

Bachelorarbeit

Hochschule der Medien Stuttgart

Fakultät Electronic Media

Studiengang Audiovisuelle Medien

Binauralsynthese von Mehrspurproduktionen auf mobilen Endgeräten

Untersuchung und Definition der Anforderungen

vorgelegt von: Dominik Litfin (Matrikel-Nr.: 23117)
am: 28.02.2014
Erstprüfer: Prof. Oliver Curdt, HdM
Zweitprüfer: Jeffrey Strößner, Lawo AG



Erklärung

Hiermit versichere ich, Dominik Litfin, an Eides Statt, dass ich die vorliegende Bachelorarbeit mit dem Titel: „**Untersuchung und Definition der Anforderungen zur Binauralsynthese von Mehrspurproduktionen auf mobilen Endgeräten**“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden. Ich habe die Bedeutung der eidesstattlichen Versicherung und die prüfungsrechtlichen Folgen (§26 Abs. 2 Bachelor-SPO der HdM) sowie die strafrechtlichen Folgen (gem. § 156 StGB) einer unrichtigen oder unvollständigen eidesstattlichen Versicherung zur Kenntnis genommen.

Stuttgart, den 28.02.2014

Ort, Datum

Unterschrift

Kurzfassung

Die vorliegende Bachelorarbeit richtet sich an alle Leser mit fundiertem Wissen im Bereich der Audiotechnik, setzt jedoch keine Vertrautheit mit der exakten Thematik voraus.

Sie beschäftigt sich in Theorie und Praxis mit der Synthese von Mehrspurproduktionen und deren Wiedergabe über Kopfhörer auf einem Mobilgerät. Ziel ist es, mit Hilfe der Binauraltechnik eine 5.1 Surroundproduktion unter Erhalt aller räumlichen und klanglichen Eigenschaften in der Digitalen Audio Workstation für die zweikanalige Wiedergabe aufzubereiten. Anschließend wird aus den gewonnenen Erkenntnissen abgeleitet, ob und wie eine solche Synthese auch auf einem Mobilgerät durchgeführt werden kann.

Zu diesem Zweck wurde mit dem Realiser A8 der Firma Smyth Research im Tonstudio der Hochschule der Medien eine Messung von Außenohrübertragungsfunktionen durchgeführt. Die daraus errechneten Impulsantworten wurden anschließend durch Faltung mit den einzelnen Signalen der Mehrspurmischung kombiniert, um das Originalschallfeld zu reproduzieren.

Das letzte Kapitel behandelt die Konzeption einer Applikation, welche die Faltung auf einem Smartphone durchführen, und so die ursprüngliche Abhörumgebung synthetisieren kann. Zusätzlich wird auf die Problematik der Musikwiedergabe via Kopfhörer eingegangen und eine Abgrenzung von Kopfhörer- und Lautsprecherstereophonie vorgenommen.

Abstract

This bachelor thesis addresses readers with in-depth knowledge in the field of audio technology, but does not require further acquirement of the exact issue this paper examines.

It theoretically and practically deals with the synthesis of multitrack-productions and their playback via headphones on a mobile device. The aim is to edit a 5.1 Surroundsound production in a Digital Audio Workstation for two-channel playback without affecting neither spatial nor tonal characteristics of the signals. It is then derived from the results obtained as to whether and how such a synthesis can be carried out on a mobile device.

For this purpose, a measurement of the Head Related Transfer Function was made in the studio of Stuttgart Media University using Smyth Research's Realiser A8. The impulse responses received were used afterwards for convolution with each channel of a multitrack-production in order to reproduce the original sound field.

The last chapter deals with the design of an application, which is able to perform the convolution on a smartphone and therefore can synthesize the original listening environment. In addition, issues of music playback via headphones are discussed and a differentiation between speaker- and binaural recordings will be conducted.

Inhaltsverzeichnis

Erklärung.....	3
Kurzfassung.....	4
Abstract.....	5
Abbildungsverzeichnis.....	8
Tabellenverzeichnis.....	9
Abkürzungsverzeichnis.....	10
1. Einleitung.....	12
2. Grundlagen zum menschlichen Gehör.....	13
2.1 Aufbau des Hörorgans.....	13
2.2 Räumliches Hören.....	14
2.2.1 Lokalisation.....	15
2.2.1.1 Richtungshören in der Horizontalebene.....	16
2.2.1.2 Richtungshören in der Medianebene.....	18
2.2.2 Entfernungswahrnehmung	20
3. Mobiles Wachstum und seine Auswirkungen auf den Musikkonsum.....	23
3.1 Technischer Fortschritt.....	23
3.2 Marktentwicklung.....	24
3.3 Auswirkungen.....	26
3.4 Zusammenfassung.....	27
4. Mehrkanalstereophonie.....	29
4.1 Zweikanalstereophonie.....	29
4.2 Surround.....	31
4.3 Auro 3D.....	33
4.4 Bedeutung für die Musikindustrie.....	34
5. Der Kopfhörer als Schallwandler.....	36
5.1 Abgrenzung von Kopfhörer- und Lautsprecherstereophonie.....	36
5.2 Stärken und Schwächen des Kopfhörers.....	37
5.3 Aktuelle Marktsituation.....	39
6. Binauraltechnik.....	40

