



Hochschule für Angewandte Wissenschaften Hamburg
Hamburg University of Applied Sciences

Bachelorarbeit

Alexander Witek

**Virtuelle Raumakustik der WFS-Anlage mit Hilfe von
Soundfield-Mikrofonaufnahmen**

*Fakultät Technik und Informatik
Studiendepartment Informatik*

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

Alexander Witek

**Virtuelle Raumakustik der WFS-Anlage mit Hilfe von
Soundfield-Mikrofonaufnahmen**

Bachelorarbeit eingereicht im Rahmen der Bachelorprüfung

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

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

Eingereicht am: 16. Oktober 2014

Alexander Witek

Thema der Arbeit

Virtuelle Raumakustik der WFS-Anlage mit Hilfe von Soundfield-Mikrofonaufnahmen

Stichworte

Wellenfeldsynthese, Soundfield-Mikrofon, Ambisonics, B-Format, 3D-Audio, Impulsantwort, Raumakustik, virtuelles Mikrofon

Kurzzusammenfassung

In dieser Arbeit wird eine virtuelle Raumakustik für die Wellenfeldsynthese-Anlage der HAW Hamburg mit Hilfe eines Soundfield-Mikrofons für einen Zuhörer erzeugt. Dabei werden Schritt für Schritt die nötigen Grundlagen erklärt, welche für dieses Vorhaben erforderlich sind, sowie die praktische Durchführung beschrieben. Zum Schluss wird das Ergebnis bewertet und eine Aussicht für Verbesserungen vorgegeben.

Alexander Witek

Title of the paper

Virtual room acoustics at wave field synthesis with Soundfield recordings

Keywords

wave field synthesis, Soundfield-Microphone, Ambisonics, B-Format, 3D-Audio, impulse response, room acoustics, virtual microphone

Abstract

This thesis describes an implementation of room acoustics at the wave field synthesis with Soundfield-Microphone recordings for a single listener. The principles that have to be known will be explained as well as the realization. At the end the results will be rated and some prospects given.

Inhaltsverzeichnis

| | | |
|----------|--|-----------|
| 1 | Einleitung | 1 |
| 1.1 | Motivation | 1 |
| 1.2 | Ziel der Arbeit | 1 |
| 1.3 | Themenabgrenzung | 1 |
| 1.4 | Struktur der Arbeit | 2 |
| 2 | Einführung in das 3D-Audio | 3 |
| 2.1 | Das Ambisonics-Verfahren | 3 |
| 2.2 | Wellenfeldsynthese der HAW Hamburg | 4 |
| 2.2.1 | Das Labor | 5 |
| 2.2.2 | Die Idee der Wellenfeldsynthese | 5 |
| 2.2.3 | Die Funktion | 7 |
| 2.2.4 | Die Rolle der Wellenfeldsynthese-Anlage in dieser Arbeit | 7 |
| 3 | Die Messvorrichtung | 10 |
| 3.1 | Die Interfaces | 10 |
| 3.1.1 | DBS1 Dante Breakout Box (Four Audio) | 10 |
| 3.1.2 | Liquid Saffire 56 (Focusrite) | 11 |
| 3.1.3 | US 122 (Tascam) | 11 |
| 3.2 | Das Soundfield-Mikrofon | 11 |
| 3.2.1 | Der Aufbau des Soundfield-Mikrofons | 12 |
| 3.3 | Die Lautsprecher | 13 |
| 3.3.1 | Dodekaeder-Lautsprecher | 13 |
| 3.3.2 | Adam A8X | 13 |
| 3.3.3 | Fazit | 14 |
| 4 | Messung der Raumimpulsantworten | 15 |
| 4.1 | Systeme | 15 |
| 4.2 | Methoden zur Messung der Raumimpulsantworten | 16 |
| 4.2.1 | Dirac-Stoß | 16 |
| 4.2.2 | Ermittlung durch deterministisches Rauschen | 17 |
| 4.2.3 | Ermittlung durch den Sinus-Sweep | 18 |
| 4.3 | Die Durchführung | 19 |
| 5 | Verwendung der gemessenen Daten | 20 |
| 5.1 | Umwandlung ins B-Format | 20 |

| | | |
|----------|--|-----------|
| 5.2 | Berechnung der Raumimpulsantworten für die einzelnen Kanäle | 20 |
| 5.2.1 | Umwandlung von B-Format auf einzelne Kanäle | 21 |
| 5.2.2 | Herausrechnen der Impulsantwort | 22 |
| 5.3 | Überprüfung im schallreflexionsarmen Raum | 22 |
| 5.4 | Die Faltung einer Quelle mit der Impulsantwort | 24 |
| 5.5 | Erstellen des VST-Plugins zur räumlichen Wiedergabe einer Quelle . . . | 25 |
| 6 | Vergleich zwischen Original und Simulation | 27 |
| 6.1 | Aufbau der originalen Aufnahme | 27 |
| 6.2 | Vergleich beider Aufnahmen | 27 |
| 6.3 | Auswertung | 27 |
| 7 | Schluss | 29 |
| 7.1 | Fazit | 29 |
| 7.2 | Ausblick | 29 |
| | Abbildungsverzeichnis | 30 |
| | Literaturverzeichnis | 32 |
| | Anhang MATLAB Code | 33 |
| 1 | Berechnung des Rauschens und der zugehörigen Impulsantwort für acht Kanäle | 33 |
| 2 | Faltung eines Signals mit den Impulsantworten | 34 |

1 Einleitung

Dieses Kapitel soll den Leser an die Thematik dieser Arbeit heranführen. Dabei wird erklärt, welches Ziel verfolgt wird, was nicht Bestandteil der Arbeit ist und die Struktur erläutert.

1.1 Motivation

Die Motivation dieser Arbeit besteht darin, den Eindruck eines Hörerlebnisses in einem bestimmten Raum oder Konzertsaal zu erfassen und in einem anderen Raum wiederzugeben.

1.2 Ziel der Arbeit

Das Ziel dieser Arbeit ist es, eine virtuelle Quelle auf der Wellenfeldsynthese so erklingen zu lassen, als würde sich diese, zusammen mit einer Person, in einem bestimmten Raum befinden. Dabei wird zuerst die Raumimpulsantwort gemessen und anschließend mit der virtuellen Quelle gefaltet. Verwendet werden soll dabei ein möglichst unkomplizierter und flexibler Aufbau der Messvorrichtung, in diesem Fall ein Soundfield-Mikrofon erster Ordnung zusammen mit der Wellenfeldsynthese, welche für die Acht-Kanal-Wiedergabe dient. Die Berechnung geschieht mit Hilfe von MATLAB, jedoch wird auch ein mögliches Modell eines VST-Plugins vorgestellt. Anschließend soll beurteilt werden, ob sich das Verfahren eignet, um weiterentwickelt und die Qualität der Wiedergabe dadurch gesteigert werden kann.

1.3 Themenabgrenzung

Nicht behandelt wird in dieser Arbeit die Erweiterung auf ein System mit mehr als acht Kanälen. Dafür wäre ein Soundfield-Mikrofon höherer Ordnung notwendig. Außerdem wird die Wellenfeldsynthese nur dafür verwendet um ein Lautsprecher-Setup zu simulieren.

Daher ist die Wiedergabe nur für den Sweet Spot korrekt, man kann sich daher innerhalb der Anlage nicht frei bewegen.

1.4 Struktur der Arbeit

Im zweiten und dritten Kapitel werden Grundlagen erklärt, welche für das Verständnis dieser Arbeit notwendig sind. Im vierten Kapitel wird erklärt, welche Methoden zur Messung der Raumimpulsantwort bekannt sind und die Durchführung beschrieben. Im fünften Kapitel werden die gemessenen Daten schließlich verarbeitet. Im sechsten Kapitel werden die Originalaufnahmen mit den simulierten Aufnahmen verglichen und bewertet.

2 Einführung in das 3D-Audio

Das Thema 3D-Audio ist ein sehr weites Feld, mit dem sich die HAW Hamburg seit einiger Zeit beschäftigt. Zu ihm gehört unter anderem die Wellenfeldsynthese-Anlage, welche in den folgenden Abschnitten erläutert wird [Fohl (2013)]. Als neues Gebiet soll nun die Möglichkeit realisiert werden, räumlich aufgelöste Raumimpulsantworten zu ermitteln, um Quellen räumlich erklingen lassen zu können. Um dies zu realisieren ist ein Grundwissen nötig, welches in dieser Arbeit, auf das wichtigste beschränkt, weitergegeben wird. Im folgenden Abschnitt wird das Ambisonics-Verfahren erklärt, welches im Gegensatz zur herkömmlichen Wiedergabe in der Lage ist tatsächlich dreidimensionale Signale zu erzeugen und wiederzugeben.

2.1 Das Ambisonics-Verfahren

Das Ambisonics-Verfahren ist ein Verfahren, welches ein gesamtes Klangfeld dreidimensional aufzeichnen und wiedergeben kann [Trevino u. a. (2010)]. Dabei ist es im Gegensatz zur kanalorientierten Aufnahme sehr flexibel einsetzbar. Aus dem räumlich aufgelösten Schalldruck und der Schallschnelle können durch mathematische Berechnungen die jeweiligen Signale für einzelne Kanäle berechnet werden. Somit kann man durch eine einzige Messung eines Raumes neben Monophonie unterschiedliche Mehrkanal-Tonsysteme erzeugen. Die bekanntesten sind Stereo- und Quadrophonie sowie die Formate 5.1, 7.1 und 8.0.

Signale, welche im Ambisonics-Verfahren verwendet werden liegen in unterschiedlichen Formaten vor. Das A-Format beinhaltet die rohen Signale, welche mit einem Soundfield-Mikrofon aufgenommen wurde. Das B-Format hingegen enthält Signale für bestimmte Achsen eines Raumes und den Gesamtschalldruck. Das B-Format errechnet sich aus dem A-Format. Das Prinzip dahinter ist, dass die aufgenommenen Signale in einem einzigen Punkt auftreffen sollen. Praktisch ist es jedoch nicht möglich, da die einzelnen Mikrofon-Kapseln eine bestimmte Größe haben und somit die Signale vor dem Punkt auftreffen. Außerdem müssen den einzelnen Achsen die korrekten Signale zugeteilt werden. Dies geschieht mit Hilfe der Signalverarbeitung. Der Hersteller eines

Soundfield-Mikrofon muss aus diesem Grund eine Soft- oder Hardware anbieten, welche die Signale des A-Formats in das B-Format korrekt übersetzt.

