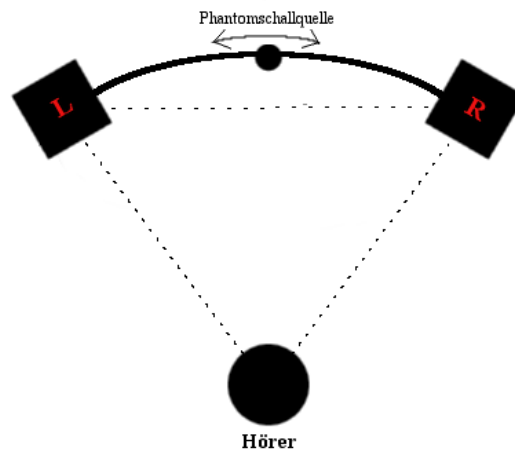


Abbildung 2.2: Die Phantomschallquelle kann durch Stereo Panning zwischen den Lautsprechern bewegt werden

Dieser wird durch Reflexionen im Außenohr beeinflusst. Zum Ermitteln der Entfernung werden Klangfarbe und spezifische Einfallsmuster in das Ohr mit bekannten Situationen verglichen. [BB08]

2.4 Wahrnehmung verschiedener Frequenzbereiche

Frequenzbereiche werden vom menschlichen Hörapparat verschieden wahrgenommen. Es sind im Allgemeinen Tonfrequenzen zwischen 20Hz und 20kHz wahrnehmbar. Mit steigendem Alter nimmt die Hörleistung meist in den Frequenzen über 1200Hz ab [Rob56]. Die Abbildung 2.3 zum ISO 226:2003 Standard [ISO03] zeigt die verschiedenen hörbaren Frequenzbereiche in Relation zu dem nötigen Schalldruck, um diese in derselben Lautstärke wahrzunehmen. Diese relativ zur Frequenz wahrgenommene Lautstärke wird als gehörrichtige Lautstärke bezeichnet. Aus der Abbildung ist erkennbar, dass tiefe Tonfrequenzen unter 500Hz in der Lautstärke deutlich schwächer wahrgenommen werden als Frequenzen, welche in höheren Frequenzbereichen liegen. Der empfindlichste Bereich im Bezug zur Lautstärke liegt demnach zwischen 600Hz und 5000Hz, darüber nimmt die Wirkung zunehmend ab.

Die Entfernung einer Schallquelle spielt bei größeren Entfernungen eine erhebliche Rolle. Hohe Frequenzen werden deutlich stärker von der Luft absorbiert als tiefe. So beträgt etwa der Luftschalldämpfungskoeffizient (dB/km) einer Frequenz von 63Hz nach DIN ISO 9613 [ISO99], $\alpha = 0,1$. Eine sehr tiefe Frequenz wird also kaum von der Luft absorbiert. Bei denselben Bedingungen und einer Frequenz von 8000Hz liegt der Koeffizient bei $\alpha = 76,6$. Pro Kilometer

Entfernung auf freier Fläche wird der Pegel also um 76,6dB abgeschwächt. Weitere Werte können aus Tabelle 2.1 entnommen werden. Bei der geringen Entfernung, die für die Hörtests genutzt wird, haben sich diese Werte als nicht relevant herausgestellt.

Tabelle 2.1: Luftschalldämmungskoeffizient (dB/km) α (20°C, 70% Luftfeuchtigkeit) [Wer09]

Frequenz in Hz	63	125	250	500	1000	2000	4000	8000
α	0,1	0,3	1,1	2,8	5,0	9,0	22,9	76,6

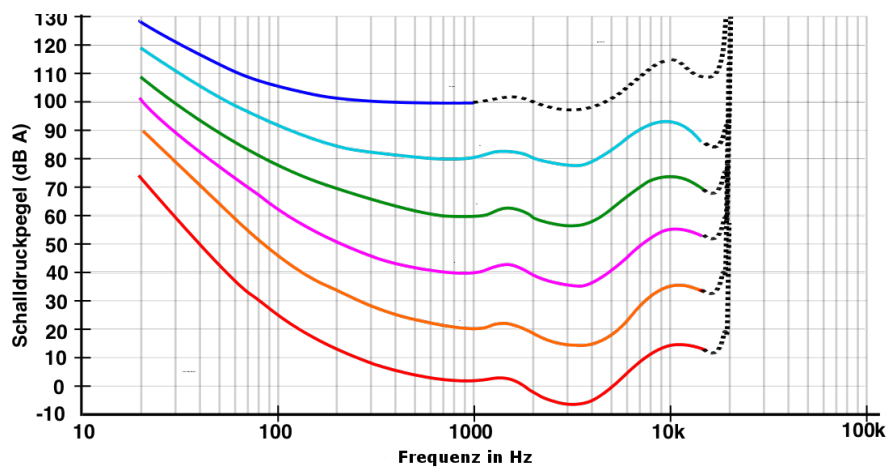


Abbildung 2.3: Gehörliche Lautstärke verschiedener Frequenzen¹

Nach Studien von Ernst [PKE14] werden bestimmte Tonfrequenzen von den Testpersonen auf verschiedenen Höhen wahrgenommen. So wurden tiefe Töne eher von unten und hohe Töne eher von oben wahrgenommen. Bei gemischten Frequenzen, wie etwa bei weißem Rauschen, welches ein sehr großes Frequenzspektrum besitzt, wurde der Ursprung deutlich präziser lokalisiert. Laut Wierstorf et al. sind hauptsächlich die Frequenzen unterhalb von 1kHz für die Lokalisation von Bedeutung [WRS13].

¹https://commons.wikimedia.org/wiki/File:Kurve_gleicher_lautst%C3%A4rke.svg [18.11.2015]

3 Audiowiedergabeverfahren

Heutzutage gibt es eine große Anzahl an Wiedergabeverfahren. Viele versuchen ein dreidimensionales Hörerlebnis zu erreichen, doch nur wenige sind in der Lage dies unabhängig von der Hörerposition umzusetzen. In diesem Abschnitt werden verschiedene Wiedergabeverfahren vorgestellt.

3.1 Monophonie

Die Monophonie stellt die ursprüngliche Form der Tonwiedergabe dar. Bei der monophonen Wiedergabe ist nur eine Tonspur sowie ein Lautsprecher involviert, sodass kein räumlicher Eindruck entstehen kann. Um einen räumlichen Klang zu erzeugen, ist es also nötig, mehr als einen Lautsprecher zu nutzen.

3.2 Stereophonie

Die Stereophonie beschreibt ein Wiedergabeverfahren mit zwei Kanälen. Für eine Stereowiedergabe wird ein rechter und ein linker Kanal benötigt. Mit Hilfe von Stereo Panning kann eine Phantomschallquelle zwischen den Lautsprechern bewegt werden, indem die jeweilige Tonspur mehr oder weniger einem Kanal zugewiesen wird. Mit Hilfe von psychoakustischen Effekten, wie der Phantomschallquelle, kann dennoch eine gewisse Tiefenwirkung erzielt werden. Probleme, die dabei auftreten können, sind etwa, dass Schallquellen nicht präzise platziert werden können und dass es schon bei kleinen Kopfbewegungen oder falscher Ausrichtung der Lautsprecher zu einer veränderten Wahrnehmung der Phantomschallquellen kommen kann. Die korrekte Wiedergabe ist also nur in Abhängigkeit mit der passenden Hörerposition zu erreichen. Als Erweiterung der Stereophonie gibt es verschiedene Mehrkanalverfahren, welche mehr Lautsprecher nutzen, aber auf demselben Prinzip basieren.

Die bisher erwähnten Wiedergabeverfahren sind meistens für den Privatanwender erschwinglich und dadurch bereits in vielen Haushalten anzutreffen. Was alle bisherigen Wiedergabeverfahren gemeinsam haben, ist die Wiedergabe von Tonmaterial, welches für einen

bestimmten Lautsprecher vorgesehen ist. Es werden Schallwellen erzeugt, welche nur den jeweiligen Lautsprechern zuzuordnen sind.

3.3 3D-fähige Wiedergabeverfahren/Stand der Technik

Die vorher benannten Verfahren basieren auf demselben Prinzip und stellen Weiterentwicklungen und Erweiterungen auf Basis der Monophonie dar. Im Folgenden werden Verfahren erwähnt, die darauf ausgelegt sind, ein möglichst reales räumliches Klangbild zu erzeugen.

3.3.1 Wellenfeldsynthese

Die Wellenfeldsynthese basiert auf der Idee der Imitation der Schallwellen einer Schallquelle mit Hilfe von Lautsprecher-Arrays und dem Huygensschen Prinzip. Durch die computergesteuerte Imitation der Schallwellen eines Objektes soll erreicht werden, dass innerhalb des Bereiches einer Wellenfeldsynthese-Anlage der Ton möglichst realitätsnah wahrgenommen wird. Durch Nutzung dieser Technik ist die Wellenfeldsynthese nicht auf psychoakustische Effekte, wie etwa Phantomschallquellen (siehe Abschnitt 2.2), angewiesen. Mit Hilfe der Wellenfeldsynthese können virtuelle Schallquellen nach Belieben inner- oder außerhalb eines Raumes platziert werden. Selbst wenn der Hörer sich innerhalb der Anlage bewegt, bleibt der Tonursprung für den Hörer am gleichen Ort.

In dem folgenden Kapitel zur **Wellenfeldsynthese** wird detaillierter auf die Funktionsweise dieser eingegangen.

3.3.2 VBAP

Das Vector Base Amplitude Panning arbeitet mit virtuellen Schallquellen, die über Phantomschallquellen reproduziert werden. VBAP ist vergleichbar mit einem dreidimensionalem Stereo Panning. Durch einen Berechnungsalgorithmus, basierend auf Triangulation, werden die wiederzugebenden Signale für jeweils benachbarte Lautsprecherpaare bestimmt. Die von VBAP erzeugten Phantomschallquellen sind deutlich präziser als bei anderen Verfahren, die diese nutzen. Die Phantomschallquellen können vertikal und horizontal zwischen den Lautsprechern platziert werden. Der größte Vorteil gegenüber der Wellenfeldsynthese ist, dass die Lautsprecher willkürlich dreidimensional angeordnet werden können. Nachteilig dagegen ist die Nutzung der Phantomschallquellen, da der Sweet Spot stark begrenzt ist. [Pul97]

3.3.3 Binaurales 3D-Audio

Eine weitere Möglichkeit, ein dreidimensionales Klangbild zu erzeugen, sind binaurale Verfahren. Der Begriff binaural steht dabei für das Hören mit beiden Ohren, wobei nur bei Verwendung von Kopfhörern eine korrekte Raumdarstellung gegeben ist. Binaurale Aufnahmen werden mit zwei Mikrofonen realisiert, deren Platzierung der Position der menschlichen Ohren entspricht. Dafür können Kunstköpfe genutzt werden, bei denen die Mikrofone in den Ohren integriert sind oder es werden spezielle Mikrofone genutzt, die in den eigenen Gehörgang eingesetzt werden. Auf diese Weise kann sogar ein individueller Gehörgang berücksichtigt werden. Bei der Wiedergabe dieser Tonspuren auf Kopfhörern entspricht die Aufnahme- und die Wiedergabeposition und die Klänge wurden in der Charakteristik so aufgezeichnet, wie sie auch von Menschen wahrgenommen werden. Dies führt zu einem starken Raumgefühl. Um ein optimales Ergebnis zu erhalten, muss die Kopfbewegung verfolgt werden, um ggf. in Echtzeit Anpassungen zu errechnen. [DDHW14]

3.3.4 Ambisonics

Der grundlegende Ansatz von Ambisonics ist vergleichbar mit dem der Wellenfeldsynthese. Bei beiden sollen Schallwellen realitätsgetreu nachgebildet werden. Die Lautsprecher sind beliebig in einer halbkugelförmigen Ausrichtung aufzustellen. Die erzeugten Schallwellen setzen sich aus den Toninformationen der genutzten Quellen sowie der Ausbreitung in verschiedene Richtungen zusammen. Erhöht man die Anzahl der Richtungen, wird von Higher Order Ambisonics gesprochen. Je höher die Anzahl der Richtungen, desto besser ist das klangliche Ergebnis sowie die räumliche Wahrnehmung. Dies ist jedoch auch mit einem erheblichen Anstieg im Rechenaufwand verbunden. Higher Order Ambisonics ist der Wellenfeldsynthese sehr ähnlich [HS05] und der Sweet Spot kann sich ab einem bestimmten Grad auflösen. Um einen Höheneffekt zu erzeugen, werden mindestens drei Lautsprecher benötigt. Mit vier oder mehr Lautsprechern kann ein Wellenfeld korrekt wiedergegeben werden. Der grundlegende Unterschied zur Wellenfeldsynthese liegt bei den mathematischen Prinzipien: Während Ambisonics auf Kugelflächenfunktionen basiert, beruht die Wellenfeldsynthese auf dem Kirchoff-Helmholtz-Integral [Wei09].

3.3.5 Zukunft der dreidimensionalen Audiowiedergabeverfahren

Dreidimensionale Audiowiedergabeverfahren gewinnen durch die großen Fortschritte in der Virtual-Reality Sparte immer mehr an Bedeutung. Vor allem im Zusammenspiel mit Virtual-Reality-Brillen, welche seit kurzem auch für die Heimanwendung erschwinglich sind, ist dies

unerlässlich um ein möglichst immersives Erlebnis zu gewährleisten. Zur Zeit konzentriert sich die Entwicklung vornehmlich auf die visuelle Wahrnehmung. Auch für die akustische Wahrnehmung gibt es schon 3D-Audio Kopfhörer, welche die gesamte Technik für binaurales 3D-Audio in deren Gehäuse integrieren [Oss16, 3DS15].

Im kommerziellen Sektor besteht großes Interesse, den Kunden ein möglichst immersives räumliches Klangbild zu bieten. Kinos zum Beispiel interessieren sich immer mehr für 3D-Audio-Verfahren. Erste Kinos sind schon mit Wellenfeldsynthese-Anlagen ausgerüstet¹. In diesem Bereich tendiert die Entwicklung immer weiter zu einer dritten Dimension.

¹<http://www.iosono-sound.com/references/cinema/>[25.04.2016]

4 Wellenfeldsynthese

Um ein Grundverständnis für die Funktionsweise der genutzten Wellenfeldsynthese-Anlage zu erlangen und mögliche Probleme bei Versuchsaufbauten besser zu verstehen, wird hier ein Einblick in die Funktionsweise der Wellenfeldsynthese gegeben.

Bei der Wellenfeldsynthese handelt es sich um ein Verfahren, dessen theoretische Grundlagen 1988 von A.J. Berkhout [Ber88] an der Technischen Universität Delft veröffentlicht wurden. Die Wellenfeldsynthese kann durch Berechnen der Wellenfront einer virtuellen Schallquelle ein realitätsnahes Schallfeld an der Lautsprecherfront erzeugen. Als Vorläufer der Wellenfeldsynthese gab es die Idee des akustischen Vorhangs (siehe Abb. 4.1). Dieser funktioniert nach demselben Prinzip wie die Wellenfeldsynthese, nimmt aber die Schallwellen für jeden Kanal einzeln auf und gibt diese auf dem entsprechenden Kanal wieder. So müsste eine riesige Bandbreite an Tonsignalen übertragen werden und es würde nur auf einem passend zu den Mikrofonen aufgestellten Wiedergabesystem funktionieren [Oel07]. Durch die enorme Bandbreite, die benötigt wird, um so viele Kanäle zu übertragen, hat sich die Idee des akustischen Vorhangs nicht durchgesetzt. Die Wellenfeldsynthese löst das Problem der Bandbreite und der starren Aufstellung durch die virtuellen Schallquellen, welche eine beliebige Aufstellung der Lautsprecher erlauben.