Da es sich um ein Soundfield-Mikrofon erster Ordnung handelt, besteht das B-Format aus den Kanälen W, X, Y, Z. W enthält den Schalldruckanteil und X, Y und Z jeweils den Schallschnellenanteil der zugehörigen Achse (siehe Abbildung 2.1). Durch Dekodieren des B-Formats lässt sich ein aufgenommenes Signal für einen bestimmten Kanal rekonstruieren. Sofern es sich um Kanalordnungen handelt, welche den regulären Polygonen entsprechen, ist die Berechnung recht unkompliziert. Soll es jedoch um eine komplexe Anordnung der Kanäle gehen, so werden bereits jede Menge fertige Decoder angeboten, wie zum Beispiel das VST-Plugin der Firma Blue Ripple Sound, welches ein 22.2-Format dekodieren kann [BlueRippleSound (2014)]. Möchte man sich dagegen einen individuellen Decoder erstellen, so existiert freie Software, welche zur Unterstützung benutzt werden kann, da sich die korrekte Berechnung der Parameter als schwierig erweist [Heller u. a. (2012)].

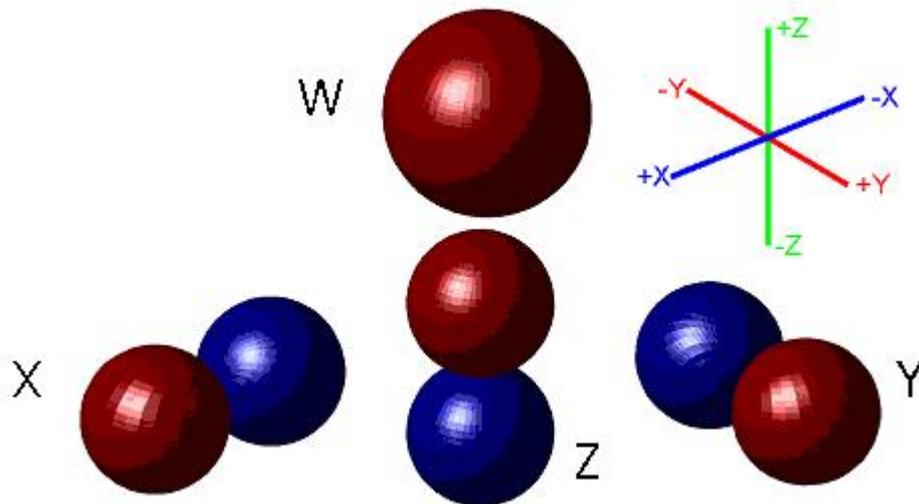


Abbildung 2.1: Das B-Format erster Ordnung

Quelle: <http://uod-true-multi-channel-mixing.wikispaces.com/Ambisonics>

2.2 Wellenfeldsynthese der HAW Hamburg

In diesem Abschnitt wird einerseits das Labor, in dem die Wellenfeldsynthese-Anlage steht, beschrieben und andererseits die Funktion dieser erläutert. Es soll ein kurzer

Überblick verschafft werden, um die Beziehung zwischen der Wiedergabe und dem Raumklang deutlich zu machen.

2.2.1 Das Labor

Die Wellenfeldsynthese-Anlage besteht aus 26 Modulen der Firma Four Audio, welche auf einer rechteckigen Fläche von etwa 5x6 Metern aufgebaut ist. Jedes dieser Module besteht aus zwei Tieftonchassis und acht Reihen mit je drei Breitbandchassis und 10 cm Abstand zueinander. Insgesamt stehen damit 208 Spuren und 676 Lautsprecher zur Verfügung.

Innerhalb der Module befindet sich außerdem ein Trackingsystem der Firma Advanced Realtime Tracking GmbH. Dieses funktioniert dank sechs Infrarotkameras, welche wegen der kurzen Reichweite leider auf ein Feld von 4x4 Metern begrenzt sind [Nogalski (2012)]. Auf der Hauptseite der Anlage befindet sich außerdem eine Powerwall, bestehend aus sechs Bildschirmen, welche durch die AMD Eyefinity Technologie zu einem einzigen Bildschirm zusammengeschlossen werden.

Weder das Trackingsystem noch die Powerwall werden in dieser Arbeit verwendet.

Im Labor stehen vier Rechner, welche für die Wellenfeldsynthese zuständig sind. Die Wellenfeldsynthese-Anlage wird über einen Rechner, der als Benutzerschnittstelle dient, gesteuert und kommuniziert mit einem Wellenfeldsynthese-Server, der wiederum mit zwei Nodes verbunden ist, welche die endgültige Berechnung der jeweiligen Signale für die einzelnen Spuren übernehmen (siehe Abbildung 2.2).

Das Labor ist mit einem Dante-Netzwerk ausgestattet. In diesem Netzwerk können alle unterstützten Geräte, wie zum Beispiel die Wellenfeldsynthese-Module, PCI-Karten oder Rechner mit Dantes virtueller Soundkarte, miteinander verbunden oder deren Signale weitergeleitet werden.

2.2.2 Die Idee der Wellenfeldsynthese

Seit vielen Jahrzehnten wird versucht, Hörerlebnisse immer realer darzustellen. Die Anfänge machte die Stereophonie, die Quadrophonie folgte etwas später, konnte sich jedoch nie wirklich durchsetzen, da sich die Hersteller nicht auf ein gemeinsames System einigen konnten. Heutzutage wird es von dem Surround-Sound ersetzt. Doch auch die moderneren Mehrkanal-Systeme haben alle einen Nachteil, denn die optimale Wirkung kann nur im Sweet Spot, welcher der Punkt ist, für den die Signale der jeweiligen Lautsprecher berechnet wurden, erlebt werden.

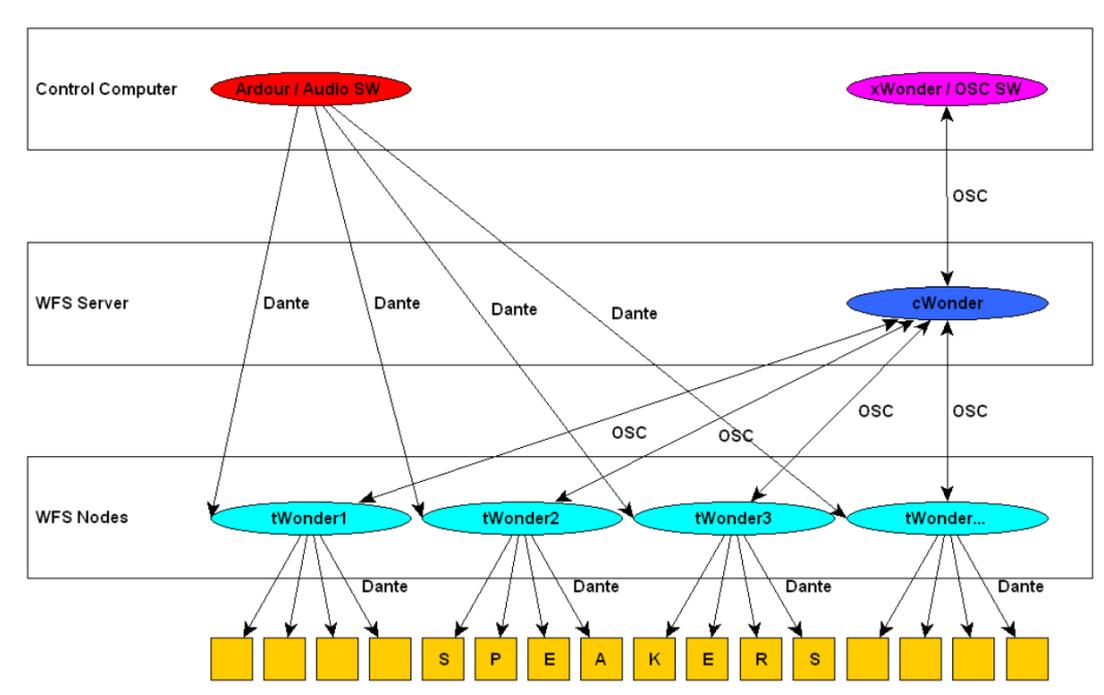


Abbildung 2.2: Netzwerkstruktur einer WFS-Anlage
[Makarski u. a. (2012)]

Anders hingegen sieht es bei der Wellenfeldsynthese aus. Sie ist darauf ausgelegt ein ganzes Feld korrekt zu beschallen. Das heißt, dass für alle Zuhörer die virtuellen Quellen an ihren echten Positionen wahrgenommen werden. Einschränkungen gibt es, wenn sich die Quelle in dem Feld befindet, denn es kann für einige Zuhörer nicht die richtige Position einer Quelle vermittelt werden. Dieser Fall tritt ein, wenn sich der Zuhörer zwischen der Quelle und den aktiven Lautsprechern befindet. Es entsteht der Eindruck, als würde der Ton aus den Lautsprechern und nicht vom Punkt der virtuellen Quelle entstehen. Dieses Problem wurde jedoch für eine einzelne Person behoben. Wie genau dies geschieht kann unter [Christoffel (2014)] nachgelesen werden. Kurz zusammengefasst wird die Position der Person mit Hilfe des Tracking-Systems erfasst und berechnet, ob die Person zwischen der Quelle und den Lautsprechern steht. Falls dieser Fall eintritt, so wird das Signal von den gegenüberliegenden Lautsprechern berechnet.

2.2.3 Die Funktion

Die Funktion der Wellenfeldsynthese beruht auf dem Huygensschen Prinzip, welches besagt, dass jeder Punkt einer Wellenfront auch Ausgangspunkt einer neuen Welle, der sogenannten Elementarwelle, gesehen werden kann [Corteel und Caulkins].

Schallquellen breiten sich kugelförmig aus. Stellt man sich nun eine Seite der Wellenfeldsynthese-Anlage als eine Wand, die acht Reihen der Lautsprecher als Löcher und die Schallquelle mittig hinter der Wand vor, so würde die Wellenfront der Quelle halbkreisförmig auf die Wand treffen und in jedem Loch eine Elementarwelle bilden. In der Mitte der Wand zuerst und nach außen hin verzögert. Durch die neuen Wellenfronten entsteht nun eine große Front, auch synthetisierte Wellenfront genannt, welche den physikalischen Verlauf der Welle ohne Wand darstellt (siehe Abbildung 2.3).

Für die korrekte Berechnung einer virtuellen Quelle sind die Rechner zuständig. Diese simulieren das Auftreffen der Wellenfront auf die Lautsprecher nach dem Huygensschen Prinzip.

Im Raum können ebenfalls Quellen positioniert werden, allerdings verhält sich das Verzögern der Signale genau andersherum als im Beispiel mit einer Quelle hinter der Wand. Es wird nach innen hin verzögert, denn die Wellenfront einer Quelle im Raum ist zu den Lautsprechern zuerst konkav. Durchläuft sie den Punkt der Quelle, so wandelt sie sich in eine konvexe Wellenfront um.

Je nachdem, wo sich die Quelle befindet, werden unterschiedliche Lautsprecher aktiv. Dafür sorgen ebenfalls die Rechner.