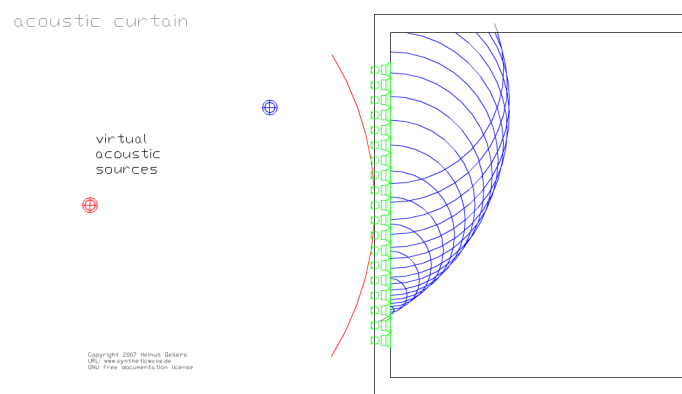


Abbildung 4.1: Der akustische Vorhang. Die von den Schallquellen erzeugten Signale werden in derselben Reihenfolge wiedergegeben, in der sie bei den Mikrofonen eintreffen.

4.1 Funktionsweise der Wellenfeldsynthese

Die Wellenfeldsynthese basiert auf dem Huygensschen Prinzip [HO03], das von Christiaan Huygens beobachtet wurde. Das Huygenssche Prinzip besagt, dass jeder Punkt an einer Wellenfront als Ursprung einer weiteren Elementarwelle betrachtet werden kann. Dadurch ist es möglich, eine eintreffende Wellenfront mit der Überlagerung von Elementarwellen nachzubilden. Dieses Prinzip macht sich die Wellenfeldsynthese in einer abgewandelten Form zunutze, indem die Wellenfront einer Schallquelle und deren Ausbreitung im Raum computergesteuert berechnet wird. Sobald diese virtuelle Wellenfront die Position der aufgestellten Lautsprecher kreuzt, werden die Elementarwellen an den jeweiligen Lautsprechern erzeugt (siehe Abb. 4.2). Damit dies funktioniert, muss die Position der einzelnen Lautsprecher im Raum bekannt sein. Als genaue Position wird die Mitte der Lautsprechermembran angegeben, da sich dort der Ursprung der Elementarwellen befindet. [DDHW14]

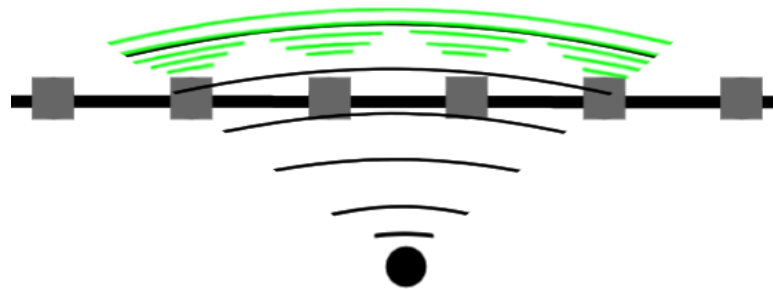


Abbildung 4.2: Erzeugung der Elementarwellen über die Lautsprecher bei Eintreffen der virtuellen Wellenfront

4.2 Aufbau einer Wellenfeldsynthese-Anlage

Eine Wellenfeldsynthese-Anlage besteht optimalerweise aus einer sehr großen Anzahl an Lautsprechern mit möglichst kleinem Abstand, um entsprechend viele Elementarwellen erzeugen zu können und somit eine realistische Wellenfront nachzubilden. Die Lautsprecher sind für das optimale Hörerlebnis meist auf Kopfhöhe oder etwas oberhalb angebracht, da auf diese Weise die Schallwellen auf dem direktesten Weg vom Menschen wahrgenommen werden können. Der Aufbau kann auf verschiedene Arten erfolgen (siehe Abb. 4.3). Es existieren runde, lineare und rechteckige Aufbauten. Die Wellenfeldsynthese wird in unterschiedlich ausgedehnten Aufbauten genutzt. Kommerzielle Anlagen werden oft mit Lautsprechermodulen ausgeliefert, in denen viele kleine Lautsprecher montiert sind. Dadurch wird gewährleistet, dass die

Wellenfront aus einer möglichst großen Menge von Elementarwellen zusammengesetzt wird und daraus resultierend eine glattere Wellenfront erzeugt werden kann. Neben Lautsprechern wird auch eine Berechnungseinheit mit genügend Leistung benötigt, um das Signal für die einzelnen Lautsprecherkanäle zu berechnen. Um Hall und Interferenzen zu vermeiden, sollte der Raum akustisch optimiert sein. Das bedeutet z.B. Schalldämmung der Wände, der Decke und des Bodens mit Akustikschaumstoff oder anderen geeigneten Dämmstoffen. Bei dem Aufeinandertreffen von unerwünschten Schallreflexionen und Elementarwellen kann es zu destruktiven oder konstruktiven Interferenzen kommen. Diese können die Elementarwellen abschwächen oder verstärken und somit die Wellenfront beeinträchtigen.

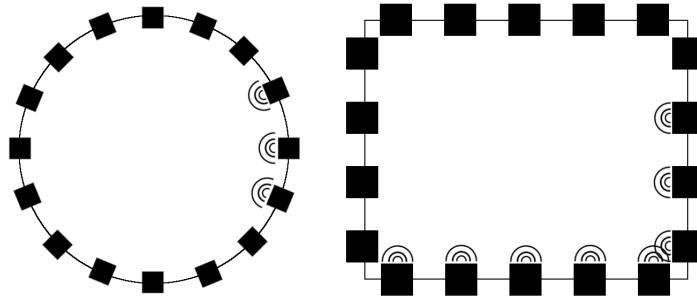


Abbildung 4.3: Mögliche Aufbauformationen einer Wellenfeldsynthese-Anlage.

4.3 Technische Umsetzung

Die Berechnungseinheit berechnet für jede Schallquelle, welche Lautsprecher für deren Schallerzeugung involviert sind und mit welchem Lautstärkepegel diese angesprochen werden. Je weiter der Lautsprecher von der Schallquelle entfernt ist, desto kleiner ist der auszubehende Schallpegel. Außerdem wird für jeden involvierten Lautsprecher die theoretische Laufzeit der Ausbreitung des Signals von der Schallquelle bis zum Lautsprecher berechnet. Daraus resultiert der Zeitpunkt, an dem der individuelle Lautsprecher mit dem Tonsignal gespeist wird. Die Kombination aus Verzögerung und Lautstärkepegel des Signals der einzelnen Lautsprecher lässt daraufhin das Schallfeld entstehen. Als ausführliche Literatur empfiehlt sich das Buch *Analytic Methods of Sound Field Synthesis* [Ahr12].

4.4 Darstellungsformen von Audioquellen

Mit Hilfe der Wellenfeldsynthese können drei verschiedene Arten von Quellen erzeugt werden. In folgendem Abschnitt werden diese kurz erläutert.

4.4.1 Punktquellen

Punktquellen stellen die üblicherweise genutzte Art von Schallquellen in der Wellenfeldsynthese dar, da die entstehenden Wellenfronten den realen gleichen. Bei Punktquellen breitet sich der Schall von der Quellenposition ausgehend, wie in [Abbildung 4.4](#) zu erkennen, ringförmig aus. Korrekt wäre eine kugelförmige Ausbreitung, WONDER berechnet diese im dreidimensionalen Raum aber zylinderförmig. Die physikalische Ausbreitung der Schallwellen durch die Lautsprecher erfolgt jedoch kugelförmig. Dies ist maßgebend für die Form der Schallwellen in [Abbildung 4.2](#).

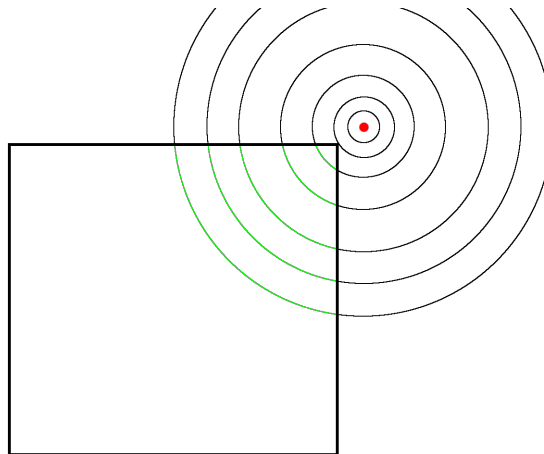


Abbildung 4.4: Die Schallwellen breiten sich von der Punktquelle ringförmig aus. Der von der WFS-Anlage zu synthetisierende Bereich ist grün markiert

4.4.2 Linearquellen

Linearquellen breiten sich gleichmäßig ohne Krümmung über eine Fläche aus, sodass nur die Richtung, aber nicht die genaue Position wahrgenommen werden kann. In der Realität können Linearquellen (siehe [Abb 4.5](#)) ähnliche Wellen bei sehr weit entfernten Quellen entstehen, wenn der Krümmungsradius sehr groß und dadurch von einem Menschen kaum mehr wahrzunehmen ist. [[SA10b](#)]

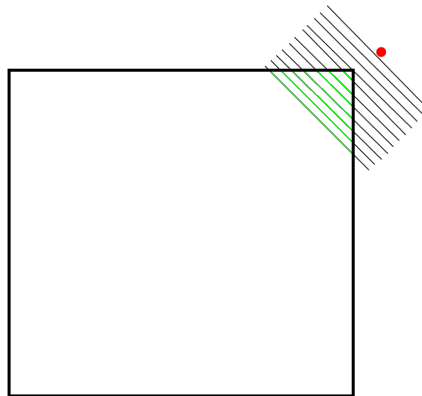


Abbildung 4.5: Die Schallwellen breiten sich von der Linearquelle linear aus. Der von der WFS-Anlage zu synthetisierende Bereich ist grün markiert

4.4.3 Fokussierte Quellen

Als fokussierte Quellen werden Punktquellen bezeichnet, die sich innerhalb des Raumes der Wellenfeldsynthese-Anlage befinden. Im normalen Betrieb funktionieren diese nur wie gewünscht, wenn sich der Hörer außerhalb des Bereiches zwischen den aktiven Lautsprechern und der Quelle befindet. Dieses Problem kann durch aktive Verfolgung der Kopfposition des Hörers umgangen werden [Chr14]. Durch die bekannte Kopfposition kann die Quelle bei Überschreitung der Grenze zwischen Lautsprecher-Array und Quelle vom gegenüberliegenden Array erzeugt werden. Fokussierte Quellen werden an der Lautsprecherfront gebildet wie inverse Punktquellen.

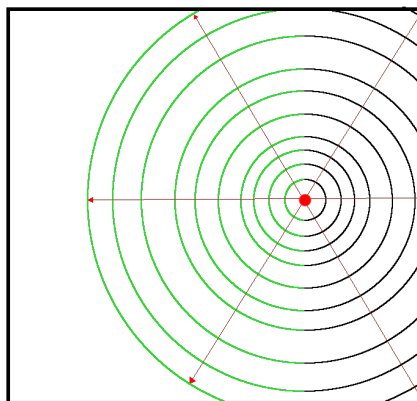


Abbildung 4.6: Die Wellenfeldsynthese-Anlage erzeugt eine konkave Wellenfront (schwarz), welche an der fokussierten Quelle umbricht und konvex wird (grün). Die roten Linien deuten den Pfad der Schallwellen an.

Es wird ein konkaves Wellenfeld erzeugt, welches an den Koordinaten der Quelle in eine konvexe Wellenfront umschlägt. Die entstehende konvexe Wellenfront entspricht der Wellenfront einer Punktquelle im Raum (siehe auch Abb 4.6). [SA10a]

4.5 Vor und Nachteile der Wellenfeldsynthese

Folgend finden sich Vor- und Nachteile der Wellenfeldsynthese gegenüber herkömmlichen Wiedergabeverfahren.

4.5.1 Vorteile

Durch Wellenfeldsynthese können ortsstabile Quellen erzeugt werden, welche unabhängig von der Position im Raum an einer bestimmten Position wahrgenommen werden. Der Sweet Spot erstreckt sich fast über die gesamte Fläche der Wellenfeldsynthese-Anlage. Dies ermöglicht freies Bewegen des Hörers innerhalb dieser Fläche und eine gleichzeitige Teilnahme von einer größeren Anzahl von Hörern an der korrekten Darstellung der Audiosignale. Durch den Verzicht auf psychoakustische Effekte ist eine exakte Positionierung der Quellen möglich. Dies wiederum wirkt sich positiv auf die Lokalisierung der Quellen aus. Durch fokussierte Quellen können Töne auch innerhalb der Wellenfeldsynthese-Anlage positioniert werden.

4.5.2 Nachteile

Auch wenn die Wellenfeldsynthese viele Vorteile in der Räumlichkeit und Lokalisierbarkeit bietet, gibt es vor allem baulich bedingt viele Nachteile. Durch die große Anzahl benötigter Hardware in Form von Lautsprechern, Recheneinheiten, Dämm- und Befestigungsmaterial sind die Kosten sowie der benötigte Platz sehr groß. In der Theorie wird von einem unendlichen Band von Lautsprechern mit unendlich kleinem Abstand ausgegangen. Da ein unendliches Lautsprecherband nicht realisierbar ist, existiert in Relation zum Lautsprecherabstand und der Hörposition eine Grenzfrequenz. Die Grenzfrequenz wird als Nyquist-Frequenz bezeichnet und entspricht der doppelten Signalfrequenz des abzutastenden Signals. Liegen Frequenzanteile des abzutastenden Signals über der Nyquist-Frequenz, entstehen räumliche Alias-Effekte. Räumliche Alias-Effekte können sich durch eine unregelmäßige Wellenfront oder falsch ausgerichtete Elementarwellen erkennbar machen [Wei09]. Es existiert ein Patent, welches die Alias-Effekte durch unregelmäßige Anordnung der Lautsprecher vermindern soll [Oel10]. Beugungseffekte an den Enden des Lautsprecherbandes können die Wellenfront beeinträchtigen. Dies geschieht durch die unvollständige Wellenfront, die mit der Elementarwelle des letzten Lautsprechers endet [DDHW14].

5 Hardware

Die Hauptkomponenten des Wellenfeldsynthese-Labors der HAW Hamburg sind die Wellenfeldsynthese-Anlage selbst sowie das Tracking-System (siehe Abb. 5.1). Dieses Kapitel gibt einen Einblick in die genutzten Systeme des Wellenfeldsynthese-Labors, in dem die Versuche zu dieser Arbeit durchgeführt wurden.