2.2.4 Die Rolle der Wellenfeldsynthese-Anlage in dieser Arbeit

Die Wellenfeldsynthese-Anlage wird in dieser Arbeit benutzt, um ein Acht-Kanal-Setup virtuell zu erstellen (siehe Abbildung 2.4). Es handelt sich um acht ebene Wellen, welche die verschiedenen Raumrichtungen darstellen sollen. Die Wellenfeldsynthese-Anlage eignet sich bestens dazu, um verschiedene Setups, ohne großes Umbauen, wie es ohne die Anlage der Fall wäre, zu erstellen, zu erweitern oder zwischen ihnen zu wechseln.

Der Zuhörer soll sich anschließend in die Mitte des Raumes, in den Sweet Spot, stellen, um die bestmögliche Wirkung der Wiedergabe zu erfahren.

In der Zukunft soll ermittelt werden, ob und wie sich die Wellenfeldsynthese-Anlage erweitern lässt, um auch Höheninformationen wiedergeben zu können.

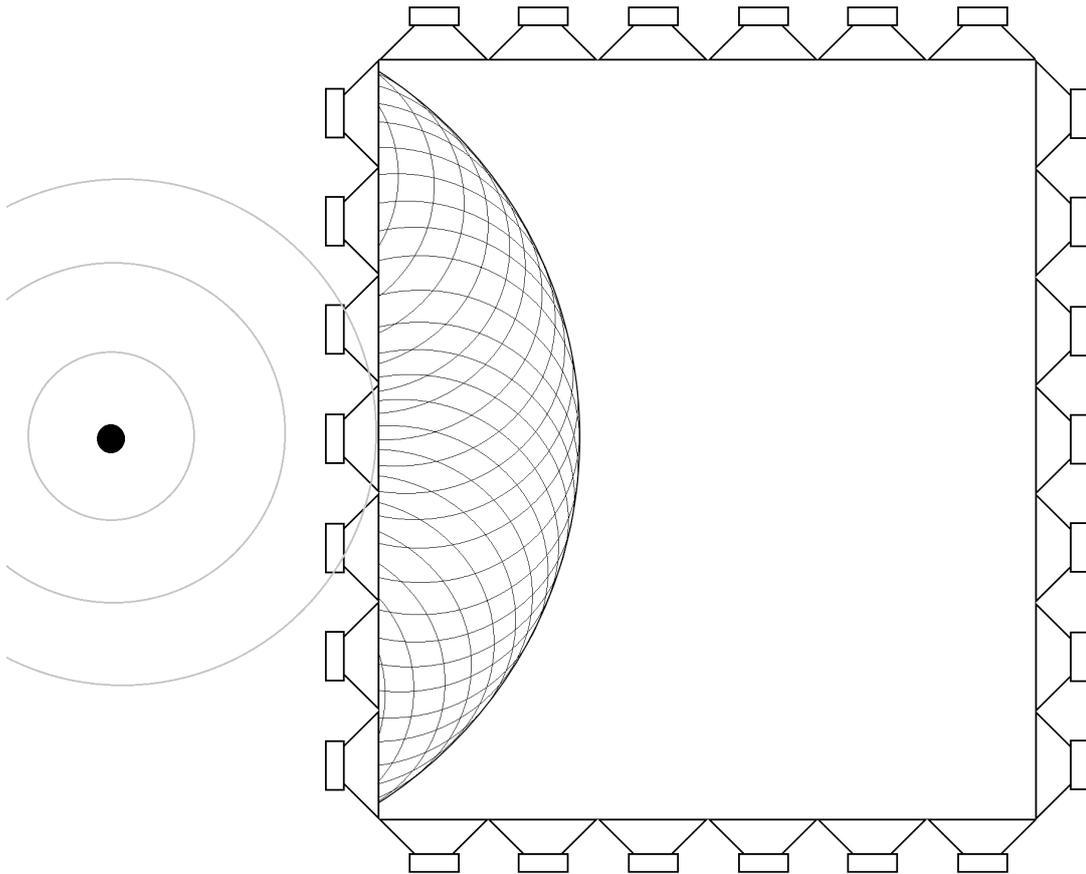


Abbildung 2.3: Ausbreitung der synthetisierten Wellenfront einer virtuellen Schallquelle
[Christoffel (2014)]

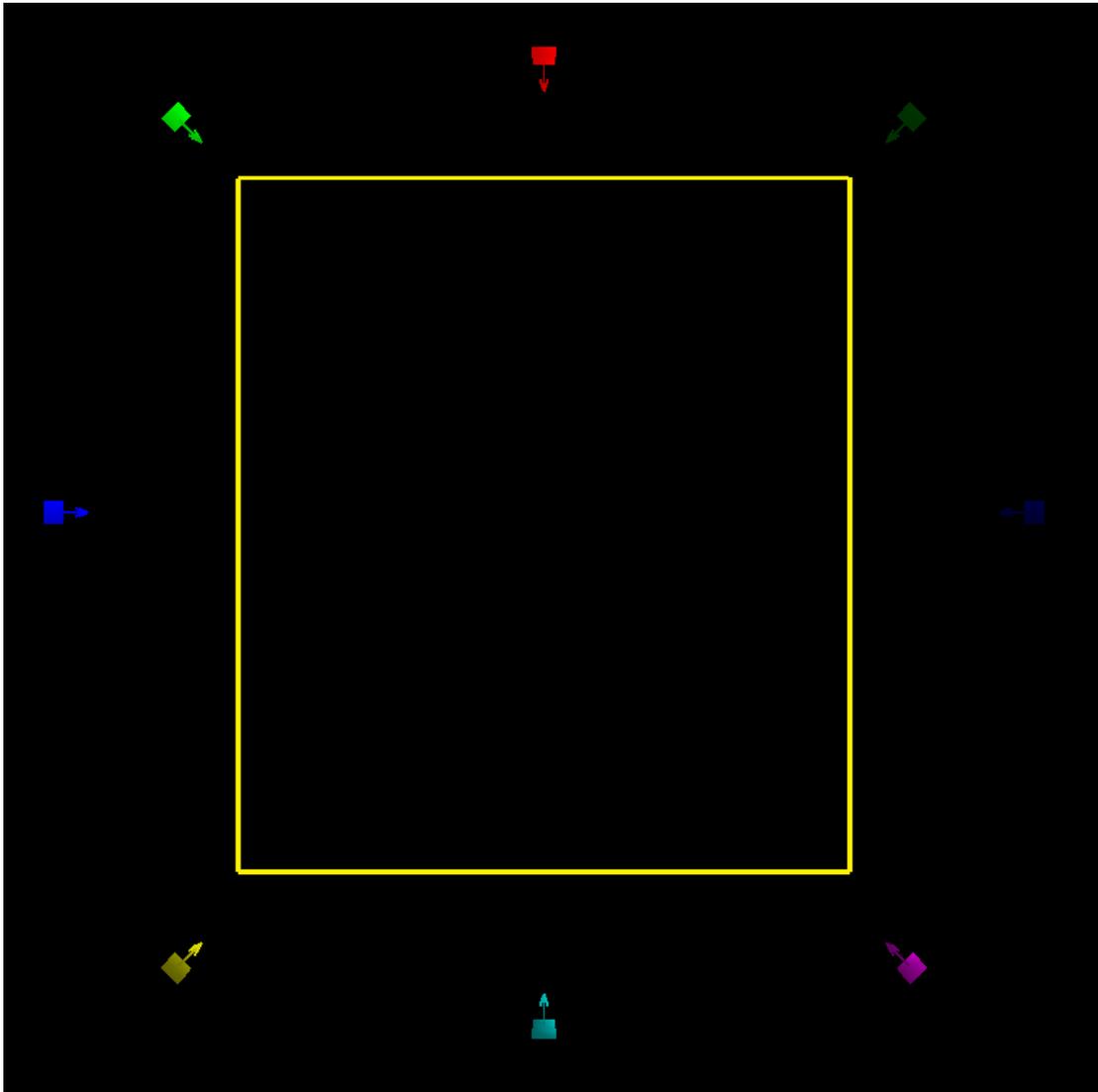


Abbildung 2.4: Darstellung der Raumakustik der Wellenfeldsynthese (gelbes Rechteck) in Xwonder

3 Die Messvorrichtung

Dieses Kapitel stellt die Geräte vor, die zur Messung der Raumimpulsantworten, in dieser Arbeit verwendet werden. Die Durchführung findet allerdings erst im nächsten Kapitel statt.

Anzumerken sei, dass mehr Interfaces und Lautsprecher vorgestellt werden, als am Ende nötig sind, da sich im Laufe der Messungen herausstellte, dass einige Geräte nicht wie erwartet funktionierten.

3.1 Die Interfaces

Die Interfaces spielen eine wichtige Rolle in der Messvorrichtung, da je besser ein Interface mit den restlichen Komponenten harmoniert, desto einfacher ist später eine korrekte Verarbeitung der aufgenommenen Daten. Zunächst wurden mehrere Interfaces gleichzeitig benutzt, was jedoch zu sehr ungenauen Ergebnissen geführt hatte. Inzwischen ist es möglich die Raumimpulsantworten mit nur einem Interface zu ermitteln.

3.1.1 DBS1 Dante Breakout Box (Four Audio)

Die Dante Breakout Box von Four Audio, auch DBS1 genannt, ist ein kleines Interface mit je zwei analogen Ein- und Ausgängen, welches per Ethernet an ein Dante-Netzwerk angeschlossen wird. Es können zeitgleich nur ein bis zwei Ein- oder ein bis zwei Ausgänge verwendet werden. Für diese Arbeit bedeutet das einen höheren Aufwand, da für die Verarbeitung der Daten ein Kanal mehr benötigt wird.

Wichtig für das Soundfield-Mikrofon ist, dass das Interface Phantomspeisung bietet. Auch, dass die Empfindlichkeit digital eingestellt werden kann, ist von Vorteil. Analoges Einstellen gleicher Pegel ist ziemlich schwer, allerdings sehr wichtig für das Vorhaben mit dem Mikrofon.

Probleme bereitet das DBS1 im Zusammenspiel mit den restlichen Geräten im Dante-Netzwerk der HAW Hamburg, da es sich nicht korrekt synchronisieren kann. An diesem Problem wird aktuell gearbeitet. Schließt man das DBS1 an ein Netzwerk ohne andere

Geräte, bereitet es keine weiteren Probleme mehr. Dass das Gerät über Ethernet verbunden wird ist einerseits ein Vorteil, da jedes vernünftige Notebook und jeder Desktop-Rechner einen Ethernet-Anschluss hat. Andererseits benötigt man zusätzliche Hardware, wie zum Beispiel einen Switch und Ethernet-Kabel.