5.1 Die Wellenfeldsynthese-Anlage der HAW Hamburg

Die Wellenfeldsynthese-Anlage der HAW Hamburg besteht aus 26 Lautsprechermodulen der Firma Four Audio, Modell 28-243, welche jeweils mit 26 Lautsprechern, in drei Reihen, zwei davon Tieftöner und 24 Breitbandlautsprecher, bestückt sind. Die Lautsprecher werden zu 208 Kanälen zusammengefasst, pro Modul acht Kanäle. Zur Steuerung wird eine Four Audio WFS Rendering Unit genutzt. Diese besteht aus drei Berechnungseinheiten, zwei davon als Node für das Rendering und die dritte als zentraler Server zur Steuerung des Systems. Diese Rechenleistung wird benötigt, um alle 208 Kanäle in Echtzeit zu berechnen. Zusätzlich ist ein Mac Pro als benutzerfreundliches Interface vorhanden. Über diesen können Aktionen auf dem Server ausgelöst und Multimediaanwendungen bedient werden. Zur Übertragung der Audiosignale ist das System über das Dante Netzwerk verbunden. Dante ist ein Produkt der Firma Audinate¹ und steht für Digital Audio Network Through Ethernet. Dieses Netzwerkprotokoll ist durch die niedrige Latenz und hohe Bandbreite besonders gut für die simultane Übertragung vieler Audiosignale geeignet. Ein weiterer Vorteil gegenüber analoger Übertragung ist die geringere Störanfälligkeit. Zur Nutzung dieses Protokolls wird ein spezielles Dante-fähiges Netzwerkinterface benötigt. [Foh13]

¹<https://www.audinate.com/>[05.03.2016]

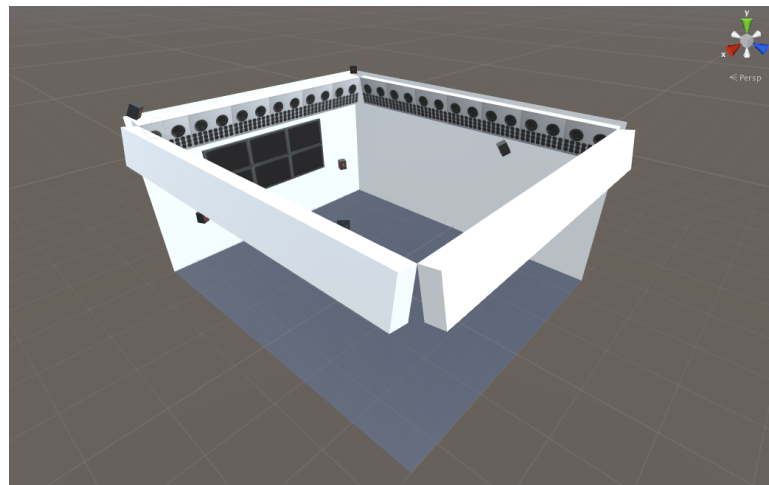


Abbildung 5.1: Aufbau der Wellenfeldsynthese-Anlage der HAW Hamburg.

5.1.1 Die dritte Dimension der Wellenfeldsynthese-Anlage

Das Prinzip der Wellenfeldsynthese ist von Anfang an auf eine dreidimensionale Nutzung ausgelegt. Aus Kosten- und Platzgründen wird diese jedoch oft vernachlässigt. Eine Erweiterung durch zusätzliche Wellenfeldsynthese-Module ist aus diesen Gründen nicht praktikabel. In der Software, welche in dieser Arbeit genutzt wird, entstehen bei der Darstellung der Wellenfront kleine Fehler in der Höhe, da die Ausbreitung der Schallwellen als Kreis berechnet wird und somit zylinderförmig wiedergegeben wird. Um ein zutreffenderes Ergebnis zu erhalten, muss die Ausbreitung der Schallwellen als Kugel berechnet werden, da sonst eine Wellenfront entsteht, die nur horizontal korrekt dargestellt wird. Dieser Effekt wird in dieser Arbeit nicht korrigiert. Es handelt sich also um eine Annäherung an die Wellenfeldsynthese.

5.2 Das Tracking-System der Wellenfeldsynthese-Anlage

Das Tracking-System der Firma ARTTRACK, im Wellenfeldsynthese-Labor der HAW Hamburg, besteht aus sechs Kameramodulen und einer Berechnungseinheit. Bei dem System handelt es sich um ein markerbasiertes System, welches mit Infrarotstrahlung arbeitet. Markerbasierte Systeme sind optische Trackingverfahren, welche spezielle Kameras nutzen, um Reflexionen im Raum zu erkennen. Bei passiven Markern handelt es sich um besonders beschichtete Objekte, welche eintreffendes Licht gut reflektieren. Aktive Marker strahlen die benötigte Wellenlänge des Infrarotlichtes mit Hilfe von Leuchtdioden eigenständig aus. Die von den Kameras aufgenommenen Bilder werden auf Punkte untersucht, die durch die Reflexion oder Ausstrahlung

von Infrarotlicht entstehen. Daraufhin werden die Bilder der Kameras abgeglichen und auf Übereinstimmungen untersucht. Wird ein Marker von mehr als zwei Kameras erkannt, kann mittels eines Triangulations-Algorithmus die Position im Raum bestimmt werden. Die Position wird genauer je mehr Kameras denselben Punkt erkennen. Durch Nutzung von Markern in eindeutigen geometrischen Konstellationen können diese individuell zugeordnet werden. Mit Hilfe dieser sogenannten Targets kann auch die Ausrichtung der Marker-Konstellation bestimmt werden. Diese wird über die Lage in einem dreidimensionalen Koordinatensystem bestimmt. Jedes Target wird dabei wie ein eigenständiges dreidimensionales Koordinatensystem gehandhabt, welches sich im Raumkoordinatensystem bewegt. Die errechneten Daten werden per WLAN-Multicast versendet.

Für ausführlichere Informationen zu dem Tracking-System der HAW Hamburg eignet sich das Kapitel „Trackingverfahren und -systeme“ aus der Bachelorthesis von Malte Nogalski [Nog12].

5.3 Weitere genutzte Hardware

Zur Verfügung standen das ARTTRACK-Tracking-System und die Four Audio-Wellenfeldsynthese-Anlage. Folgende Hardware wurde zusätzlich zu den im Wellenfeldsynthese-Labor vorhandenen Anlagen für den Testaufbau genutzt:

- 16x Adam Audio A7x 2-Wege Aktiv-Bassreflexlautsprecher ² mit Verkabelung
- 8x Stative mit einer einstellbaren Höhe von 1,27 Meter
- 2x Focusrite RedNet 1 Ethernet Audio-Interfaces ³
- AKG C3000B Großmembran-Kondensatormikrofon ⁴
- Focusrite Scarlett 18i8 USB 2.0 Audio-Interface mit Mikrofon-Vorverstärker ⁵
- Logitech R400 Presenter ⁶

Für das Tracking-System wurden folgende Targets genutzt:

- TT1 als Zeiger, siehe Abb. 5.2
- HT23 als Kalibrierungsziel, siehe Abb. 5.2
- GT04 zur Bestimmung der Kopfposition, siehe Abb. 5.2

²<http://www.adam-audio.com/de/pro-audio/products/a7x/description>[20.02.2016]

³<http://focusrite.de/ethernet-audio-interfaces/rednet-1>[20.02.2016]

⁴<http://www.ake.com/de/p/c3000>[20.02.2016]

⁵<http://focusrite.de/usb-audio-interfaces/scarlett-18i8>[20.02.2016]

⁶<http://business.logitech.com/de-de/product/wireless-presenter-r400-business>[20.02.2016]



Abbildung 5.2: TT1, HT23 und GT04 Marker

6 Software

In diesem Kapitel wird ein Überblick über die für diese Arbeit relevanten Softwarekomponenten gegeben.

6.1 WONDER

Die Open-Source Wellenfeldsynthese-Software WONDER berechnet die benötigten Informationen zur Erzeugung eines Schallfeldes und leitet Toninformationen gezielt an die Lautsprecher im System weiter. Bei WONDER handelt es sich um eine verteilte Anwendung, sodass mehrere Recheneinheiten innerhalb eines Netzwerkes genutzt werden können. Durch die Verteilung kann die Rechenlast gegenüber einem einzelnen Gerät verringert werden. WONDER wurde von Marije Baalman an der TU Berlin entwickelt und 2004 vorgestellt [BP04]. WONDER stellt die zentrale Softwarekomponente der Wellenfeldsynthese-Anlage der HAW Hamburg dar. Um die Vorgänge im System, die Anpassungen im System zur Erweiterung um die zusätzlichen Lautsprecher und einige Aufrufe aus dem Programmcode nachvollziehen zu können, folgt ein Einblick in die Funktionsweise der genutzten WONDER Komponenten.

6.1.1 Bestandteile von WONDER

Die Wellenfeldsynthese-Software WONDER besteht aus folgenden Bestandteilen:

- **twonder** • **cwonder** • jfwonder • scoreplayer
- **xwonder** • fwonder • qfwonder • tracker

An der Wellenfeldsynthese-Anlage der HAW Hamburg werden nur die folgenden Bestandteile von WONDER genutzt:

6.1.1.1 tWonder

tWonder ist für die Berechnungen innerhalb des WONDER Systems zuständig. Die extern zugeführten Tonspuren werden verarbeitet und die Laufzeit und Lautstärke der einzelnen Tonsignale werden berechnet. Die so berechnete Verzögerung und Lautstärke des Tonsignals

wird daraufhin auf die jeweils passenden Lautsprecher übertragen, um die Wellenfront zu erzeugen.

6.1.1.2 cWonder

cWonder dient als Kommunikations- und Kontrollserver. Über cWonder läuft die gesamte Kommunikation zwischen den einzelnen Komponenten. Die Kommunikation erfolgt über Open Sound Control-Nachrichten. Die Komponenten melden sich mit einer OSC-Nachricht bei cWonder an und werden mit den für sie relevanten Informationen versorgt.

6.1.1.3 xWonder

xWonder ist die graphische Oberfläche (siehe Abb 6.1), mit welcher die Quellen erstellt, entfernt oder positioniert werden können. Außerdem kann die Art der Quelle sowie weitere Einstellungen zum Tonverhalten ausgewählt werden. Durch die in xWonder vergebene ID ist jede Quelle eindeutig identifizierbar. Auch die Targets aus dem Tracking-System werden als Quellen behandelt und sind somit in xWonder sichtbar.

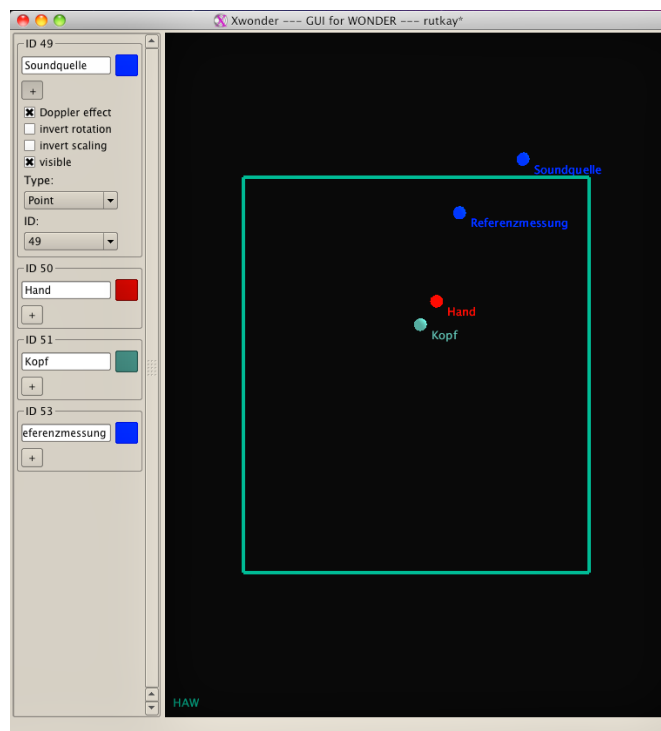


Abbildung 6.1: Screenshot der xWonder Oberfläche

6.1.2 Konfiguration

Die Einstellungen im Programm werden über Konfigurationsdateien vorgenommen, welche üblicherweise im XML- oder conf-Format vorliegen. Diese liegen zentral auf dem Server, der cWonder ausführt. Dort werden unter anderem die Lautsprecheranzahl, die Aufstellung der Lautsprecher und die zuständigen Knoten für die Berechnungen der jeweiligen Informationen für die Lautsprecher auf den Nodes festgelegt. Die einzelnen Komponenten können sich von dort die benötigten Dateien abrufen. Als Beispiel ist in Listing 6.1 ein Auszug aus der Datei cluster.conf zu sehen, welche Angaben zum Aufbau des Systems und der Audiokonfiguration enthält.

Listing 6.1: cluster.conf

```
1 # this file contains shared information for the wonder modules
2 USER_NAME='wfs'                #Name des Systems
3 CLUSTER_NODES='n101 n102';    #Namen der Knoten
4 NUM_SOURCES=64;               #Max. Anzahl an Tonquellen
5 CWONDER_IP='192.168.3.1'      #IP Adresse des cWonder Server
6 JACK_BLOCKSIZE=128;          #Blockgröße des Audiostreams
7 JACK_SAMPLERATE=48000;       #Audio-Samplerate
```

Die twonder_speakerarray.xml Dateien sind wie in 6.2 aufgebaut und enthalten die Informationen zu den Lautsprecherpositionen. Im ersten Segment mit der ID 27 wird ein Lautsprecher angelegt, welcher auf der X-Achse des Raumes 2.34 Meter nach links verschoben wurde, wobei die Membranmitte auf der Z-Achse bei 0,2 Meter steht und genau an der Y-Achse positioniert ist. Es ist auch möglich innerhalb eines Segmentes mehrere Lautsprecher in die Konfiguration einzutragen, sowie es schon für die vorhandenen Module umgesetzt wurde.

Listing 6.2: twonder_speakerarray.xml Aufbau

```
1 <!DOCTYPE speakerarray SYSTEM "twonder_speakerarray.dtd">
2
3 <speakerarray>
4   <segment id="27"
5     numspeak="1" winwidth="0"
6     startx= "-2.3385935" starty= "0.0" startz= "0.2"
7     endx=   "-2.3385935" endy=   "0.0" endz=   "0.2"
8     normalx= "0.0"         normaly="1.0" normalz="0.0"/>
9 </speakerarray>
```

6.1.2.1 Hinzufügen der Einzellautsprecher

Um die 16 Einzellautsprecher im System bekannt zu machen, wurden zwei weitere `twonder_speakerarray.xml` Dateien erzeugt. Diese enthalten jeweils die Positionen zu acht der Lautsprecher im Format wie es aus Listing 6.2 zu entnehmen ist. Sowohl aus Mangel an Anschlüssen für die Lautsprecher als auch um die Last im System gleichmäßig zu verteilen, wird jeweils die Hälfte der Lautsprecher von einem Node verwaltet. Des Weiteren müssen die `twonder_n101.conf` und `twonder_n102.conf` Dateien entsprechend um die Anzahl der Lautsprecher erweitert werden.

6.1.3 Kommunikation und Steuerung

WONDER wird zum Großteil über Shell-Skripte gesteuert, welche auf dem Server liegen. Diese Skripte enthalten Abfolgen von Anweisungen, um bestimmte Module zu starten oder zu beenden und andere Aktionen auszuführen. Die interne Kommunikation erfolgt über OSC-Nachrichten (siehe Abb. 6.2). Als Beispielskript befindet sich anbei ein verkürzter Auszug aus dem `start-wonder.sh` Skript mit Kommentaren zum Ablauf:

Listing 6.3: `start-wonder.sh`

```
1 #!/bin/bash                                #Interpreter auswählen
2 SCRIPTDIR='dirname $0'                      #Verzeichnis ermitteln
3 source ${SCRIPTDIR}/configs/cluster.conf    #Einlesen der conf Datei
4
5 ${SCRIPTDIR}/jack/start-all-jackd.sh        #Ausführen des Jack Startskripts
6
7 ${SCRIPTDIR}/cwonder/start-cwonder.sh      #cWonder Startskript
8 sleep 1                                     #Eine Sekunde warten
9
10 ${SCRIPTDIR}/twonder/start-all-twonder.sh #tWonder Startskript
```

Die verwendeten OSC-Nachrichten sind folgendermaßen aufgebaut:

Listing 6.4: Aufbau von OSC-Nachrichten

```
1 NetAddr ("IP ADRESSE", PORT) .sendMsg ("/PFAD/FUNKTION", VARIABLEN);
```

Eine übliche Anwendung, wie z.B. die Neupositionierung einer Quelle, kann wie folgt aussehen:

Listing 6.5: OSC-Nachricht zum Neupositionieren einer Quelle

```
1 NetAddr("192.168.3.1", 58100) //ID,X,Y,Z (Variablen)
2 .sendMsg("/WONDER/source/position3D", 1, 3, 0, 2);
```

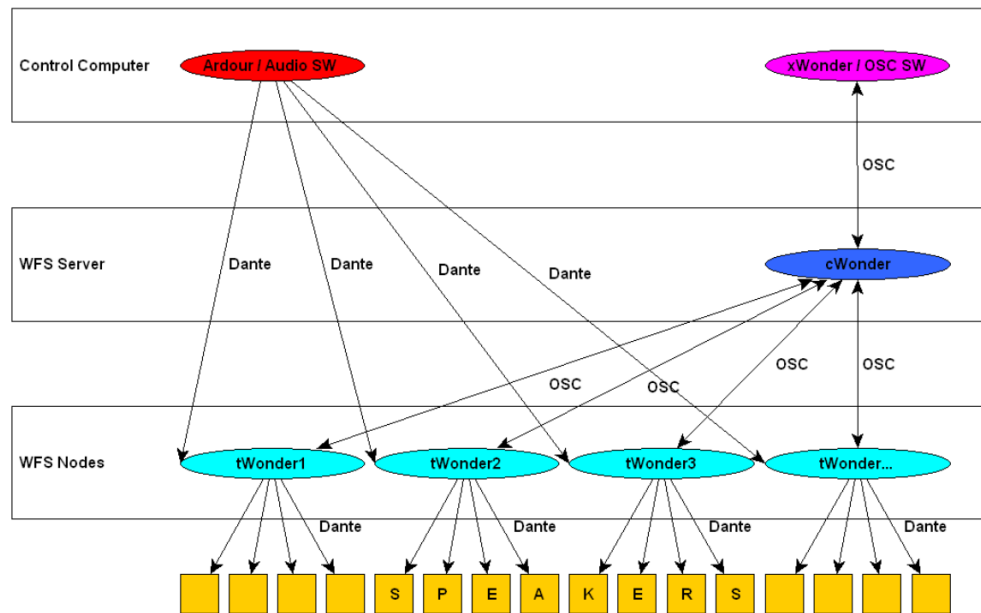


Abbildung 6.2: Netzwerkstruktur der WFS-Anlage [Mic12]

6.2 MoWeC

MoWeC steht für „Motion Tracking Wavefield Synthesis Connector“ und wurde ursprünglich von Malte Nogalski entwickelt, um eine gestengesteuerte Positionierung von Schallquellen einer Wellenfeldsynthese-Anlage mit Hilfe eines kamerabasierten 3D-Tracking-Systems zu ermöglichen. Durch die vorhandene Anbindung an das genutzte Tracking-System und das Empfangen von Daten, welche für die Messungen relevant sind, bot sich die Entwicklung der Messsoftware auf Basis von MoWeC an.

6.2.1 Funktion

Die Hauptaufgabe des MoWeC ist die Verbindung des Tracking-Systems und der Wellenfeldsynthese-Anlage. Dies geschieht durch die Analyse der OSC-Nachrichten, die von WONDER verschickt werden, und der Pakete, welche das Tracking-System per Multicast versendet. Quellen können mit Hilfe von MoWeC einem Tracking-Target zugeordnet werden. Durch diese Verbindung können Koordinaten in WONDER wie Quellenpositionen genutzt werden und werden in der xWonder Oberfläche als Quelle angezeigt. Weitere Funktionen zum Umrechnen der Einheiten und zur Vereinheitlichung der verschiedenen Koordinatensysteme sind ebenfalls vorhanden. Für weitere Informationen empfiehlt sich die Bachelor Thesis von Malte Nogalski [Nog12]

7 Entwicklung des Hörtests zur Lokalisationsgenauigkeit

Im Rahmen der Entwicklung des automatisierten Hörtests wurde der MoWeC erweitert, um die benötigten Werte zu bestimmen, zu berechnen und verwertbar abzuspeichern. Der Hörtest bestimmt bei jedem Durchlauf die Position der genutzten Targets und berechnet die Abweichungen zum Ziel. Diese Werte werden in externe Dateien geschrieben, um diese später zur Auswertung zu nutzen. Gleichzeitig werden die aktuellen Werte in der GUI ausgegeben, um Unregelmäßigkeiten sofort zu erkennen. Nach jeder Betätigung des Auslösers werden die aktuellen Daten erfasst und die Schallquelle wechselt zur nächsten zufälligen Position.

7.1 Zielsetzung

Die Genauigkeit, mit welcher Probanden eine Schallquelle lokalisieren, soll automatisiert ermittelt werden. Durch Nutzung des Tracking-Systems soll die genaue Position lokalisiert werden, an der die Testperson die Schallquelle vermutet. Es sollen nacheinander Schallquellen an verschiedenen Positionen auftauchen. Die Position der Schallquelle soll nach jeder Bestätigung der Probanden zur nächsten Position springen. Die Bestätigung soll über ein externes Eingabegerät abgewickelt werden, sodass ein komfortables Bestätigen seitens der Probanden möglich ist. Die Abweichung zur tatsächlichen Position soll berechnet und in ein Dokument abgespeichert werden, welches später der Auswertung dient.

7.2 Berechnung der Abweichung vom Ziel

Um verwertbare Daten aus den ermittelten Werten des Tracking-Systems zu erhalten, müssen die relevanten Koordinaten extrahiert und die Abstände zwischen diesen berechnet werden. Die Position der verschiedenen Targets des Tracking-Systems werden durch den MoWeC erfasst und ausgelesen. Mit der `targetSourceMatcher`-Funktion werden die physischen Targets jeweils einer Quelle in WONDER zugeordnet. Im MoWeC werden die OSC-Nachrichten zu diesen Targets empfangen und verarbeitet. Dort setzt die Entwicklung des Testprogramms an.

In der Hauptschleife, die bis zur Beendigung des Programms weiter läuft, werden mit jedem Durchlauf die aktuellen Positionen der Quellen zu Berechnungszwecken zwischengespeichert. Die genutzten Quellen stellen die Position der Hand, des Kopfes und des Tonursprungs dar. Es werden jeweils die X-, Y- und Z-Werte genutzt. Eine Ausnahme gibt es bei dem Tonursprung. Da MoWeC dreidimensionale Quellen nicht vollständig unterstützt, wurden die Positionsdaten zu diesem direkt aus dem WLAN des Tracking-Systems abgefangen. Für die Lage des Zeiger-Targets im Raum werden die Positionsdaten der Z-Achse aus der Rotationsmatrix extrahiert. Die Berechnungen laufen nach folgendem Schema ab:

1. Berechnung des optimalen Vektors, der von der Hand zur Schallquelle zeigt.

$$\text{OptimalerVektor} = \text{VektorSoundPosition} - \text{VektorHandPosition}$$

2. Bestimmen der Vektorlänge.

$$\text{Vektorlänge} = \sqrt{\text{OptimalerVektor}X^2 + \text{OptimalerVektor}Y^2 + \text{OptimalerVektor}Z^2}$$

3. Normieren des Vektors.

$$\text{NormierterOptimalerVektor} = \text{OptimalerVektor} / \text{Vektorlänge}$$

4. Zeigerichtung der Z-Achse des Zeiger-Targets aus der wie folgt aufgebauten Rotationsmatrix bestimmen:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$\text{ZeigerichtungX} = a_{13}$$

$$\text{ZeigerichtungY} = a_{23} * -1 // \text{Zur Anpassung an das Koordinatensystem}$$

$$\text{ZeigerichtungZ} = a_{33}$$

Das Koordinatensystem des Tracking-Systems nutzt im Vergleich zum Koordinatensystem der Wellenfeldsynthese-Anlage eine invertierte Y-Achse. Um diesen Effekt auszugleichen, wird der Y-Wert invertiert.

5. Berechnung der Abweichung sowie Umrechnung der Werte in Radiant und Konversion dieser in Grad.

$$\begin{aligned} \text{AbweichungX} &= (\text{Degree}) \arcsin(\text{ZeigerichtungX}) \\ &- (\text{Degree}) \arcsin(\text{NormierterOptimalerVektor}X) \end{aligned}$$

$$\begin{aligned} \text{AbweichungZ} &= (\text{Degree}) \arcsin(\text{ZeigerichtungZ}) \\ &- (\text{Degree}) \arcsin(\text{NormierterOptimalerVektor}Z) \end{aligned}$$

7 Entwicklung des Hörtests zur Lokalisationsgenauigkeit

Die Abweichung auf der Y-Achse benötigt eine gesonderte Berechnung, um die Z-Komponente herauszurechnen, da diese sonst verfälscht wird.

6. Bestimmung der Vektorlänge von NormierterOptimalerVektorX und -Y sowie ZeigerichtungX und -Y.

$$\text{Vektorlänge} = \sqrt{\text{NormierterOptimalerVektorX}^2 + \text{NormierterOptimalerVektorY}^2}$$

$$\text{NormierterOptimalerVektorX} = \text{NormierterOptimalerVektorX} / \text{Vektorlänge}$$

$$\text{NormierterOptimalerVektorZ} = \text{NormierterOptimalerVektorY} / \text{Vektorlänge}$$

$$\text{Vektorlänge} = \sqrt{\text{ZeigerichtungX}^2 + \text{ZeigerichtungY}^2} \hat{=} \sqrt{a_{13}^2 + a_{23}^2}$$

$$\text{NormierteZeigerichtungX} = \text{ZeigerichtungX} / \text{Vektorlänge}$$

$$\text{NormierteZeigerichtungY} = \text{ZeigerichtungY} / \text{Vektorlänge}$$

7. Berechnung des Skalarproduktes des eigens erstellten 2D-Vektors, um den Winkel zwischen den beiden Vektoren zu ermitteln sowie Umrechnung in Grad.

$$\text{SkalarProdukt} = (\text{NormierterOptimalerVektorX} * \text{NormierteZeigerichtungX}) \\ + (\text{NormierterOptimalerVektorY} * \text{NormierteZeigerichtungY})$$

$$\cos \varphi = \arccos(\text{SkalarProdukt})$$

Mit $\cos \varphi$ ist somit der gesuchte Wert für die Y-Achse berechnet. Allerdings ist dieser noch vorzeichenlos.

Berechnung des Vorzeichens:

Zur Berechnung des Vorzeichens wird eine gedachte Linie vom Hand-Target aus in dessen Zeigerichtung gezogen. Diese Linie ist so lang wie der Abstand zwischen dem Hand-Target und der Soundquelle. Die Position des Endpunktes der Linie wird mit der Position der Soundquelle verrechnet, dadurch wird ermittelt, ob das Vorzeichen positiv oder negativ ist. Bei der genutzten Methode kann es zu Ungenauigkeiten nahe des Übereinstimmungspunktes kommen, die jedoch in einem für diesen Versuch irrelevant kleinen Bereich liegen. Um ein genaueres Ergebnis zu erreichen, könnte eine Fläche in der Größe der sichtbaren Fläche aufgespannt werden. Mit Hilfe der Fläche können daraufhin die Schnittpunkte der Soundquellenposition und dem Zielpunkt berechnet werden.

8. Bestimmung der Entfernung zwischen Hand- und Soundquelle mit Hilfe der *distanceToXY()* Funktion aus MoWeC sowie Normierung des Vektors.

$$\text{HandSoundDistanz} = (\text{Meter})\text{Handposition.distanceToXY}(\text{Soundposition})$$

$$\text{Vektorlänge} = \sqrt{\text{ZeigerichtungX}^2 + \text{ZeigerichtungY}^2}$$

$$\text{NormierteZeigerichtungX} = \text{ZeigerichtungX} / \text{Vektorlänge}$$

9. Streckung des Vektors um die berechnete Distanz.

$$\text{GestreckteZeigerichtungX} = \text{NormierteZeigerichtungX} * \text{HandSoundDistanz}$$

10. Berechnung des Zielpunktes.

$$\text{ZielpunktX} = \text{HandPositionX} + \text{ZeigerichtungX}$$

11. Bestimmung und Zuweisung des Vorzeichens.

$$\text{Vorzeichenvariable} = \text{ZielpunktX} - \text{SoundPositionX}$$

$$\text{if}(\text{Vorzeichenvariable} < 0) \{ \cos \varphi = \cos \varphi * -1 \}$$

$$\text{AbweichungY} = \cos \varphi$$

Zu Zwecken der Referenzmessung wird dieselbe Berechnung benutzt. Das bisherige Ziel, die Soundquelle, wird in dem genannten Fall durch ein Tracking-Target ersetzt.

Die durchschnittliche Laufzeit der Berechnung beträgt bei 1000 Durchläufen 0,457 Millisekunden pro Durchlauf. Dementsprechend ist die Latenz für die Testzwecke vernachlässigbar klein, da während dieser Zeit keine Bewegungen von bedeutender Größe ausgeführt werden können. Laut Forschungen an der Harvard Universität [RVRL13] ist die schnellste Bewegung eines Menschen die Wurfbewegung, bei der die Schulter eine Rotationsbewegung mit bis zu 9000° pro Sekunde durchführt. Auf die durchschnittliche Laufzeit heruntergerechnet, kann eine maximale Bewegung von $(9000^\circ / 1000) * 0,457ms = 4,113^\circ$ zwischen zwei Durchläufen erreicht werden. Dieser Wert ist vertretbar, da die Geschwindigkeit des Zeigens den Test nicht beeinflusst. Des Weiteren handelt es sich hierbei um nicht vorgesehene Geschwindigkeiten, bei denen ein präzises Zeigen nicht möglich ist.