Für das Soundfield-Mikrofon würden zwei dieser Interfaces verwendet werden. Ein weiteres wäre nötig, um den Lautsprecher anzusteuern. Leider lieferte der Hersteller bisher kein weiteres Interface, womit die Verwendung dieser Geräte aufgeschoben wurde.

3.1.2 Liquid Saffire 56 (Focusrite)

Das Liquid Saffire 56 der Firma Focusrite ist ein Audio-Interface, welches acht analoge Ein- und zehn Ausgänge bietet. Daneben bietet es zahlreiche weitere Anschlüsse, welche hier nicht verwendet werden. Angeschlossen wird das Interface per Firewire an einen Rechner.

Nachdem es einige Schwierigkeiten mit der Synchronisierung zwischen Wiedergabe und Aufnahme gab, da zunächst zu wenige DBS1-Module vorhanden waren und auf alternative Interfaces gesetzt werden musste, wurde dieses Interface in diese Arbeit eingeführt. Mit ihm lässt sich gleichzeitig etwas wiedergeben und aufnehmen. Genau das wird benötigt, um die Raumimpulsantwort zu bestimmen.

Ein Nachteil dieses Interfaces ist, dass es per Firewire angeschlossen wird. Möchte man mobil arbeiten und hat ein Notebook ohne einen solchen Anschluss, gibt es kaum eine Möglichkeit ihn nachzurüsten. Allerdings gleicht die einfache und problemlose Bedienung diesen Punkt wiederum aus.

3.1.3 US 122 (Tascam)

Das Tascam US 122 ist ein kleines Audio/MIDI Interface, welches per USB an einen Computer angeschlossen wird. Es bietet zwei analoge Ein- und Ausgänge und zusätzlich einen Kopfhörerausgang an. Ebenso hat das Interface Phantomspeisung und einen MIDI Ein- und Ausgang. Beides wird in dieser Arbeit jedoch nicht benötigt.

Insgesamt spielt das Interface keine große Rolle, da es nur dafür verwendet wird, um Aufnahmen einer E-Gitarre zu erzeugen.

3.2 Das Soundfield-Mikrofon

Das Soundfield-Mikrofon ist das Kernelement in dieser Arbeit. Zuerst wird erklärt wie das Soundfield-Mikrofon aufgebaut ist und wie es sich von anderen Aufnahmevorrichtungen

gen unterscheidet. Das Ambisonics-Verfahren, welches für die aufgenommenen Signale verwendet wird, wurde bereits in 2.1 erläutert.

3.2.1 Der Aufbau des Soundfield-Mikrofons

Das Soundfield-Mikrofon erster Ordnung, in diesem Falle das SPS200 von TSL-Products, besteht aus vier Mikrofon-Kapseln mit Nierencharakteristik, welche sehr nahe aneinander in einem Tetraeder angeordnet sind.

Das Mikrofon benötigt 48 Volt Phantomspannung, da es sich um ein Kondensatormikrofon handelt. Angeschlossen wird es über vier XLR-Steckverbinder. Jede Kapsel hat einen eigenen Anschluss und einen fest definierten Namen. Es handelt sich dabei um die Abkürzungen LF, LB, RF und RB. Das **L** steht jeweils für **left**, das **R** für **right** und **F** und **B** für **front** und **back**. Das Logo auf dem Mikrofon gibt die Richtung an, in die es zeigt.

Das SPS200 ist eine sehr einfache Form des Soundfield-Mikrofons, da es im A-Format aufnimmt. Genauer gesagt heißt das, dass wir tatsächlich nur die rohen Signale erhalten, wie sie an den einzelnen Kapseln ankommen.



Abbildung 3.1: Das TSL Products SoundField SPS200 Mikrofon
Quelle: <http://www.tslproducts.com/>

3.3 Die Lautsprecher

Für die Wiedergabe der Messsignale zur Bestimmung der Impulsantwort werden zwei Lautsprecher verwendet und anschließend miteinander verglichen. Es handelt sich um einen passiven Dodekaeder Lautsprecher und einen Adam A8X Aktivmonitor. Beide haben ihre Vor- und Nachteile.

3.3.1 Dodekaeder-Lautsprecher

Der Dodekaeder-Lautsprecher kann, zusammen mit der passenden Endstufe, in dem Medientechnikdepartement der HAW Hamburg ausgeliehen werden.

Der Verstärker bietet zwei mal 250 Watt Ausgangsleistung, welche zwar auch zu 500 Watt zusammengeschaltet werden kann, jedoch dafür die Impedanz des Lautsprechers nicht mehr kompatibel ist. Im schlimmsten Fall würde die Endstufe Schaden nehmen. Doch selbst bei 250 Watt möchte man nicht zusammen in einem Raum mit dem Lautsprecher sitzen.

Der Dodekaeder-Lautsprecher hat auf jeder Fläche einen Breitbandlautsprecher, insgesamt sind es zwölf. Der Lautsprecher hat durch seine Form eine kugelförmige Richtcharakteristik. Der Nachteil ist das hohe Gewicht und das korrekte Anschließen, da er erst nach vielen Versuchen mit mehreren Interfaces einen Ton von sich gab. Durch die Verkettung der Geräte miteinander entstand ein dauerhafter Pfeifton, welcher die Aufnahme störte.

3.3.2 Adam A8X

Eine Alternative für die Wiedergabe bietet der Adam A8X Aktivmonitor. Es ist ein 2-Wege Lautsprecher im Bassreflexgehäuse mit insgesamt 200 Watt Ausgangsleistung. Die Richtcharakteristik entspricht in etwa einem Gitarrenverstärker.

Die Vorteile dieses Lautsprechers liegen deutlich auf der Hand. Einerseits sind das geringe Gewicht und die Größe von Vorteil, wenn man mobil sein möchte. Andererseits ist es von Vorteil, Lautsprecher und Verstärker in einem Gerät zu haben, da diese besser aufeinander abgestimmt sind und störende Signale auf dem Signalweg vermieden werden können.

3.3.3 Fazit

Insgesamt ist der Dodekaeder-Lautsprecher für diese Arbeit nicht geeignet, da ein Gitarrenverstärker sowieso keine kugelförmige Richtcharakteristik hat. Möchte man jedoch klassische Instrumente als Referenz verwenden, so ist diese Variante eindeutig besser.

Der Adam A8X ist vor allem wegen seiner Richtcharakteristik für diese Arbeit besser geeignet.

4 Messung der Raumimpulsantworten

In diesem Kapitel werden die Methoden zur Messung der Raumimpulsantworten vorgestellt. Im Anschluss daran wird eine davon umgesetzt. Der Verarbeitung der gemessenen Daten erfolgt dagegen erst im nächsten Kapitel. Eine kurze Einleitung, was ein System ist folgt nun.

4.1 Systeme

Da die System- und Signaltheorie ein sehr weites Feld ist, wird nur auf das Nötigste eingegangen, um die folgenden Vorgänge nachvollziehen zu können. Bei größerem Interesse an diesem Gebiet ist das Lehrbuch **Tontechnik** [Görne (2008)] zu empfehlen. Ein System ist etwas, was Signale überträgt. Sei es ein Mikrofon, Lautsprecher oder jedoch auch ein Raum. Ketten von Systemen bilden wiederum ein neues System. Gibt man ein Eingangssignal in ein System ein, so erhält man ein Ausgangssignal, auch genannt Systemantwort. In dieser Arbeit geht es genau um dieses Prinzip, nur ist das Ziel, die Systemantwort zu simulieren. Wie sich ein Signal in einem Raum verhält, kann man vorhersagen, indem man die Raumimpulsantwort bestimmt. Ähnlich verhält es sich, wenn man räumlich aufgelöste Impulsantworten ermitteln möchte. Dafür ist es nötig, entsprechende Mikrofone so zu platzieren, wie später die Kanalanordnung sein soll. Einfacher und unabhängiger geht es jedoch mit einem Soundfield-Mikrofon, welches wie bereits beschrieben, ein ganzes Klangfeld aufnehmen kann. Erst hinterher bestimmt man, welche Kanalanordnung man haben möchte.

Das Ausgangssignal für einen Kanal erhält man, indem man das Eingangssignal mit der zugehörigen Impulsantwort faltet [Görne (2008)]. Faltung kann man sich bildlich vorstellen, indem man das zweite Signal, in dem Fall die Raumimpulsantwort, spiegelt, gegen das erste Signal, das Eingangssignal, schiebt und die sich überschneidenden Werte miteinander multipliziert und anschließend summiert (siehe Abbildung 4.1).

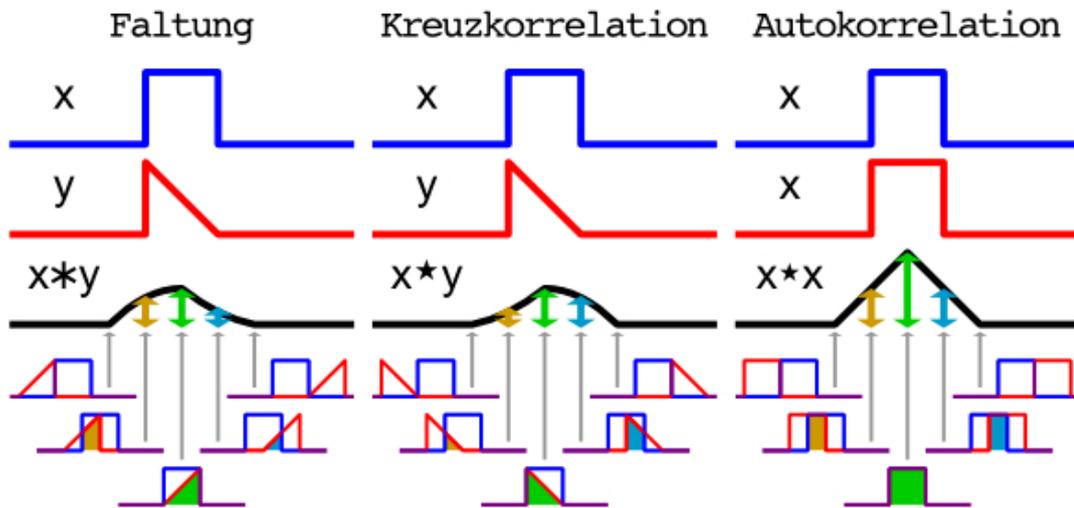


Abbildung 4.1: Faltung, Kreuzkorrelation und Autokorrelation im Vergleich
 Quelle: <http://de.wikipedia.org/wiki/Kreuzkorrelation>

4.2 Methoden zur Messung der Raumimpulsantworten

Es gibt mehrere Methoden, um die Raumimpulsantwort zu bestimmen. Leider ist die beste Methode nur in der Theorie umsetzbar. Ansonsten gibt es eine gängige Methode zur Messung der Raumimpulsantwort, welche sehr zuverlässig funktioniert. Auch eine weitere, etwas ungenauere, Methode wird kurz vorgestellt.

4.2.1 Dirac-Stoß

Theoretisch kann man die Raumimpulsantwort bestimmen [Holters u. a. (2009)], indem man einen Dirac-Stoß dem System zuführt. Es handelt sich um einen Impuls mit unendlich hoher Amplitude, der allerdings in kürzester Zeit erfolgen müsste. Praktisch ist dies jedoch nicht umsetzbar. Näherungen wie zum Beispiel ein Pistolenschuss oder ein platzender Luftballon sind zwar möglich, liefern jedoch kein qualitativ hochwertiges Ergebnis. Oft entstehen dabei auch unerwünschte Verzerrungen, die durch die Nichtlinearität von Bauteilen, in etwa in Mikrofonen, hervorgerufen werden [Görne (2008)]. In der Theorie ist diese Methode sehr wichtig, wie im folgenden Abschnitt gezeigt wird. Faltet man ein Signal mit dem Dirac-Stoß, so erhält man wiederum das Ausgangssignal.

4.2.2 Ermittlung durch deterministisches Rauschen

Eine umsetzbare Methode zur Ermittlung der Raumimpulsantwort ist das Messen mit Hilfe von deterministischem Rauschen.

Ähnlich wie die Faltung funktioniert die Kreuzkorrelation zweier Signale. Der einzige Unterschied besteht darin, dass das zweite Signal nicht gespiegelt wird (siehe Abbildung 4.1). Das Ergebnis zeigt, wie ähnlich die beiden Signale zueinander sind. Führt man die Kreuzkorrelation auf zwei identischen Signalen durch, so spricht man von der Autokorrelationsfunktion. Nimmt man nun ein Zufallssignal, näherungsweise gilt das auch für das weiße Rauschen, so erhält man bei der Autokorrelation einen Dirac-Stoß (siehe Abbildung 4.2). Steigern kann man die Qualität, indem man immer länger werdende Messsignale verwendet, da je mehr Samples miteinander korreliert werden, desto besser gleichen sie sich aus.

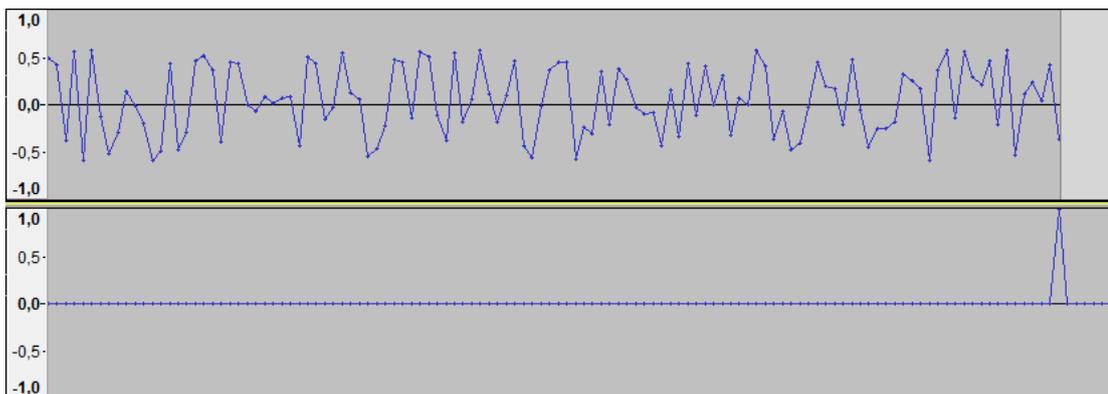


Abbildung 4.2: Ausschnitt weißes Rauschens (oben) und der daraus durch Autokorrelation erzeugte Dirac-Stoß (unten) im zeitlichen Verlauf

Der Ansatz lässt uns nun sehr einfach die Raumimpulsantwort bestimmen, indem man die Kreuzkorrelation von Eingangssignal und Ausgangssignal berechnet.

Sei $m(t)$ das Messsignal zum Zeitpunkt t , in dem Fall somit das Rauschen, $a(t)$ die Systemantwort, $\delta(t)$ der Dirac-Stoß und $i(t)$ die Impulsantwort, so kann die Systemantwort dargestellt werden als [Zerva und Petropulu (2000)]:

$$a(t) = m(t) * i(t). \quad (4.1)$$

Führt man nun die Kreuzkorrelation der Systemantwort mit dem Messsignal durch, so folgt aus Formel 4.1:

$$[m(t) * i(t)] * m(-t) = [m(t) * m(-t)] * i(t) = \delta(t) * i(t) = i(t). \quad (4.2)$$

Wie bereits erklärt, erhält man einen Dirac-Stoß, sobald man auf das weiße Rauschen, welches bei uns das Messsignal ist, die Autokorrelation durchführt. In der Systemantwort ist bereits die Impulsantwort sowie das Messsignal enthalten. Um deutlich zu machen, dass ein Dirac-Stoß entsteht, wird das Messsignal mit der Impulsantwort vertauscht. Wie ebenfalls schon erwähnt, ergibt die Faltung von Impulsantwort mit dem Dirac-Stoß die Antwort oder eine verschobene Form dieser.

Mehr braucht es nicht, um die Impulsantwort zu berechnen. Jedoch ist es ratsam, bei langen Aufnahmen und aus Performance-Gründen mit Hilfe der schnellen Fourier-Transformation die Impulsantwort zu ermitteln. Näheres dazu findet man unter 5.4.

4.2.3 Ermittlung durch den Sinus-Sweep

Anstelle von weißem Rauschen eignet sich auch der Sinus-Sweep zur Berechnung einer Impulsantwort [Holters u. a. (2009)]. Der Sinus-Sweep ist ein Signal, welches einen bestimmten Frequenzbereich abdeckt, vorzugsweise etwa 16 - 20000 Hz. Es handelt sich um ein Sinus-Signal mit konstanter Amplitude und exponentiell steigender Frequenz. Dank diesem Signal wird jede gewünschte Frequenz im Raum angeregt. Durch die Kreuzkorrelation des Messsignals und der Systemantwort erhält man ebenfalls eine Raumimpulsantwort.

Die Autokorrelation eines Sinus-Sweeps ergibt ein Signal, welches dem der Autokorrelation von rosa Rauschen ähnelt (siehe Abbildung 4.3).

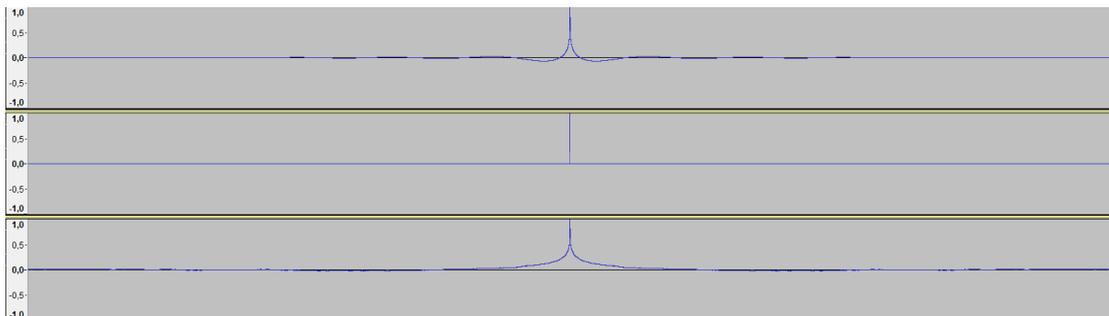


Abbildung 4.3: Autokorrelation des Sinus-Sweeps (oben), des weißen Rauschens (Mitte) und des rosa Rauschens (unten) im Vergleich

4.3 Die Durchführung

Die Durchführung findet im Raum 1060 statt. Er ist ca. 20,35 Meter lang und 9,60 Meter breit. Der Adam A8X Lautsprecher befindet sich mit Abstand von etwa 2,60 Metern von der Tafel mittig zu den Seitenwänden und wird über das Saffire 56 Interface vom Laptop aus angesteuert. Das Soundfield-Mikrofon, welches ebenfalls an dem Interface angeschlossen ist, steht in der Mitte des Raumes etwa 8,35 Meter von der hinteren Wand entfernt.

Das über Audacity generierte, 20 Sekunden lange, weiße Rauschen wird über Nuendo abgespielt und gleichzeitig aufgenommen, welches als Referenz dienen soll, nachdem die Verzögerung zwischen dem Abspielen und Aufnehmen beseitigt wurde. Nuendo bietet an, die Aufnahmeverzögerung zu unterbinden. Geprüft wird dies, indem ein Mikrofon so nah wie nur möglich vor den Lautsprecher gestellt wird. Anschließend wird ein Signal gleichzeitig wiedergegeben und aufgenommen. Die Verzögerung betrug im Schnitt neun Samples bei einem Abstand von etwa drei Zentimetern zwischen Lautsprecher und Mikrofon. Auf diese Entfernung dürften jedoch nur vier Samples Verzögerung entstehen. Die Differenz beträgt folglich fünf Samples. Diese kann ausgeglichen werden, indem man das gemessene Signal fünf Samples vorzieht. Auf die Verzögerung von vier Samples kommt man, indem man berechnet, wie lange Schall benötigt, um eine Strecke von drei Zentimetern zurückzulegen. Bei einer Schallgeschwindigkeit von etwa 343 Metern pro Sekunde ergibt sich für drei Zentimeter eine Laufzeit von $8,746 \cdot 10^{-5}$ Sekunden. Bei 44100 Samples pro Sekunde ergibt sich für diesen Wert eine Verzögerung von $8,746 \cdot 10^{-5}$ Sekunden \cdot 44100 Hz \approx 4 Samples.

Die aufgezeichneten Daten werden im folgenden Kapitel weiterverarbeitet, da ansonsten bei der Durchführung nichts weiter zu beachten ist.

5 Verwendung der gemessenen Daten

In diesem Kapitel werden die aufgezeichneten Daten verarbeitet. Es folgt eine Beschreibung, wie die vier Rohsignale des Soundfield-Mikrofons schließlich zur Raumimpulsantwort werden. Anschließend wird auch erklärt, wie eine Quelle mit Hilfe von MATLAB gefaltet wird. Zum Schluss wird ein mögliches Modell eines VST-Plugin beschrieben, welches in Echtzeit einen Raumklang für eine Quelle berechnen soll.

5.1 Umwandlung ins B-Format

Da die Raumimpulsantworten nur ein einziges Mal berechnet werden müssen, kann die Vorarbeit auch etwas umständlicher, dafür jedoch viel genauer geleistet werden. Die Umrechnung vom A-Format ins B-Format wird nicht in MATLAB realisiert, sondern geschieht direkt in Nuendo.