7.3 GUI

Die GUI wurde während der Entwicklung genutzt, um in Echtzeit die Daten analysieren und auf Fehler überprüfen zu können. Später wurde diese umgestaltet, sodass verschiedene Positionswerte und die Abweichung ausgegeben wurden. Dadurch kann der Testlauf aktiv überwacht werden. Durch Einfärbung des Hintergrundes der Abweichungswerte kann einfach erkannt werden, in welchem Rahmen sich die Abweichung um die verschiedenen Achsen ungefähr bewegt.

Tabelle 7.1: Einfärbung der Abweichungswerte

Grün	-6° bis 6°
Gelb	-12° bis -6° und 6° bis 12°
Rot	< -12° und >12°

Die Werte aus Tabelle 7.1 basieren auf Annahmen, die aus Testläufen stammen. In Abbildung 7.1 sind verschiedene Zustände der GUI zu sehen, in welchen verschieden präzise gezielt wurde. In der ersten Zeile ist die Position des Kopfes angegeben, die zweite Zeile enthält die aktuelle Position der Soundquelle und die dritte Zeile die aktuelle Abweichung von der gezeigten Richtung zur Soundquelle. Angegeben sind in den ersten beiden Zeilen jeweils die X-, Y- und Z-Koordinaten im Koordinatensystem der Wellenfeldsynthese-Anlage und in der letzten Zeile die Abweichung für die jeweiligen Achsen in Grad.

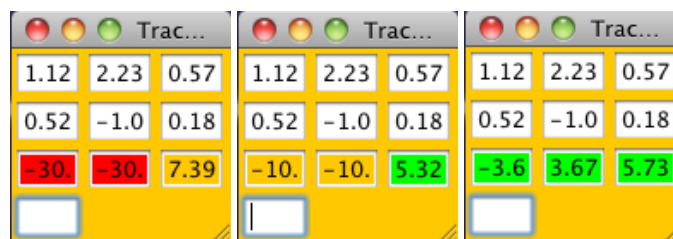


Abbildung 7.1: Verschiedene Zustände der GUI

7.4 Erfassen der Eingaben

Das untere linke Eingabefeld in der GUI (siehe Abb. 7.1) dient als Eingabefeld zur Aktivierung des Logs. Auf diesem Feld erwartet ein *KeyListener* das Aktivieren der *Punkt*-, *SEITE AUF*-, *SEITE AB*- oder *F5*-Taste. Diese Tasten entsprechen dem vorgegebenen Layout des genutzten Presenters. Somit kann jede beliebige Taste des Presenters (siehe Abschnitt 5.3) genutzt werden, um den Log zu aktivieren. Andere Tasteneingaben werden ignoriert. Damit der Tastendruck erkannt werden kann, muss das Textfeld im Vordergrund und aktiviert sein (Cursor blinkt). Sobald eine der vier möglichen Tasten gedrückt wird, werden die zuletzt berechneten Werte, weitere Positions- und Richtungsdaten sowie persönliche Daten der Testperson in eine txt-Datei mit dem Namen der Testperson geschrieben. Diese wiederum liegt in einem Ordner mit selbigem Namen. Im Laufe der Referenzmessung werden die jeweiligen Werte in eine zusätzlichen Datei geschrieben, die ebenfalls in dem Ordner liegt, dem Namen folgt bei dieser

Datei ein „Ref“. Beim jeweils ersten Aktivieren der Logs wird einmalig eine Überschrift in die Datei eingefügt, die später der Identifikation der Werte dient. Ein genauer Aufbau der Log-Datei wird in Abschnitt 7.6 beschrieben.

7.5 Ändern der Quellenposition

Die 63 Quellenpositionen werden bei jedem Programmstart mithilfe des Fisher-Yates Zufallsalgorithmus [Bla15] erneut in eine zufällige Reihenfolge gebracht. Somit können Probanden sich nicht an einem festen Ablauf orientieren. Nach jeder erfassten Eingabe über den Presenter wird eine OSC-Nachricht gesendet, welche WONDER veranlasst die Schallquelle an die entsprechende Position zu setzen. Die OSC-Nachrichten werden folgendermaßen erzeugt:

```
OSCMessage("/WONDER/source/position3D", new Object[]{id, x, y, z})
```

Die Variablen der Nachricht werden mit den jeweiligen Informationen zur Schallquelle beschrieben. Anschließend wird die Nachricht in ein OSCMessage-Objekt gespeichert, welches daraufhin durch MoWeC versandt wird.

7.6 Messdaten

Die Messdaten werden in einer txt-Datei gespeichert. Diese enthält die Werte aus der Messung und einen Header mit zugehöriger Beschriftung im CSV-Format. Der Header ist wie folgt aufgebaut:

```
"DX: ", "DY: ", "DZ: ", "GX: ", "GY: ", "GZ: ", "TX: ", "TY: ", "TZ: ", "SPX: ", "SPY: ",  
"SPZ: ", "HPX: ", "HPY: ", "HPZ: ", "KPX: ", "KPY: ", "KPZ: ", "ID: ", "Order: ",  
"Age: ", "Sex: ", "Audio: ", "WFS: ", "Time: "
```

Diese Werte werden in der Tabelle 7.2 genauer erläutert. Einige Daten, wie etwa $G_{X,Y,Z}$ und $T_{X,Y,Z}$, wurden nur zum Zweck der Nachvollziehbarkeit aufgenommen, um im Falle von besonders auffälligen Werten eine Fehlfunktion ausschließen zu können. Die restlichen Daten werden zu Auswertungszwecken oder für die Zuordnung der Werte zu bestimmten Positionen benötigt.

Tabelle 7.2: Erklärung der Abkürzungen der Log-Dateien

$D_{X,Y,Z}$	Differenz/Abweichung von der Soundquelle in Grad
$G_{X,Y,Z}$	Ist-Vektor der gezeigten Richtung
$T_{X,Y,Z}$	Soll-Vektor der gezeigten Richtung
$SP_{X,Y,Z}$	Koordinaten der Soundposition
$HP_{X,Y,Z}$	Koordinaten der Handposition
$KP_{X,Y,Z}$	Koordinaten der Kopfposition
ID	Identifikationsnummer der Soundposition
Order	Position in der zufälligen Abspielreihenfolge
Age	Alter der Testperson
Sex	Geschlecht der Testperson
Audio	Audioerfahrung der Testperson
WFS	WFS-Anlage bekannt oder nicht
Time	Zeit bis zur Auslösung des Log

8 Versuch zur Lokalisierungsgenauigkeit im dreidimensionalen Raum

Die Wellenfeldsynthese-Anlage ist bisher auf eine bestimmte Höhe beschränkt. Dieser Versuch soll zeigen, ob diese Beschränkung mit geringem Aufwand und zufriedenstellendem Ergebnis überwunden werden kann. Mit Hilfe von 16 Lautsprechern soll herausgefunden werden, inwiefern sich eine Erweiterung der Wellenfeldsynthese-Anlage um einzelne Lautsprecher in der Höhe auf die Präzision der Lokalisation von Schallquellen einer Testperson auswirkt. Die technische Grundlage für diese Arbeit lieferte die Abschlussarbeit von Marvin Geitner [Gei15]. Im Zuge dessen wurde die Software WONDER so modifiziert, dass Quellen nun auch in der Höhe positioniert werden können.

8.1 Erwartete Ergebnisse

Da der Versuchsaufbau in der Theorie zu große Lücken zwischen den Lautsprechern aufweist, ist zu erwarten, dass die Genauigkeit durch das stärkere räumliche Aliasing auf der Z-Achse deutlich geringer ausfallen wird als auf der X-Achse. In der Horizontale sollte die Genauigkeit kaum Einbußen gegenüber den WFS-Modulen haben, da die Lücken dort wesentlich kleiner sind.

8.1.1 Beeinflussende Faktoren

Faktoren, welche sich negativ auf den Test auswirken können, sind:

- Der Raum ist nur wenig gegen Schallreflexionen gedämmt
- Ungenauigkeit der Lautsprechermodule durch bauliche Trennung der Hoch- und Tief-tonchassis sowie drei Reihen an Lautsprechern in den Four Audio-Modulen
- Softwareseitige, zylinder- statt kugelförmige Ausbreitung des Signals, sodass die Laufzeiten auf verschiedenen Höhen beeinflusst werden

- Auftretende Alias-Effekte
- Die Sichtbarkeit der Lautsprecher
- Die Z-Achse ist schmaler als die X-Achse, sodass Fehler in Relation zur X-Achse verstärkt werden

8.2 Erweiterung der Wellenfeldsynthese-Anlage

Der grundlegende Testaufbau erfolgte durch Hinzufügen zweier Lautsprecher-Arrays (siehe Abb. 8.1 u. 8.2) mit jeweils acht Lautsprechern in den vorhandenen Aufbau der Wellenfeldsynthese-Anlage der HAW Hamburg (siehe Abb. 5.1). Bedingt durch den Aufbau und um horizontal eine möglichst große Lautsprecherdichte zu erhalten, wurden die Lautsprecher auf einer Breite von 3,15 Meter aufgeteilt. Das entspricht einem Abstand von 45 Zentimeter zwischen den Membranmitten. Das obere Array steht auf einer Höhe von 1,29 Meter, das untere ist mit 18 Zentimeter angegeben, was jeweils der Membranmitte der Breitbandchassis entspricht. Die Mitte der Arrays ist an der Mitte der darüberliegenden Wellenfeldsynthese-Module ausgerichtet.

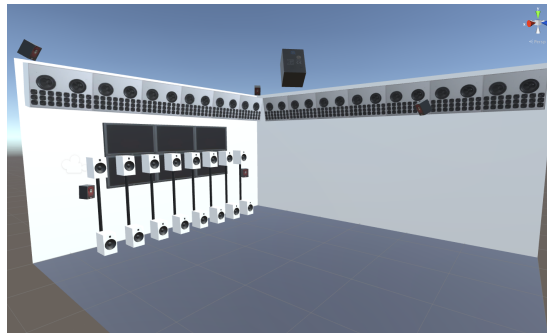


Abbildung 8.1: Der Versuchsaufbau von der Seite gesehen (in Unity modelliert)

Jeweils acht Lautsprecher werden per XLR-Kabel an ein RedNet-Interface angeschlossen. Durch ein herkömmliches Netzkabel wird eine Verbindung zu dem Dante-Netzwerk hergestellt. Damit die neuen Lautsprecher angesteuert werden können, müssen diese über den Dante Controller, einer Software zur virtuellen Verbindung von Ton-Ein- und Ausgängen, mit dem System verknüpft werden.

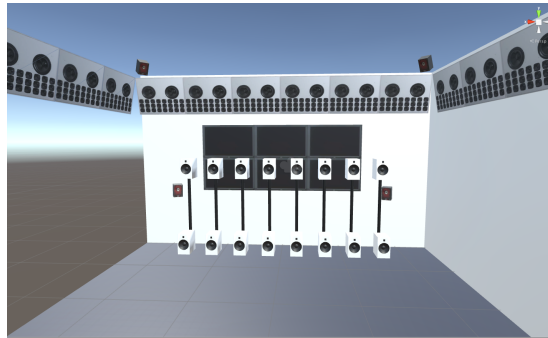


Abbildung 8.2: Der Versuchsaufbau frontal gesehen (in Unity modelliert)

8.2.1 Anpassung der Lautstärke und Klangfarbe

Das Anpassen der Lautstärke und der Klangfarbe ist ein wichtiger Schritt, um korrekte Ergebnisse in den Messungen zu gewährleisten. Sind die Lautstärken ungleichmäßig verteilt, wird das Geräusch eher in der Nähe des Lautsprechers mit dem höheren Geräuschpegel geortet. Durch einen Unterschied in der Klangfarbe ist die Unterscheidung von verschiedenen Lautsprechertypen möglich. Dadurch kann eine grobe Position der Schallquelle unerwünschterweise geschätzt werden. Da konstruktionsbedingt keine exakte Anpassung der Klangfarbe möglich ist, wurde diese soweit wie möglich durchgeführt.

Zum Anpassen der Lautstärke und der Klangfarbe wurde auf den Arrays und auf den Wellenfeldsynthese-Lautsprechern jeweils ein rosa Rauschen wiedergegeben, um dieses mit dem AKG C3000B Mikrofon in Verbindung mit dem Scarlett 18i8 Interface aufzunehmen. Dazu wurde das Mikrofon in 2 Metern Entfernung zu den Lautsprechern in der Mitte des Raumes und auf 1,30 Meter Höhe aufgestellt (siehe Abb. 8.4). Der Abstand von 2 Meter wurde gewählt, weil die Probanden zur Testdurchführung in derselben Entfernung sitzen. Die Höhe soll in etwa der Kopfhöhe im Sitzen entsprechen. Das aufgenommene Tonmaterial wurde später mit Hilfe der Open-Source-Audiobearbeitungssoftware Audacity in Form von Spektrogrammen untersucht. Die Spektrogramme geben Aufschluss über das Frequenzspektrum einer Audiodatei und lassen so eine Analyse der Differenzen zu. Nachdem die Lautstärke an den Adam A7x-Lautsprechern angepasst wurde, folgte der Hardware-Equalizer. Dieser sollte so eingestellt werden, dass er soweit wie möglich dem Frequenzspektrum der Four Audio-Module gleicht. Zum Vergleich wurden die Spektrogramme in Abbildung 8.3 übereinander gelegt. So werden die unterschiedlichen Frequenzspektren der Lautsprecher ersichtlich und geben Aufschluss über die nötige Einstellung des Equalizers.

8 Versuch zur Lokalisierungs-genauigkeit im dreidimensionalen Raum

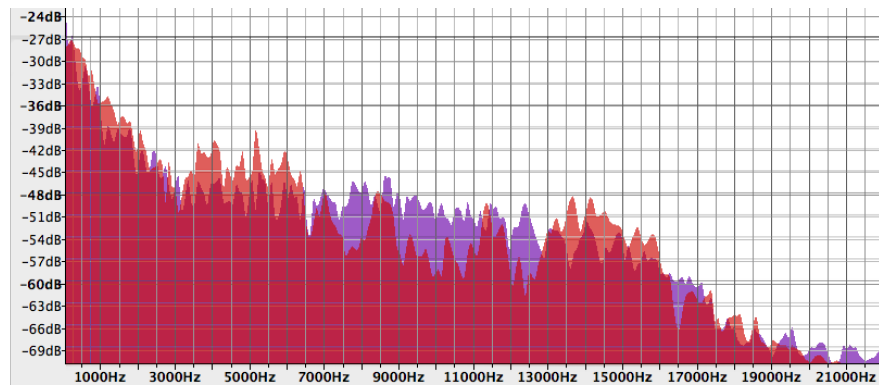


Abbildung 8.3: Spektrogramme der A7x-(lila) und Four Audio-Module (rot) übereinandergelegt. Die verwendete Equalizer Einstellung der A7x-Lautsprecher ist aus Abb. 8.5 zu entnehmen.