Das SPS200-Surround-Zone2 Plugin bietet nicht nur die Möglichkeit, vom A-Format direkt in das gewünschte Ausgangsformat umzurechnen, sondern kann auch nur das B-Format ausgeben. Die erste Version des Plugins liefert sogar eine Umrechnung in das 8.0-Format, welches für diese Arbeit benötigt wird, ist jedoch mit der aktuellen Version von Nuendo nicht mehr kompatibel.

Nuendo bietet die Möglichkeit eine Audiospur voranzuberechnen und diese anschließend zu speichern.

5.2 Berechnung der Raumimpulsantworten für die einzelnen Kanäle

Wurde die Datei, welche das B-Format beinhaltet, erfolgreich erzeugt, so werden aus diesen Signalen zunächst die Signale für die entsprechenden Kanäle berechnet. Im Anschluss daran kann für jeden einzelnen Kanal die Raumimpulsantwort berechnet werden.

5.2.1 Umwandlung von B-Format auf einzelne Kanäle

Für die Umwandlung des B-Formats auf einzelne Kanäle existiert eine einfache Formel. Jedoch funktioniert diese nur, wenn es sich um Kanäle auf einem Kreis in der Horizontalen handelt. Vorzugsweise wählt man Figuren der regelmäßigen Polygone.

Sei d wählbar zwischen 0 und 2, wobei 0 die Richtcharakteristik einer Kugel, 1 die einer Niere und 2 die einer Acht darstellt und α der Winkel zur Mittelachse, ergibt folgende Gleichung das Signal für je einen Kanal [Wiggins (2008)]:

$$S = 0.5((2 - d)\sqrt{2}W + d(\cos(\alpha)X + \sin(\alpha)Y)) \quad (5.1)$$

Vereinfacht in Worten erklärt, werden zu dem Gesamtschalldruck Signale addiert. Möchte man nun einen Kanal berechnen, welcher vor dem Zuhörer steht, so ergibt sich ein Winkel von Null. Für den Kosinus erhält man den Wert eins, für den Sinus null. Multipliziert man diese Werte nun mit dem zugehörigen Signal, so wird schnell klar, dass die Signale, welche von den Seiten, demnach der Y-Achse kommen, komplett wegfallen (siehe Abbildung 5.1).

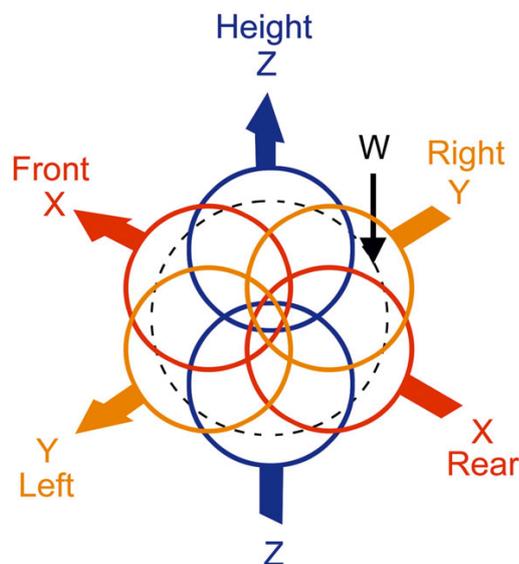


Abbildung 5.1: Das B-Format mit Hilfe von Richtcharakteristiken als einzelne Figur dargestellt

Quelle: <http://www.surround-library.com/white-papers>

Berechnet man nun das Signal für einen gewünschten Kanal, gleicht dieses einem Signal, welches mit einem Mikrofon aufgenommen wurde, das den gleichen Winkel und die gleiche Richtcharakteristik hat. Somit können beliebig viele virtuelle Mikrofone erzeugt werden, wobei zu beachten ist, dass zu viele Kanäle keinen Vorteil bringen, sofern man ein Soundfield-Mikrofon erster Ordnung verwendet, da nah aneinanderliegende Kanäle sehr ähnlich klingen würden.

5.2.2 Herausrechnen der Impulsantwort

Hat man das Signal für einen bestimmten Kanal berechnet, so kann nun die Impulsantwort für die zugehörige Richtung bestimmt werden (siehe Abbildung 5.2). Dies geschieht, wie bereits in 4.2.2 beschrieben, mit Hilfe der Kreuzkorrelation von Messsignal und dem Signal des Kanals. Anschließend wird die Raumimpulsantwort an der Stelle, in der sie im Grundrauschen des Systems untergeht, abgeschnitten [Görne (2008)].

Dieser Vorgang wird für jeden einzelnen Kanal wiederholt, sodass man am Ende für ein Acht-Kanal-Tonsystem auch acht einzelne Impulsantworten hat.

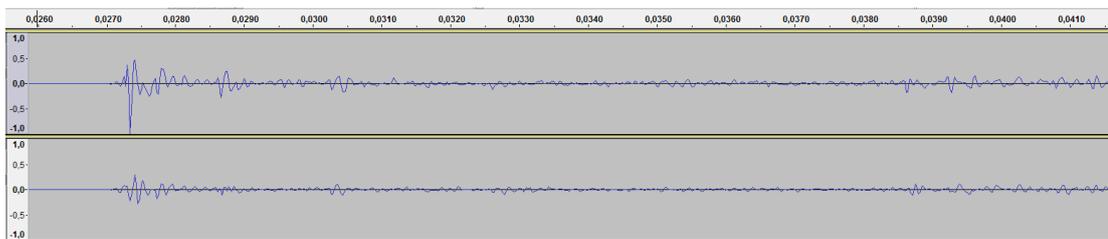


Abbildung 5.2: Ausschnitt der Raumimpulsantworten des vordersten (oben) und hintersten Kanals (unten) des Acht-Kanal-Tonsystems im Raum 1060

5.3 Überprüfung im schallreflexionsarmen Raum

Um die Richtigkeit dieses Verfahrens zu bestätigen, wird es im schallreflexionsarmen Raum der HAW-Hamburg anhand zweier Beispiele überprüft. Zum einen werden Impulsantworten für das A-Format des Soundfield-Mikrofons berechnet und zum anderen für acht virtuelle Mikrofone, welche aus dem B-Format berechnet werden. Erwartet wird, da im Raum so gut wie kein Schall reflektiert wird, dass im A-Format möglichst nur die vorderen beiden Mikrofonkapseln einen einzigen Impuls als Impulsantwort haben. In dem acht Kanal Setup wird erwartet, dass möglichst nur für die vorderen Kanäle ein Impuls als Antwort berechnet wird, für die hinteren Kanäle jedoch ein sehr

schwacher Impuls (siehe Abbildung 5.3).



Abbildung 5.3: Die Raumimpulsantworten des vordersten (oben) und hintersten Kanals (unten) des Acht-Kanal-Tonsystems im schallreflexionsarmen Raum

Die Überprüfung zeigt hingegen, dass das Ergebnis zwar nicht weit von den Erwartungen liegt, jedoch nicht ganz mit ihnen übereinstimmt. Gründe könnten sein, dass der Raum nicht komplett gedämpft ist, da sich an der Tür Rohre und in der Mitte des Raumes eine Vorrichtung zur Anbringung eines Lautsprechers montiert ist. Ein weiterer Grund ist mutmaßlich der Laufzeitunterschied vom Lautsprecher zur Mikrofonkapsel, da es sich um ein Gehäuse mit Bassreflexöffnungen und zudem einen 2-Wege Lautsprecher handelt. Ein geschlossenes Gehäuse mit Breitbandchassis könnte hier durchaus von Vorteil sein. Als letzter Grund ist zu nennen, dass in dem schallreflexionsarmen Raum mit rosa statt weißem Rauschen gearbeitet wurde, um die Hochton-Chassis der Lautsprecher zu schonen. Verwendet man rosa statt weißes Rauschen, so erhält man bei der Autokorrelation keinen Dirac-Stoß, sondern ein nur ähnliches Gebilde. Allerdings lässt sich durch einen Hochpassfilter mit 3 dB pro Oktave, welcher sowohl das Messsignal als auch die Signale der virtuellen Mikrofone filtert, ein sehr viel besseres Ergebnis erzielen, welches näher an den Erwartungen liegt (siehe Abbildung 5.4). Dies liegt daran, dass man durch Filterung der tiefen Frequenzen aus dem rosa ein annähernd weißes Rauschen erzeugt.

Insgesamt ist die Überprüfung überzeugend und zeigt, dass das Verfahren korrekt ist.

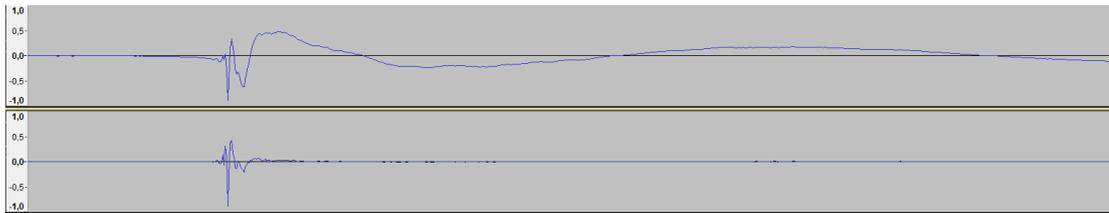


Abbildung 5.4: Die mit Hochpass gefilterte, berechnete Raumimpulsantwort (unten) im Vergleich zur nicht gefilterten, berechneten Raumimpulsantwort (oben) des selben virtuellen Mikrofons im schallreflexionsarmen Raum

5.4 Die Faltung einer Quelle mit der Impulsantwort

Nachdem für die gewünschten Kanäle eine Impulsantwort berechnet wurde, kann man eine Quelle mit ihr falten. Eine bildliche Beschreibung der Faltung ist unter 4.1 zu finden. Mathematisch gesehen wird pro Sample ein Integral der sich überschneidenden Fläche berechnet. Ein 60 Sekunden langes Stück und einem Kanal mit einer Impulsantwort von einer Sekunde benötigt bei 44100 Hertz Abtastfrequenz folglich 2.690.099 Berechnungen eines Integrals. Rechnet man das auf acht Kanäle hoch, so ist man bereits bei 21.520.792 dieser Berechnungen.

Wie man sieht, bekommt man zwar als Ergebnis das gefaltete Signal und somit das fertige Audiosignal für den jeweiligen Kanal, jedoch benötigt diese Methode sehr viel Rechenleistung und ist somit für ein Echtzeitsystem nicht geeignet.

Um aus einem Signal und einer Impulsantwort trotzdem eine Systemantwort in Echtzeit zu berechnen, wird der Zeitbereich des Signals und der Impulsantwort in den Frequenzbereich, mit Hilfe der schnellen Fourier Transformation, umgewandelt (siehe Abbildung 5.5). Anschließend werden die beiden Frequenzbereiche miteinander multipliziert und mit Hilfe der inversen Fourier Transformation zurück in den Zeitbereich umgewandelt [Girod u. a. (2007)].

Der Nachteil hierbei ist jedoch, dass man bei einer Echtzeitberechnung eine Latenz hat, die abhängig von der Fensterlänge des Signals ist. Wählt man bei einer Abtastrate von 44100 Hertz eine Fensterlänge von 1024 Samples, so errechnet sich eine Verzögerung von 23 ms.

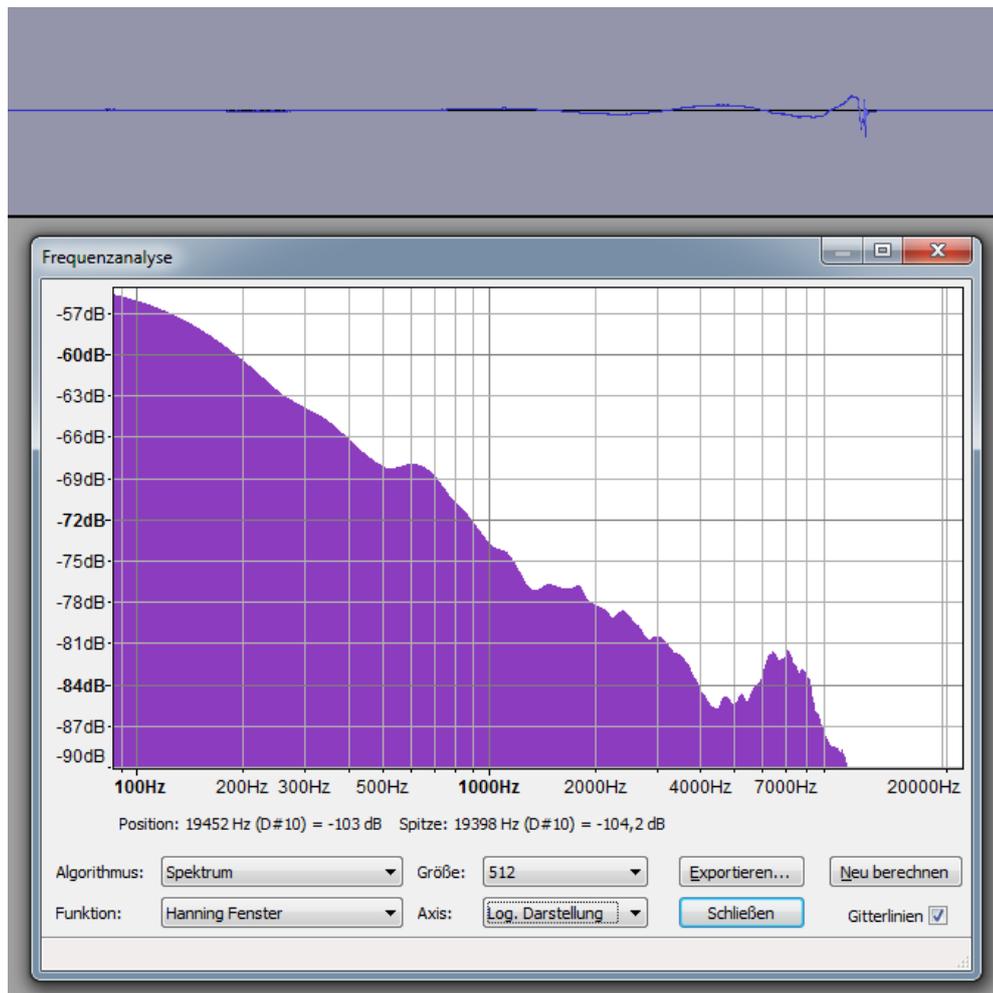


Abbildung 5.5: Ein Signal im Zeitbereich und dessen Frequenzbereich

5.5 Erstellen des VST-Plugins zur räumlichen Wiedergabe einer Quelle

Um das ganze Verfahren auch an der Wellenfeldsynthese-Anlage betreiben zu können, wird in diesem Abschnitt ein Modell eines VST-Plugins beschrieben. Dabei werden die einzelnen Impulsantworten mit MATLAB vorberechnet und deshalb als bereits gegeben betrachtet. Diese sollen im Plugin geladen werden und auf ein vielfaches der Fensterlänge gebracht werden, indem so viele Samples wie nötig mit dem Wert Null angehängt werden. Auch das Eingangssignal muss auf ein vielfaches der Fensterlänge und zusätzlich um die Länge der bereits verlängerten Impulsantwort erweitert werden. Dies ist nötig, damit

einerseits die schnelle Fourier Transformation funktioniert, damit das Frequenzspektrum gleich aufgelöst ist, andererseits muss für das letzte Sample die gesamte Impulsantwort berechnet werden.

Nachdem nun die korrekte Länge der Signale vorhanden ist, kann die schnelle Fourier Transformation fensterweise auf das Eingangssignal angewendet werden. Für die Impulsantwort kann die Transformation bereits vorher durchgeführt werden, da diese von Anfang an bekannt ist. Nun werden beide Frequenzspektren miteinander multipliziert und mit Hilfe der inversen schnellen Fourier Transformation zurück in den Zeitbereich gewandelt. Das hier gewonnene Signal ist das gesuchte Ausgangssignal.

Zu beachten ist jedoch, dass eine Fensterlänge eines Eingangssignals zwischengespeichert und auf das zugehörige Signal aufsummiert werden muss und zwar genau so lange, bis es alle Fenster der Impulsantwort durchlaufen ist.

Das Verfahren wurde mit MATLAB simuliert und man erhält tatsächlich ein gefaltetes, räumliches Signal. Jedoch muss zugegeben werden, dass der Algorithmus in MATLAB noch nicht ganz ausgereift ist, da sich an einer, noch unbekanntem, Stelle offenbar ein Fehler eingeschlichen hat, welcher das Signal verzerrt erklingen lässt.

6 Vergleich zwischen Original und Simulation

Nachdem die Gitarrenaufnahme simuliert wurde, wird in diesem Kapitel ermittelt, wie sehr sie sich von dem Original unterscheidet. Dabei wird zuerst die Aufnahme des Originals erläutert und anschließend miteinander verglichen.

6.1 Aufbau der originalen Aufnahme

Der Aufbau des Originals unterscheidet sich nicht von dem des Aufbaus zur Ermittlung der Raumimpulsantwort. Der Lautsprecher muss an genau der selben Stelle stehen, ebenso wie das Soundfield-Mikrofon. Statt des Rauschens wird nun die Gitarrenaufnahme abgespielt und gleichzeitig aufgenommen.

6.2 Vergleich beider Aufnahmen

Wurde das Original aufgenommen und über das B-Format in das Acht-Kanal-Tonsystem überführt, so kann es mit dem gefalteten Signal verglichen werden.

Die beiden Aufnahmen müssen zuerst auf den gleichen Pegel gebracht werden, um Verfälschungen des Ergebnisses zu vermeiden. Es gibt drei Verfahren, um die Aufnahmen zu vergleichen. Als erstes werden jeweils die einzelnen Kanäle beider Aufnahmen per Sichtkontrolle verglichen. Als zweites werden sie als A/B-Wiedergabe angehört und als drittes wird die Korrelationsfunktion aufgerufen, da die Ausgabe die Ähnlichkeit zweier Signale liefert.

6.3 Auswertung

Der Vergleich erweist sich als relativ schwierig, da man zwar Unterschiede im Zeitverlauf sieht, jedoch keinen großen Unterschied hören kann. Die Kreuzkorrelation beider Signale

6 Vergleich zwischen Original und Simulation

liefert dagegen ein recht gutes Ergebnis, da im Zeitpunkt $t=0$ ein Peak entsteht (siehe Abbildung 6.1), davor und danach verläuft das Signal jedoch relativ glatt.



Abbildung 6.1: Das originale Signal (oben), das gefaltete Signal (unten), sowie die Kreuzkorrelation beider Signale im Vergleich

Als Vergleich werden zwei zufällig ausgewählte Musikstücke mit der selben Methode miteinander verglichen. Das Ergebnis zeigt hier dagegen keine eindeutige Ähnlichkeit zwischen den beiden Musikstücken (siehe Abbildung 6.2).

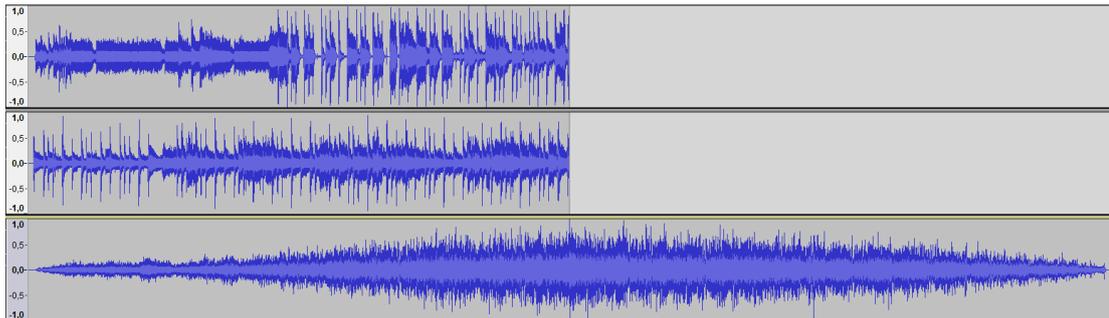


Abbildung 6.2: Zwei zufällige Musikstücke und deren Kreuzkorrelation

Als Alternative wurde mit den mit Hilfe des Sinus-Sweeps gemessenen Impulsantworten das trockene Signal gefaltet. Hierbei waren jedoch deutliche Unterschiede zum Original zu hören, da Frequenzen ab etwa 7000 Hz kaum vorhanden waren. Dies könnte an dem typischen Frequenzgang von Lautsprechern liegen, deren Schalldruck ebenfalls ab etwa 6000 - 7000 Hz kontinuierlich abnimmt. Würde man nun einen Equalizer benutzen, um die schlecht hörbaren Frequenzen anzugleichen, würde man möglicherweise ein besseres Ergebnis erzielen. Die Faltung mit der Impulsantwort, welche mit Hilfe von weißem Rauschen gemessen wurde, ist demgegenüber unkomplizierter, da sie keine Anpassung benötigt und von vornherein ein gutes Ergebnis mit sich bringt.

7 Schluss

Dieses Kapitel fasst den praktischen Teil zusammen und vermittelt einen Ausblick auf Verbesserungen und weitere Anwendungsmöglichkeiten.