Abbildung 8.4: Der Versuchsaufbau zum Ermitteln des Frequenzspektrums



Abbildung 8.5: Eingänge und Equalizer der Adam A7x Lautsprecher

8.2.2 Auswahl des Testtones und der Positionierung

Der wiedergegebene Ton muss einige Ansprüche erfüllen. Wichtige Kriterien dabei sind etwa das Frequenzspektrum und dass dieser Ton in der Realität auch aus verschiedenen Höhen stammen kann. Das Frequenzspektrum sollte optimalerweise sehr breit gefächert sein, sodass möglichst alle hörbaren Bereiche abgedeckt sind. Ein Rauschen ist für diesen Fall optimal, kann bei einem mehrminütigen Testlauf jedoch als unangenehm empfunden werden. Da sich der Mensch bei der Lokalisation von bestimmten Geräuschen oft auf bekannte Situationen beruft, kann ein Geräusch, das Assoziationen mit einem bestimmten Höhenbereich hervorruft, dazu führen, dass dieses eher in diesem assoziierten Bereich vermutet wird [PKE14]. Hört man beispielsweise ein Flugzeug, würde man sich vorzugsweise am Himmel nach diesem umschauchen. Ein Auto hingegen wird eher auf dem Boden erwartet. Getrieben durch diese Kriterien wurde eine Tonspur mit einem Wasserfall/Wasserplätschern ausgewählt. Diese Tonspur erfüllt beide Bedingungen: Es handelt sich sowohl um ein natürliches Rauschen, welches angenehmer ist als ein künstlich erzeugtes Rauschen, als auch um ein Geräusch, das aus verschiedenen Höhen stammen kann.

Eine weitere Frage warf der gewählte Abstand der Schallquelle zu den Lautsprechern auf. So sollte der Abstand groß genug sein, um mehrere Lautsprecher mit einzubeziehen, damit Wellenfeldsynthese betrieben werden kann. Er darf jedoch auch nicht zu groß sein, da sonst die Lokalisierbarkeit darunter leiden würde. Durch Berechnungen, bei denen davon ausgegangen wird, dass alle Lautsprecher innerhalb eines 60° Abstrahlwinkels ausgehend von der Schallquelle direkt involviert sind, wurde -1 Meter als Y-Wert für ausreichend befunden. Bei diesem Wert werden Lautsprecher auf einer Breite von etwa 1,08 Meter angesprochen. Bei einem Lautsprecherabstand von 45 Zentimeter, gemessen von der Membranmitte, entspricht

das minimal zwei und maximal drei aktiven Lautsprechern, welche direkt involviert sind. In der Höhe sind somit ein bis zwei der drei Arrays zeitgleich aktiv. Laut Wierstorf et al. [Bla13] zeigte ein ähnlicher jedoch zweidimensionaler Versuchsaufbau, dass sich die Lokalisationsgenauigkeit bei einem Lautsprecherabstand von 20 Zentimeter gegenüber einem Abstand von 40 Zentimeter nur um etwa 1° unterscheidet.

Als Quellenart kommt aufgrund der besten Ortbarkeit nur die Punktquelle in Frage.

8.2.3 Versuchsablauf

Nachdem der Proband gebeten wurde sich auf den Stuhl zu setzen, wird ihm der Ablauf kurz erklärt. Nach der Erklärung wird das Messprogramm gestartet, wodurch nach Eingabe der persönlichen Daten des Probanden die Referenzmessung gestartet wird. Als Ziel dient bei der Referenzmessung das HT23 Target, welches an fünf ausgewählten Positionen an der Front der Anlage gehalten wird. Die Positionen werden in folgender Reihenfolge abgearbeitet: oben links, unten links, mittig, oben rechts und unten rechts. Der Proband wird gebeten, jeweils mit dem TT1 Target auf das HT23 Target zu zeigen, welches der Versuchsleiter in der Hand hat, und zum Bestätigen der Eingabe eine beliebige Taste auf dem Logitech R400 Presenter zu betätigen. Nach jeder Bestätigung ertönt ein Ton, um dem Versuchsleiter zu signalisieren die nächste Position einzunehmen. Nachdem alle Positionen durchlaufen sind, wird dem Probanden das weitere Vorgehen erklärt. Er wird gebeten mit dem TT1 Target dort hinzuzeigen, wo er den Ton vermutet, und wieder mit dem Presenter zu bestätigen. Daraufhin wird der Wasserfall-Ton in Ardour auf Dauerschleife aktiviert und der Proband gebeten, nun zu beginnen. Nach jeder Bestätigung durch den Presenter springt die Schallquelle an die nächste Position. Die Schallquelle durchläuft insgesamt 63 Positionen in zufälliger Reihenfolge. Die aktuelle Nummer der Position wird in der Konsole ausgegeben. Der Test ist beendet, wenn die Quelle nicht mehr weiterläuft, bzw. alle 63 Positionen durchlaufen sind.

8.3 Ergebnisse

Die Auswertung soll aufzeigen, wie genau die Lokalisation einer Schallquelle im gegebenen Versuchsaufbau möglich ist. Des Weiteren sollen Zusammenhänge verschiedener Faktoren untersucht werden. Für die Auswertung standen 1260 Messwerte zur Verfügung.

8.3.1 Ergebnisse des Versuches

Der Versuch wurde mit 20 Personen durchgeführt, 2 davon weiblich und 18 männlich. Das Durchschnittsalter liegt bei 26 Jahren, teilgenommen haben Personen im Alter zwischen 19

und 50 Jahren. 95% der Probanden sind zwischen 19 und 32 Jahre alt sind. Aus Abbildung 8.6 ist kein Zusammenhang der Präzision mit dem Alter feststellbar.

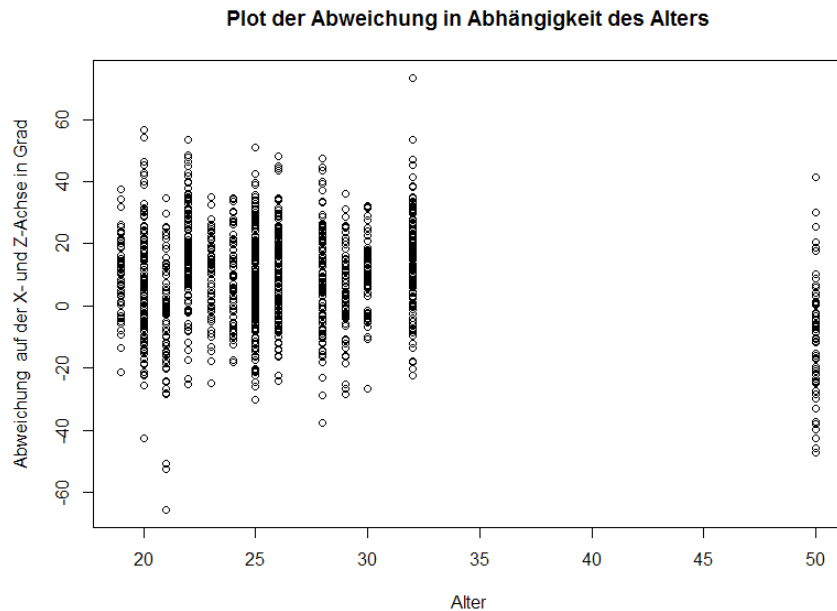


Abbildung 8.6: Die Abweichung auf der X- und Z-Achse im Zusammenhang mit dem Alter der Probanden dargestellt

50% der Probanden war die Wellenfeldsynthese-Anlage schon vor dem Versuch bekannt, bei der Auswertung gab es jedoch keine auffälligen Unterschiede in der Abweichung gegenüber den Probanden, welchen die Anlage nicht bekannt war. Als audiophil bezeichneten sich 12 der 20 Probanden. Bei der Auswertung lassen sich jedoch keine großen Unterschiede feststellen. Probanden welche mit „ja“ geantwortet haben, schneiden insgesamt minimal schlechter ab (siehe Abb. 8.7). Der Unterschied befindet sich aber im Rahmen einer zu erwartenden Varianz der Ergebnisse.

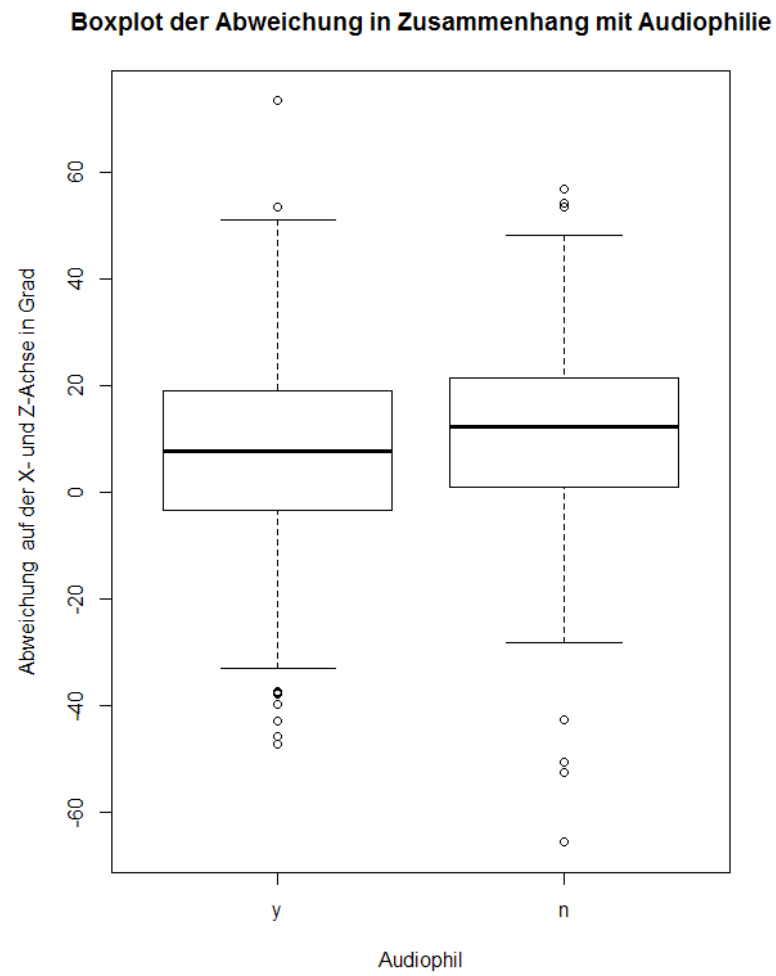


Abbildung 8.7: Die Abweichung auf der X- und Z-Achse in Relation zur Audiophilie als Boxplot dargestellt

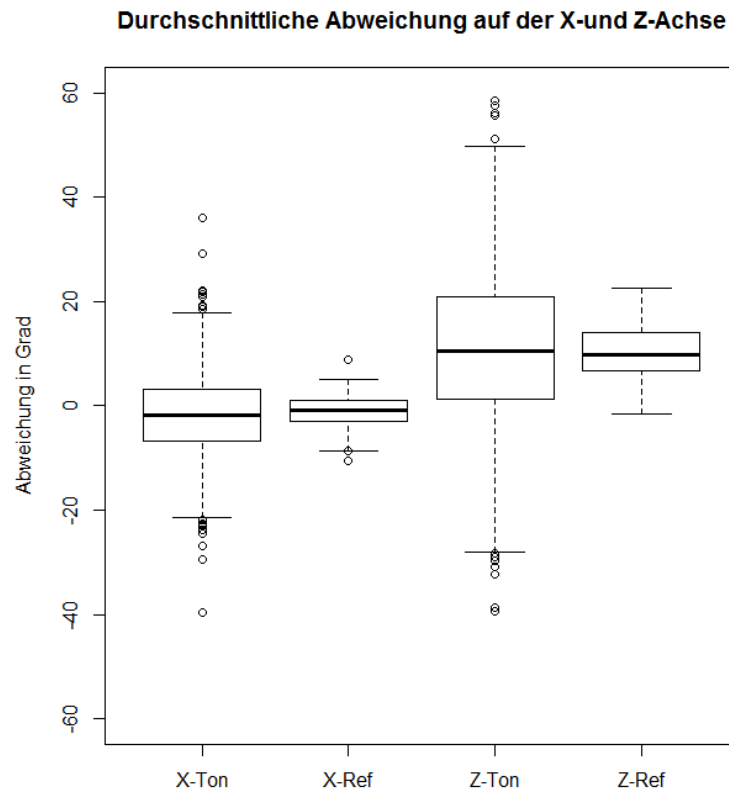


Abbildung 8.8: Die Abweichungen auf den verschiedenen Achsen als Boxplot dargestellt, jeweils für die Referenzmessung mit sichtbarem Ziel und für den eigentlichen Versuch mit Schallquelle als Ziel

Die durchschnittliche Abweichung auf den verschiedenen Achsen zeigt deutlich, dass auf der Z-Achse, wie zu erwarten war, die größten Abweichungen auftreten. Auffällig ist, dass der Großteil der Werte im Boxplot der Z-Achse (siehe Abb. 8.8) im positiven Bereich liegt. Die Quellen wurden folglich eher zu weit oben wahrgenommen. Bei der X-Achse sind hingegen keine auffälligen Ergebnisse sichtbar. Der Median liegt sehr nahe am Idealwert und die Quartile sind etwa gleichmäßig verteilt. Auf der X-Achse reichen die Quartile von 3° bis -6° , was einem guten Wert entspricht. Die Quartile der Z-Achse reichen dagegen von 21° bis 1° , die Streubreite der präzisesten Hälfte der Ergebnisse liegt gegenüber dem Wert der X-Achse etwa doppelt so hoch. Die absolute durchschnittliche Abweichung, welche der Tabelle 8.1 entnommen werden kann, zeigt bei dem Wert der Z-Achse eine etwas mehr als doppelt so große durchschnittliche Abweichung als auf der X-Achse. Weiter auffallend ist, dass die Ausreißer sehr starke Abweichungen aufweisen. Die durchschnittliche Abweichung aus der Referenzmessung weist zwar, wie in Abbildung 8.8 und 8.9 zu sehen, eine geringere Streuung der Werte auf, zeigt

jedoch dieselbe Charakteristik wie bei den Messungen mit Schallquellen als Ziel. Die genauen Werte zu den Abweichungen der Referenzmessung sind aus Tabelle 8.2 zu entnehmen.

Tabelle 8.1: Absolute durchschnittliche Abweichung mit akustischem Ziel

Achse	Absolut	Relativ	Median
X	6,12°	-1,78°	-1,81
Z	14,74°	10,69°	10,58

Tabelle 8.2: Durchschnittliche Abweichung mit visuellem Ziel

Achse	Absolut	Relativ	Median
X	2,51°	-1,11°	-0,79°
Z	10,28°	10,18°	9,81°

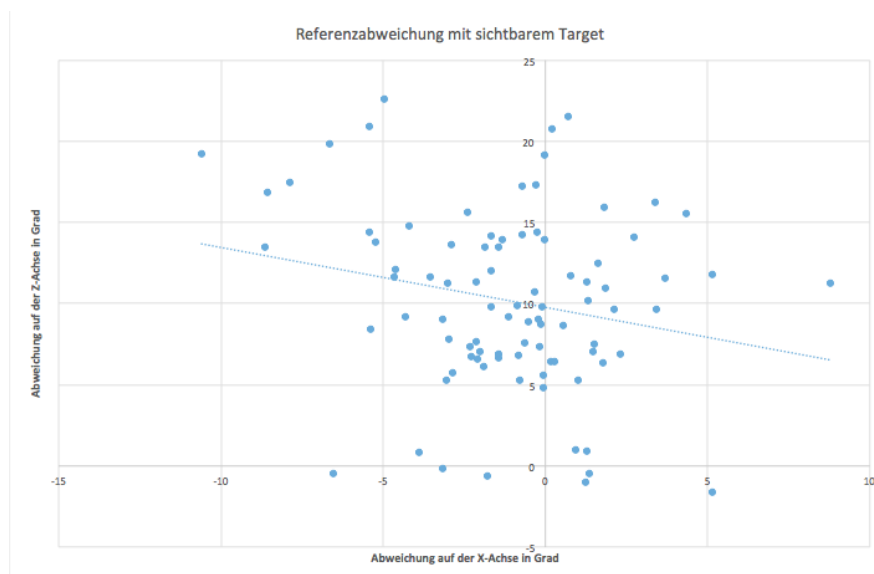


Abbildung 8.9: Die Abweichungen der Referenzmessung auf der X- und Z-Achse

Bei dem Vergleich der durchschnittlichen absoluten Abweichung je ID fällt auf, dass auf der X-Achse neun der dreizehn am schlechtesten lokalisierten (mehr als 7° Abweichung (siehe Abb. 8.10)) Quellenpositionen knapp unter oder über dem mittleren Lautsprecher-Array liegen. Des Weiteren fällt auf, dass die dreizehn Positionen eher an den Rändern des Messbereiches verteilt sind. Als Rand des Messbereiches werden hier die jeweils zwei äußeren Reihen und Spalten der ID-Matrix bezeichnet (siehe Abb. 8.12). Zwölf der vierzehn Positionen liegen bei

der X-Achse im Randbereich, sodass lediglich zwei Positionen bleiben, welche sich im Zentrum des Messbereiches befinden.

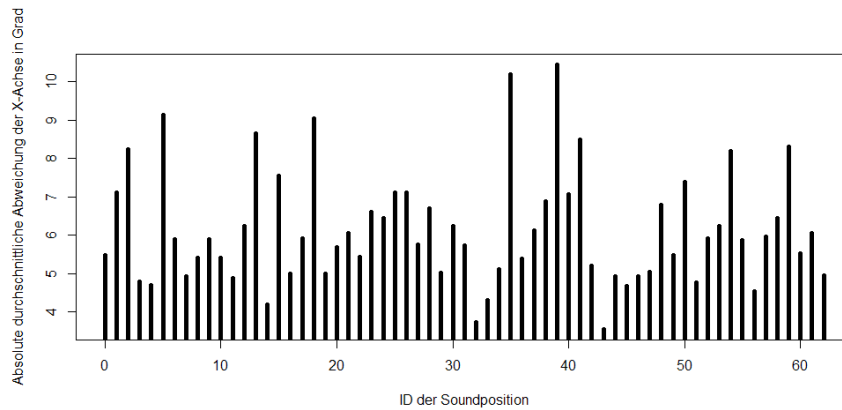


Abbildung 8.10: Durchschnittliche absolute Abweichung der einzelnen Quellenpositionen der X-Achse

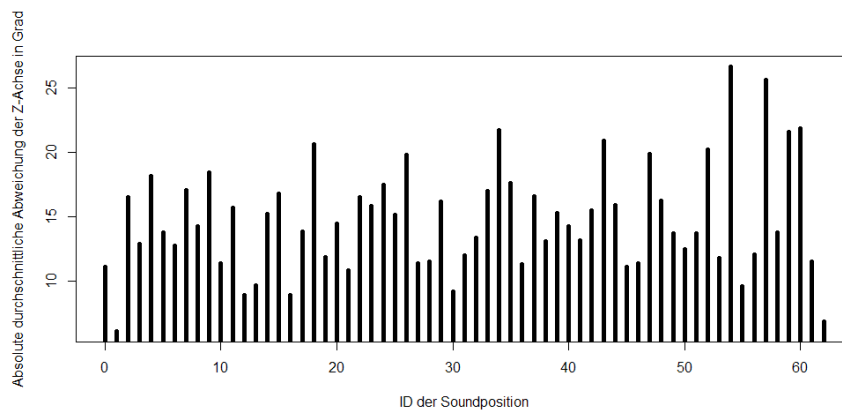


Abbildung 8.11: Durchschnittliche absolute Abweichung der einzelnen Quellenpositionen der Z-Achse

Auf der Z-Achse verteilen sich die zehn am schlechtesten lokalisierten Quellenpositionen (mehr als 20° Abweichung(siehe Abb. 8.11)) fast gleichmäßig auf den rechten und oberen Rand. Durch die hohe durchschnittliche Abweichung musste der Grenzwert für die am schlechtesten lokalisierten Quellenpositionen der Z-Achse deutlich nach oben korrigiert werden. Die am schlechtesten lokalisierten Quellenpositionen der X-Achse zeigen keine auffallenden Gemein-

samkeiten. Lediglich zwei der Werte, Position 54 und 59, treten bei beiden Achsen an derselben Stelle auf.

8	17	26	35	44	53	62	R1
7	16	25	34	43	52	61	R2
6	15	24	33	42	51	60	R3
5	14	23	32	41	50	59	R4
4	13	22	31	40	49	58	R5
3	12	21	30	39	48	57	R6
2	11	20	29	38	47	56	R7
1	10	19	28	37	46	55	R8
0	9	18	27	36	45	54	R9

Abbildung 8.12: Darstellung der IDs als Matrix an der ungefähren Position im Versuchsaufbau mit Reihenbeschriftung an der rechten Seite

Die Auswertung der einzelnen Reihen von Quellenpositionen aus Abbildung 8.13 zeigt mit absteigender Höhe der Quellenposition einen deutlichen Anstieg der Abweichung auf der Z-Achse. Die Abweichung auf der X-Achse bleibt konstant über alle Reihen.

Die genutzte Zeit wirkt sich in den Ergebnissen nicht positiv auf die Lokalisierungs-genauigkeit aus, da die Verteilung der präzisesten Werte, wie in Abbildung 8.14 sichtbar, bei beiden Achsen unter zehn Sekunden erzielt wurden. Betrachtet man die mittleren 50% der Werte für die jeweiligen Achsen, wurden diese auf der X-Achse in einem Zeitraum zwischen ca. 3 und 7 Sekunden und auf der Z-Achse zwischen ca. 2 und 5 Sekunden erreicht.

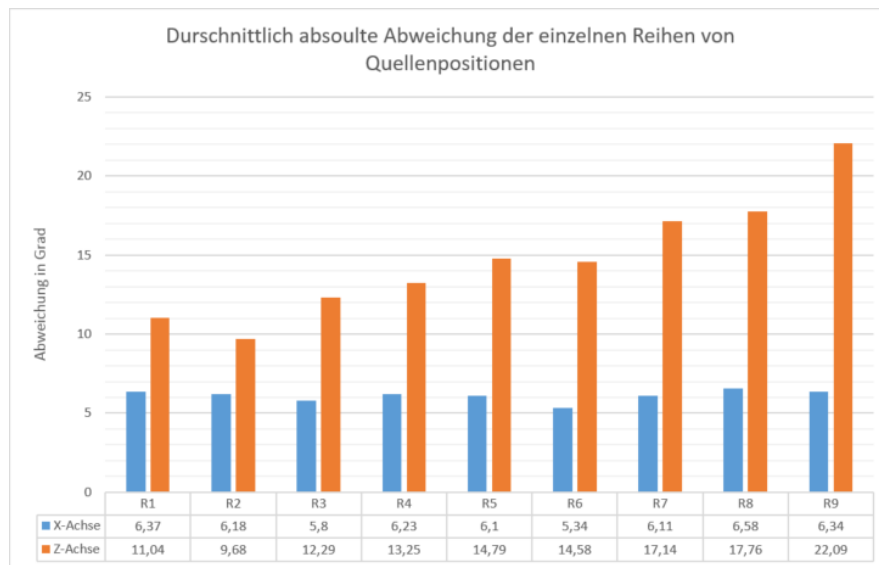


Abbildung 8.13: Auswertung der absoluten durchschnittlichen Abweichung pro Reihe von Quellenpositionen (siehe Abb. 8.12)

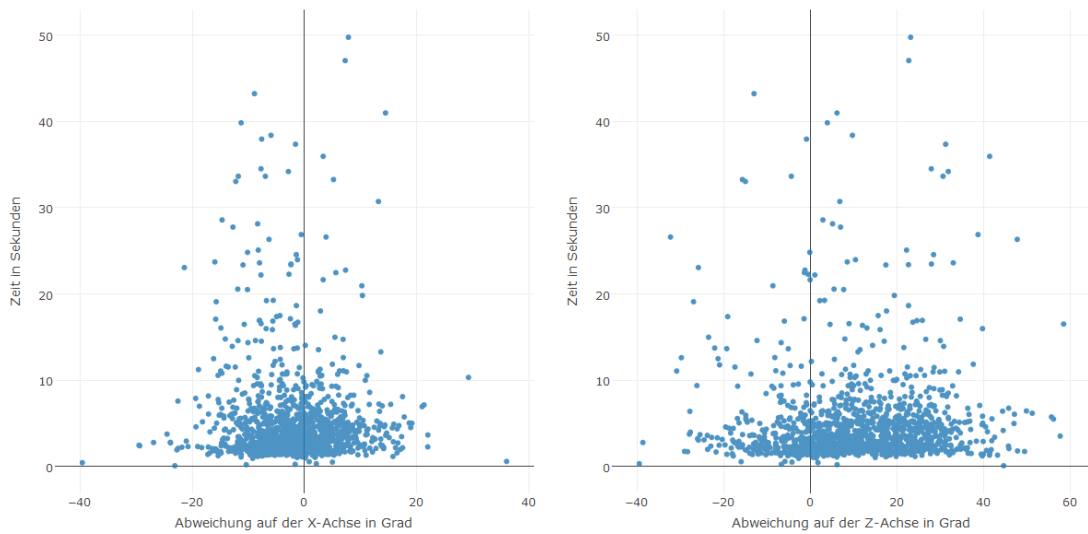


Abbildung 8.14: Abweichungen gegenüber der Zeit bis zum Auslösen der X- und Z-Achse

8.3.2 Diskussion der Ergebnisse

Matthew N. Montag führte einen ähnlichen Versuch mit zwei Wellenfeldsynthese-Lautsprecher-Arrays und vertikalem VBAP durch [Mon11]. Bis auf den Aufwärtsdrift der Z-Achse decken sich die Ergebnisse bezüglich der Abweichung mit den hier erzielten Ergebnissen. Auch bei Montag stellte sich heraus, dass die Abweichung in der Vertikale ca. doppelt so groß ist wie die in der Horizontale. Da auch Marvin Geitner [Gei15] schon festgestellt hat, dass ein zusätzliches Verwenden von Panning in der Vertikale kaum Vorteile bringt, bestärkt dies noch einmal das Absehen von einer Implementierung einer solchen gemischten Wiedergabemethode. Die Unterschiede in der Wahrnehmung decken sich laut Montag mit dem Grundsatz, dass der Mensch Ton in der Horizontale besser lokalisieren könne als in der Vertikale [Mon11]. Da in dem Versuch zwischen den Lautsprecher-Arrays ein Abstand von über einem Meter gegeben ist, sollten weitere Versuche mit einer Wand aus Wellenfeldsynthese-Arrays durchgeführt werden, um diese Aussage zu bestätigen. Sowohl bei Matthew Montag als auch in dieser Arbeit wurde der Versuch mit einem geringen Abstand des Hörers zu den Lautsprechern von unter 3 Meter sowie einer mittigen Sitzposition durchgeführt. Bei Versuchsaufbauten mit nicht mittigen Sitzpositionen oder sehr großen Abständen des Hörers zu den Lautsprechern können die Ergebnisse abweichen, weil das räumliche Aliasing in diesen Bereichen stärker ausfallen kann.

Bei Betrachtung der Ergebnisse der Referenzmessung relativiert sich der Aufwärtsdrift, da auch bei einem sichtbaren Ziel durchschnittlich um $10,28^\circ$ (siehe Tab. 8.2) über das Ziel hinweg gezeigt wird. Zieht man die absolute durchschnittliche Abweichung der Referenzmessung von den Werten der Messung mit Schallquellen (Tab. 8.1) ab, so ergeben sich Werte von $3,61^\circ$ absoluter durchschnittlicher Abweichung auf der X-Achse und $4,46^\circ$ absoluter durchschnittlicher Abweichung auf der Z-Achse. Die erzielten Ergebnisse zeigen einen guten Wert in der Lokalisierung auf und ähneln den Ergebnissen aus Montags Arbeit, in der Werte von $3,3^\circ$ absoluter durchschnittlicher Abweichung auf der X- und $6,5^\circ$ absoluter durchschnittlicher Abweichung auf der Z-Achse ermittelt worden sind [Mon11]. Bei der vertikalen Abweichung kann „kognitive Verzerrung“ eine Rolle spielen. Diese bezeichnet systematische, fehlerhafte Neigungen beim Wahrnehmen, Erinnern, Denken und Urteilen [Lüb15]. Eine kognitive Verzerrung kann bei der Zielerfassung dieses Versuches aufgrund von Wahrnehmungsfehlern bei der Ausrichtung des Zeigers aufgetreten sein. Eine korrekte Lokalisierung auf der X-Achse ist mit einer Abweichung von etwa 5 Grad scheinbar möglich. Auf der Z-Achse beträgt die Abweichung mit etwa 10 Grad ungefähr doppelt so viel. Die Annahme, dass die Lokalisation auf der X-Achse besser funktioniert als auf der Z-Achse, hat sich demnach bestätigt.

Bei der genutzten Zeit ist es, anders als erwartet, so, als hätten die schwer lokalisierbaren Positionen mehr Zeit in Anspruch genommen, ohne dabei zu einem präzisen Ergebnis zu führen. Dies lässt sich möglicherweise auf den Versuch der Probanden zurückführen, die Position der Schallquelle, trotz Unsicherheit, genauer zu bestimmen. Durch die suboptimale Dämmung des Labors können Schallreflexionen die Lokalisation zusätzlich beeinträchtigen. Somit lässt sich die Vermutung anstellen, dass Quellenpositionen, die schnell bestimmt wurden, auch einfacher zu lokalisieren waren, wohingegen ein längerer Zeitraum für eine schwierige Lokalisierbarkeit der Tonherkunft spricht.

Die Abweichung an bestimmten Tonpositionen zeigt, dass vor allem einzelne Positionen zwischen den Arrays schlechter lokalisiert wurden. Dies kann auf eine nicht exakte Lautstärkeanpassung der Lautsprecher oder auf die unterschiedlichen Klangfarben der Lautsprechermodelle zurückzuführen sein. Dadurch würde es zu einer Verzerrung der Position zugunsten des laueren Lautsprechers kommen. Eine weitere Vermutung bezieht sich auf die variierende Kopfhöhe. Die Werte variieren bei einer durchschnittlichen Kopfhöhe von 1,09 Meter in einem Bereich zwischen 0,95 Meter und 1,21 Meter nicht sonderlich stark, sodass sich der Kopf der Testpersonen immer etwas unterhalb des mittleren Arrays befand.