7.1 Fazit

Nachdem alle Grundlagen, welche nötig sind, um das Verhalten eines Signals in einem System vorherzusagen, angeeignet wurden, konnte der praktische Teil realisiert werden. Der Aufbau erwies sich nach mehreren Versuchen als sehr unkompliziert, ebenso wie die Durchführung des praktischen Teils. Dies ist jedoch nur dank des Soundfield-Mikrofons und des zugehörigen VST-Plugins möglich, da durch Software ein Großteil der Arbeit abgenommen wird. Im Prinzip benötigt man bei sehr guten Bedingungen nur eine einzige Messung. Die restlichen Schritte lassen sich bequem am Rechner durchführen. Außerdem ist man mit dem Mikrofon sehr flexibel, da man sich nicht auf ein bestimmtes Format beschränken zu braucht.

Das Ergebnis ist auf jeden Fall zufriedenstellend und zeigt, dass die Theorie auch in der Praxis tatsächlich funktioniert. Natürlich existieren kleinere Abweichungen, jedoch lassen sich deren Ursachen kaum vermeiden. Insgesamt kann behauptet werden, dass das Ziel erreicht wurde.

7.2 Ausblick

Die Ergebnisse sind zwar zufriedenstellend, jedoch könnte die Qualität gesteigert werden, indem man ein Soundfield-Mikrofon höherer Ordnung verwenden würde, da virtuelle Mikrofone genauer gerichtet werden könnten. Der Unterschied zwischen nah aneinanderliegenden Kanälen wäre größer. Auch ein geschlossener Lautsprecher mit Breitbandchassis und sehr linearem Frequenzgang könnte das Ergebnis noch genauer werden lassen.

Das vorgestellte Modell eines VST-Plugins könnte realisiert werden, um auch in Echtzeit Raumakustik für eine Quelle oder die Mikrofonwiedergabe zu erzeugen.

Abbildungsverzeichnis

| | | |
|-----|--|----|
| 2.1 | Das B-Format erster Ordnung | 4 |
| 2.2 | Netzwerkstruktur einer WFS-Anlage | 6 |
| 2.3 | Ausbreitung der synthetisierten Wellenfront einer virtuellen Schallquelle | 8 |
| 2.4 | Darstellung der Raumakustik der Wellenfeldsynthese (gelbes Rechteck) in Xwonder | 9 |
| 3.1 | Das TSL Products SoundField SPS200 Mikrofon | 12 |
| 4.1 | Faltung, Kreuzkorrelation und Autokorrelation im Vergleich | 16 |
| 4.2 | Ausschnitt weißes Rauschens (oben) und der daraus durch Autokorrelation erzeugte Dirac-Stoß (unten) im zeitlichen Verlauf | 17 |
| 4.3 | Autokorrelation des Sinus-Sweeps (oben), des weißen Rauschens (Mitte) und des rosa Rauschens (unten) im Vergleich | 18 |
| 5.1 | Das B-Format mit Hilfe von Richtcharakteristiken als einzelne Figur dargestellt | 21 |
| 5.2 | Ausschnitt der Raumimpulsantworten des vordersten (oben) und hintersten Kanals (unten) des Acht-Kanal-Tonsystems im Raum 1060 | 22 |
| 5.3 | Die Raumimpulsantworten des vordersten (oben) und hintersten Kanals (unten) des Acht-Kanal-Tonsystems im schallreflexionsarmen Raum | 23 |
| 5.4 | Die mit Hochpass gefilterte, berechnete Raumimpulsantwort (unten) im Vergleich zur nicht gefilterten, berechneten Raumimpulsantwort (oben) des selben virtuellen Mikrofons im schallreflexionsarmen Raum | 24 |
| 5.5 | Ein Signal im Zeitbereich und dessen Frequenzbereich | 25 |
| 6.1 | Das originale Signal (oben), das gefaltete Signal (unten), sowie die Kreuz- korrelation beider Signale im Vergleich | 28 |
| 6.2 | Zwei zufällige Musikstücke und deren Kreuzkorrelation | 28 |

Literaturverzeichnis

- [BlueRippleSound 2014] : *TOA DECODING VST*. 2014. – URL <http://www.blueripplesound.com/products/toa-decoding-vst>
- [Christoffel 2014] CHRISTOFFEL, Carola: *Modifikation der Software einer Wellenfeldsyntheseanlage zur Wiedergabe fokussierter Quellen in Abhängigkeit der Zuhörerposition*, Diplomarbeit, 2014
- [Corteel und Caulkins] CORTEEL, Etienne ; CAULKINS, Terence: *Sound Scene Creation and Manipulation using Wave Field Synthesis*.
- [Fohl 2013] FOHL, Wolfgang: *The Wave Field Synthesis System at the HAW Hamburg*. In: BADER, Rolf (Hrsg.): *Sound - Perception - Performance* Bd. 1. Springer, 2013
- [Girod u. a. 2007] GIROD, Bernd ; RABENSTEIN, Rudolf ; STEGNER, Alexander: *Einführung in die Systemtheorie*. B.G. Teubner Verlag / GWV Fachverlage GmbH, Wiesbaden 2007, 2007
- [Görne 2008] GÖRNE, Thomas ; SCHMIDT, Prof. Dr. U. (Hrsg.): *Tontechnik*. Carl Hanser Verlag München, 2008
- [Heller u. a. 2012] HELLER, Aaron J. ; BENJAMIN, Eric M. ; LEE, Richard: *A Toolkit for the Design of Ambisonic Decoders*, 2012
- [Holters u. a. 2009] HOLTERS, Martin ; CORBACH, Tobias ; ZÖLZER, Udo: *Impulse Response Measurement Techniques and their Applicability in the Real World*, 2009
- [Makarski u. a. 2012] MAKARSKI, Michael ; GOERTZ, Anselm ; THADEN, Rainer ; KLEBER, Jochen: *Entwicklung einer WFS-Anlage mit FIR-Entzerrung und Dante Audionetzwerk*, 2012
- [Nogalski 2012] NOGALSKI, Malte: *Gestengesteuerte Positionierung von Klangquellen einer Wellenfeldsynthese-Anlage mit Hilfe eines kamerabasierten 3D-Tracking-Systems*, Diplomarbeit, 2012

- [Trevino u. a. 2010] TREVINO, Jorge ; OKAMOTO, Takuma ; IWAYA, Yukio ; SUZUKI, Yôiti: *High order Ambisonic decoding method for irregular loudspeaker arrays*. 2010
- [Wiggins 2008] WIGGINS, Dr. B.: Has Ambisonics Come of Age? 30 (2008)
- [Zerva und Petropulu 2000] ZERVA, A. ; PETROPULU, A.P.: Blind System Identification, 2000

MATLAB Code

1 Berechnung des Rauschens und der zugehörigen Impulsantwort für acht Kanäle

```
1 %Einlesen des Messsignals und des W, X und Y Kanals des B-Formats
2 r = wavread('20secrauschenweiss.wav');
3 w = wavread('bformatW.wav');
4 x = wavread('bformatX.wav');
5 y = wavread('bformatY.wav');
6
7 d = 1; %S=0.5*((2-d)\sqrt{2}W+d(\cos(\alpha)X+\sin(\alpha)Y))
8
9 cha = zeros(1, 8);
10
11 for i = 0:1:7
12     cha(i+1) = i*45 %Winkel fuer den jeweiligen Kanal wird gesetzt
13 end
14
15 %Berechnung des jeweiligen Signals fuer ein virtuelles Mikrofon
16 ch1 = 0.5*((2-d)*sqrt(2)*w + d*(x*cos(cha(1)) + y*sin(cha(1))));
17 ch2 = 0.5*((2-d)*sqrt(2)*w + d*(x*cos(cha(2)) + y*sin(cha(2))));
18 ch3 = 0.5*((2-d)*sqrt(2)*w + d*(x*cos(cha(3)) + y*sin(cha(3))));
19 ch4 = 0.5*((2-d)*sqrt(2)*w + d*(x*cos(cha(4)) + y*sin(cha(4))));
20 ch5 = 0.5*((2-d)*sqrt(2)*w + d*(x*cos(cha(5)) + y*sin(cha(5))));
21 ch6 = 0.5*((2-d)*sqrt(2)*w + d*(x*cos(cha(6)) + y*sin(cha(6))));
22 ch7 = 0.5*((2-d)*sqrt(2)*w + d*(x*cos(cha(7)) + y*sin(cha(7))));
23 ch8 = 0.5*((2-d)*sqrt(2)*w + d*(x*cos(cha(8)) + y*sin(cha(8))));
24
25 channels = [ch1, ch2, ch3, ch4, ch5, ch6, ch7, ch8];
26
27 wavwrite(channels, 44100, 'signaledervirtuellenmikrofone.wav');
```

```
28
29 %Kreuzkorrelation zwischen dem berechneten Signal und dem
30 %Messignal
31 cor1 = xcorr(ch1, r);
32 cor2 = xcorr(ch2, r);
33 cor3 = xcorr(ch3, r);
34 cor4 = xcorr(ch4, r);
35 cor5 = xcorr(ch5, r);
36 cor6 = xcorr(ch6, r);
37 cor7 = xcorr(ch7, r);
38 cor8 = xcorr(ch8, r);
39
40 impulsantworten = [cor1, cor2, cor3, cor4,
41 cor5, cor6, cor7, cor8];
42
43 impulsantwortenohneclipping = impulsantworten/20000;
44
45 wavwrite(impulsantwortenohneclipping, 44100,
46 'impulsantwortendervirtuellenmikrofone.wav');
```

2 Faltung eines Signals mit den Impulsantworten

```
1 %Einlesen des Signals welches gefaltet werden soll und der
2 %Impulsantworten der jeweiligen Kanäle
3 sig = wavread('signal.wav');
4 ria1 = wavread('raumimpulsantwort1.wav');
5 ria2 = wavread('raumimpulsantwort2.wav');
6 ria3 = wavread('raumimpulsantwort3.wav');
7 ria4 = wavread('raumimpulsantwort4.wav');
8 ria5 = wavread('raumimpulsantwort5.wav');
9 ria6 = wavread('raumimpulsantwort6.wav');
10 ria7 = wavread('raumimpulsantwort7.wav');
11 ria8 = wavread('raumimpulsantwort8.wav');
12
13 %Faltung des Signals mit den Impulsantworten und Vermeidung von
14 %Clipping
15 ch1 = conv(sig, ria1)/2;
16 ch2 = conv(sig, ria2)/2;
17 ch3 = conv(sig, ria3)/2;
```

MATLAB Code

```
18 ch4 = conv(sig, ria4)/2;
19 ch5 = conv(sig, ria5)/2;
20 ch6 = conv(sig, ria6)/2;
21 ch7 = conv(sig, ria7)/2;
22 ch8 = conv(sig, ria8)/2;
23
24 convchannels = [ch1, ch2, ch3, ch4, ch5, ch6, ch7, ch8];
25
26 wavwrite(convchannels, 44100,
27 'gefaltetessignal.wav');
```

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

Hamburg, 16. Oktober 2014 Alexander Witek