Werden die jeweiligen Reihen der Quellenpositionen so betrachtet, dass die Ergebnisse in neun Höhenstufen aufgeteilt werden, fällt auf, dass mit abnehmender Höhe der Quellenposition die Präzision der Lokalisierung in der Höhe deutlich abnimmt. Vermutlich ist der Grund dafür der Wirkungsbereich der Wellenfeldsynthese-Module. Je niedriger die Quellenposition auf der Z-Achse liegt, desto geringer ist die Auswirkung der Module auf das Wellenfeld. Eine weitere Möglichkeit ist die Annahme, dass die Testpersonen aus psychologischen Gründen keine Töne nahe am oder unter dem Boden erwarten. Die niedrigste Quellenposition befindet sich bei diesem Versuch auf 18 Zentimetern Höhe und 1 Meter hinter dem unteren Lautsprecherarray. Die gleichmäßigen Werte auf der X-Achse lassen darauf schließen, dass die horizontale Lokalisierung auf allen Reihen gut funktioniert, was darauf hindeutet, dass die Lautsprecherdichte in der Vertikale zu gering ist, um Wellenfeldsynthese zu betreiben.

9 Schluss

9.1 Zusammenfassung

Ziel dieser Arbeit war es, den Bereich der Wellenfeldsynthese-Anlage der HAW Hamburg um die dritte Dimension zu erweitern. Um die Präzision der Lokalisation in der dritten Dimension zu ermitteln, wurde ein Versuch entwickelt. Zunächst wurden Grundlagen zur auditiven Wahrnehmung des Menschen und der Wellenfeldsynthese sowie weitere, übliche oder ähnliche, 3D-Audio ermöglichenden Wiedergabeverfahren dargestellt. Anschließend wurde ein Überblick über die genutzten Anlagen aus dem Wellenfeldsynthese-Labor der HAW Hamburg gegeben. Im darauf folgenden Abschnitt wurde die Wellenfeldsynthese-Software WONDER und MoWeC sowie die Erweiterung um den Hörtest vorgestellt, welche die Hauptbestandteile des Versuches darstellen. Im letzten Teil der Arbeit wurden der Aufbau, die Durchführung und die Ergebnisse des Hörtestes dargestellt, um herauszufinden, wie genau Positionen von Menschen visuell oder akustisch lokalisiert werden können.

9.2 Fazit

In dieser Arbeit wurde gezeigt, dass die einfache Erweiterung von Wellenfeldsynthese-Anlagen ein durchaus vielversprechendes Thema ist, dessen weitere Verfolgung sich lohnt. Die Ergebnisse in der Abweichung sind auf der Z-Achse, wie erwartet, etwas schlechter als auf der X-Achse ausgefallen, befinden sich aber dennoch in einem Bereich den man als präzise bezeichnen kann. Marvin Geitner beschreibt das Ergebnis seines Hörtestes als einen ansatzweise wahrnehmbaren, dreidimensionalen Höreindruck, der durch Alias-Effekte und durch die geringe Anzahl an zusätzlichen Lautsprechern gestört ist [Gei15]. Dass die Ergebnisse mit einer entsprechenden Anzahl an zusätzlichen Lautsprechern deutlich besser ausfallen, konnte in dieser Arbeit belegt werden. Ein exakter Vergleich erweist sich aufgrund der unterschiedlichen Methodik und dem ungleichen Umfang der Versuche als schwierig. Im Vergleich zu Montags Thesis fallen die Ergebnisse ähnlich positiv aus, obwohl kein VBAP in der Vertikale genutzt wurde [Mon11].

9.3 Ausblick

Um ein besseres Ergebnis im Höreindruck zu erreichen, könnten einige Veränderungen vorgenommen werden.

Eine Idee basiert auf einem spezifischen Panning über die drei vertikal angeordneten Lautsprecher-Arrays, während innerhalb der einzelnen Arrays auf horizontaler Ebene weiter Wellenfeldsynthese angewandt wird. Bei dem aktuellen Aufbau kommt es zu einem Lautstärkeabfall sobald sich die Quelle zwischen den Arrays befinden, da das System versucht, Wellenfeldsynthese zu betreiben, obwohl in der Vertikale eine zu geringe Dichte an Lautsprechern besteht. Diesen Lautstärkeabfall würde das Panning wieder ausgleichen und die jeweils über und unter der Quelle liegenden Lautsprecher in der Lautstärke regulieren. Somit würden in der Vertikale Phantomschallquellen genutzt werden, wohingegen in der Horizontale Wellenfeldsynthese betrieben wird.

Wie schon in Abschnitt 8.2.1 erwähnt, kam es auch zu Unterschieden im Frequenzspektrum der verschiedenen Lautsprechertypen. Um eine Lokalisation anhand des unterschiedlichen Klangs zu vermeiden, könnten Anpassungen am Software-Equalizer vorgenommen werden. So könnten entweder zwei Spuren mit demselben Tonmaterial und unterschiedlichen Equalizer-Einstellungen oder eine Tonspur, in welcher alle unterschiedlichen Frequenzen herausgefiltert werden, verwendet werden. Dabei ist jedoch fraglich, ob die Qualität des Tonmaterials während der Wiedergabe nicht zu sehr leidet.

Versuche mit einer Wand aus Wellenfeldsynthese-Lautsprecher-Arrays wären interessant, um einen Vergleich der Höhenwahrnehmung anstellen zu können.

Denkbar wäre eine Implementierung von VBAP. Dafür könnten 24 virtuelle Soundquellen - drei Arrays mit jeweils acht Lautsprechern - erzeugt werden, welche jeweils einem Lautsprecher zugeordnet sind. Bei diesem Ansatz handelt es sich zwar nicht mehr um Wellenfeldsynthese, da VBAP aber in der Lage ist, präzise Phantomschallquellen zu erzeugen, ist es dennoch eine mögliche Alternative.

Abbildungsverzeichnis

2.1	Sweet Spot Dreieck	4
2.2	Phantomschallquelle bei Stereo-Wiedergabe	5
2.3	Gehörrichtige Lautstärke	6
4.1	Der akustische Vorhang	11
4.2	Nachbildung der Wellenfront aus Elementarwellen	12
4.3	Mögliche Aufbauformationen einer Wellenfeldsynthese-Anlage.	13
4.4	Punktquelle	14
4.5	Linearquelle	15
4.6	Fokussierte Quelle	15
5.1	Aufbau der Wellenfeldsynthese-Anlage der HAW Hamburg.	18
5.2	TT1, HT23 und GT04 Marker	20
6.1	Screenshot der xWonder Oberfläche	22
6.2	Netzwerkstruktur der WFS-Anlage [Mic12]	25
7.1	Versuchs-GUI in verschiedenen Farbzuständen	30
8.1	Der Versuchsaufbau von der Seite gesehen (in Unity modelliert)	34
8.2	Der Versuchsaufbau frontal gesehen (in Unity modelliert)	35
8.3	Spektrogramm der Four Audio- und A7x-Lautsprecher	36
8.4	Versuchsaufbau zum Ermitteln des Frequenzspektrums	36
8.5	Eingänge und Equalizer der Adam A7x Lautsprecher	37
8.6	Die Abweichung auf der X- und Z-Achse im Zusammenhang mit dem Alter der Probanden dargestellt	39
8.7	Die Abweichung auf der X- und Z-Achse in Relation zur Audiophilie als Boxplot dargestellt	40
8.8	Boxplot der Abweichung von akustischen- und sichtbaren Quellen	41
8.9	Abweichung bei sichtbarer Quelle	42

8.10	Durchschnittliche absolute Abweichung der einzelnen Quellenpositionen der X-Achse	43
8.11	Durchschnittliche absolute Abweichung der einzelnen Quellenpositionen der Z-Achse	43
8.12	Darstellung der IDs als Matrix an der ungefähren Position im Versuchsaufbau mit Reihenbeschriftung an der rechten Seite	44
8.13	Auswertung der absoluten durchschnittlichen Abweichung pro Reihe von Quellenpositionen	45
8.14	Abweichungen gegenüber der Zeit bis zum Auslösen der X- und Z-Achse . . .	45

Tabellenverzeichnis

2.1	Tabelle des Luftschalldämmungskoeffizienten	6
7.1	Bedeutung der Farben in der Versuchs-GUI	30
7.2	Log-Datei Format	32
8.1	Absolute durchschnittliche Abweichung der Messwerte auf der X- und Z-Achse mit Ton als Ziel	42
8.2	Durchschnittliche Abweichung der Messwerte auf der X- und Z-Achse mit sichtbarem Ziel	42

Literaturverzeichnis

- [3DS15] 3D Sound Labs - 3D Sound One binaural synthesis headphones. <http://www.3dsoundlabs.com/>, 2015. Zugriff: 14.04.2016.
- [Ahr12] Jens Ahrens. *Analytic Methods of Sound Field Synthesis*. Springer Science & Business Media, Berlin Heidelberg, 2012. aufl. edition, 2012.
- [BB08] Jens Blauert and Jonas Braasch. *Handbuch der Audiotechnik*, chapter Räumliches Hören, pages 87–121. Springer Berlin Heidelberg, Berlin, Heidelberg, 2008.
- [Ber88] A. J. Berkhout. A holographic approach to acoustic control. *J. Audio Eng. Soc.*, 36(12):977–995, 1988.
- [Bla13] Jens Blauert. *The Technology of Binaural Listening*. Springer Science & Business Media, Berlin Heidelberg, 2013. aufl. edition, 2013.
- [Bla15] Paul E. Black. Dictionary of Algorithms and Data Structures - Fisher-Yates shuffle. <https://xlinux.nist.gov/dads//HTML/fisherYatesShuffle.html>, 2015. Zugriff: 08.05.2016.
- [BP04] M.A.J. Baalman and D. Plewe. Wonder - a software interface for the application of wave field synthesis in electronic music and interactive sound installations. In *International Computer Music Conference 2004, Miami, 1-6 November 2004*, 2004.
- [Chr14] Carola Christoffel. Modifikation der Software einer Wellenfeldsyntheseanlage zur Wiedergabe fokussierter Quellen in Abhängigkeit der Zuhörerposition. Bachelorarbeit, Hochschule für Angewandte Wissenschaften Hamburg, 2014.
- [DDHW14] Michael Dickreiter, Volker Dittel, Wolfgang Hoeg, and Martin Wöhr. *Handbuch der Tonstudioteknik*. Walter de Gruyter GmbH & Co KG, 2014.
- [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.

- [Gei15] Marvin Geitner. Erweiterung der Wellenfeldsynthese-Software WONDER um 3D-Quellenpositionen und zusätzliche Lautsprecherkomponenten. Bachelorarbeit, Hochschule für Angewandte Wissenschaften Hamburg, 2015.
- [HO03] Christiaan Huygens and Wilhelm Ostwald. *Abhandlung über das licht - Worin die ursachen der vorgänge bei seiner zurückwerfung und brechung und besonders bei der eigenthümlichen brechung des isländischen spathes dargelegt sind*. W. Engelmann, Leipzig, 1903.
- [HS05] R. Höldrich and A. Sontacchi. Wellenfeldsynthese - erweiterungen und alternativen. In *DAGA 2005, 31. Deutsche Jahrestagung für Akustik, München*, 2005.
- [ISO99] DIN ISO. Akustik - Dämpfung des Schalls bei der Ausbreitung im Freien - Teil 2: Allgemeines Berechnungsverfahren (ISO 9613-2:1996). DIN ISO DIN ISO 9613-2:1999-10, Deutsches Institut für Normung, International Organization for Standardization, 1999.
- [ISO03] ISO. Acoustics – Normal equal-loudness-level contours. ISO 226:2003, International Organization for Standardization, 2003.
- [Lüb15] Lübeck, Präsident der Fachhochschule. *Design & Research*. kassel university press GmbH, Kassel, 2015.
- [Mic12] Michael Makarski, Anselm Goertz, Rainer Thaden, Jochen Kleber. Entwicklung einer WFS-Anlage mit FIR-Entzerrung und Dante Audionetzwerk. *Institut für Akustik und Audiotechnik and Four Audio GmbH & Co. KG*, 2012.
- [Mon11] Matthew N. Montag. Wave Field Synthesis In Three Dimensions By Multiple Line Arrays. Masterthesis, University of Miami, 2011.
- [Nog12] Malte Nogalski. Gestengesteuerte Positionierung von Klangquellen einer Wellenfeldsynthese-Anlage mit Hilfe eines kamerabasierten 3D-Tracking-Systems. Bachelorarbeit, Hochschule für Angewandte Wissenschaften Hamburg, 2012.
- [Oel07] Helmut Oellers. Der akustische Vorhang. <http://www.syntheticwave.de/akustischer%20Vorhang.htm>, 2007. Zugriff: 02.03.2016.
- [Oel10] Helmut Oellers. Reduktion von Aliasingeffekten. Patent, August 2010. DE 102009006762 A1.

- [Oss16] OSSIC auto-calibrating 3d audio headphones. <http://www.ossic.com/>, 2016. Zugriff: 14.04.2016.
- [PKE14] Cesare Parise, Katharina Knorre, and Marc O. Ernst. Natural auditory scene statistics shapes human spatial hearing. *PNAS*, 111(16):6104–6108, 2014.
- [Pul97] Ville Pulkki. Virtual sound source positioning using vector base amplitude panning. *Journal of the Audio Engineering Society*, 45(6):456–466, June 1997.
- [Rob56] R. S. Robinson, D. W. und Dadson. A re-determination of the equal-loudness relations for pure tones. *British Journal of Applied Physics*, 7:166–181, May 1956.
- [RVRL13] Neil T Roach, Madhusudhan Venkadesan, Michael J Rainbow, and Daniel E Lieberman. Elastic energy storage in the shoulder and the evolution of high-speed throwing in homo. *Nature*, 498:483–6, 2013 Jun 27 2013.
- [SA10a] Sascha Spors and Jens Ahrens. Reproduction of Focused Sources by the Spectral Division Method. 4th international symposium on communications, control and signal processing (isccsp), Deutsche Telekom Laboratories, Technische Universität Berlin, 2010.
- [SA10b] Sascha Spors and Jens Ahrens. Sound Field Reproduction Using Planar and Linear Arrays of Loudspeakers. *Ieee transactions on audio, speech, and language processing*, vol. 18, no. 8, 2010.
- [Wei09] Stefan Weinzierl. *Handbuch der Audiotechnik*. Springer-Verlag, Berlin Heidelberg New York, 1. aufl. edition, 2009.
- [Wer09] Ulf-J. Werner. *Schallschutz und Raumakustik - Handbuch für Theorie und Praxis*. Beuth Verlag GmbH, Frankfurt am Main, 1. aufl. edition, 2009.
- [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.

Hiermit versichere ich, dass ich die vorliegende Arbeit ohne fremde Hilfe selbständig verfasst und nur die angegebenen Hilfsmittel benutzt habe.

Hamburg, 12. Mai 2016 Rutkay Küpelikilinc