6.1 Grundlagen.....	40
6.2 HRTF.....	41
6.3 Kunstkopfstereophonie.....	42
6.3.1 Probleme.....	43
6.4 Akustische Raumsimulation.....	45
6.4.1 Exkurs: Das LTI-System.....	45
6.4.2 Die binaurale Raumimpulsantwort BRIR.....	45
6.4.3 Die personalisierte Raumimpulsantwort PRIR.....	47
6.4.4 Faltung.....	48
6.4.5 Zusammenfassung.....	49
7. Messung einer PRIR mit dem Smyth Realiser A8.....	51
7.1 Das Gerät im Überblick.....	51
7.2 Vorbereitungen.....	52
7.3 Die Messung.....	53
7.4 Ergebnis.....	54
7.5 SVS-Dateiformat.....	54
7.6 Workaround.....	55
8. Konzeption einer Applikation zur Faltung auf Mobilgeräten.....	57
8.1 Externe Faltung in einer DAW.....	57
8.2 Anforderungen.....	59
8.2.1 Skalierbarkeit.....	60
8.2.1.1 Qualitätsparameter.....	60
8.3 Entwurf der Rendering Engine.....	61
8.4 Leistungsmessung.....	62
9. Fazit.....	64
Literaturverzeichnis.....	66
Anhang.....	69

Abbildungsverzeichnis

Abbildung 1: Aufbau des menschlichen Gehörs.....	13
Abbildung 2: Kopfbezogenes Koordinatensystem nach Blauert.....	16
Abbildung 3: Einfallswinkel ψ einer ebenen Schallwelle beim Auftreffen auf den Kopf.....	17
Abbildung 4: Interaurale Zeitdifferenzen von Trägerschwingung und Hüllkurve.....	17
Abbildung 5: Relative Häufigkeit der Versuchspersonen, die mit 95% Sicherheit eine der drei Antworten „h“, „o“, „v“ häufiger angeben als die beiden anderen zusammen.	19
Abbildung 6: Lokalisation zw. Schallquellen- und Hörereignisentfernung.....	22
Abbildung 7: Abhängigkeit der Hörereignisentfernung vom Schalldruckpegel am Hörerort.....	22
Abbildung 8: Kurven gleicher Lautstärke.....	22
Abbildung 9: Häufigkeit des mobilen Musikkonsums, 159 befragte Personen.....	26
Abbildung 10: Stereodreieck mit Lautsprechern L1 und L2, Hörer H, Phantomschallquelle S, Basisbreite b und Lokalisationswinkel α	30
Abbildung 11: Lautsprecheranordnung nach ITU-R BS.775-1.....	32
Abbildung 12: Basiskonfiguration "Auro 3D 9.1", kompatibel mit ITU-R BS.775-1. Rechts: Elevationswinkel der Höhenlautsprecher.....	33
Abbildung 13: Vergleich der HRTF-Kurven des linken Ohres. Die weißen Kurven stellen die bewerteten Mittelwerte dar.....	42
Abbildung 14: KU100 der Firma Neumann.....	43
Abbildung 15: LTI-System im Zeitbereich.....	45
Abbildung 16: Raumimpulsantwort $h(n)$ und vereinfachte Aufteilung der Raumimpulsantwort in Direktsignal, erste Reflexionen und diffusen Nachhall (Darstellung als $ h(n) $).....	46
Abbildung 17: Schnelle Faltung.....	49
Abbildung 18: Zusammenfassung Binauraltechnik.....	50
Abbildung 19: Screenshot des Workaround-Projektes in Logic Pro 9.....	55
Abbildung 20: Das Hallplugin "Space Designer".....	57
Abbildung 21: Rendering Engine zur Synthese einer 5.1 Aufnahme.....	61
Abbildung 22: Ergebnisse für die einzelnen Qualitätsstufen der Synthese: maximale Rechenzeit = 35449 μ s, maximale Arbeitsspeicherbelastung = 8,38 MB, maximale Kanalanzahl = 28.....	63

Tabellenverzeichnis

Tabelle 1: Attribute der räumlichen Abbildung und prinzipielle Möglichkeiten der Übertragungsverfahren.....	41
Tabelle 2: Die einzelnen Qualitätsstufen und ihre Ressourcennutzung in der Übersicht.....	62

Abkürzungsverzeichnis

AIFF	Audio Interchange File Format
ARM	englischer Prozessorhersteller
BRIR	Binaural Room Impulse Response
CD	Compact Disc
DAW	Digital Audio Workstation
dB	dezibel
DS	Dynamic Sources
DVD-A	Digital Versatile Disc Audio
ER	Early Reflections
FC	Filter Crossfading
FFT	Fast Fourier Transformation
FFT ⁻¹	inverse Fast Fourier Transformation
FIF	Filter Interpolation in the Frequency domain
FIT	Filter Interpolation in the Time domain
FLAC	Free Lossless Audio Codec
FUR	Filter Update Rate
Gfu	Gesellschaft für Unterhaltungs- und Kommunikationselektronik
HEQ	Headphone Equalization
HiFi	High Fidelity
HRTF	Head Related Transfer Function
ILD	Interaural Level Difference
IR	Impulse Response

ITD	Interaural Time Difference
ITU	International Telecommunication Union
KH	Kopfhörer
LA	Limited angular range in Azimut
LE	Limited angular range in Elevation
LFE	Low Frequency Effects Channel
LTI-System	Linear Time-Invariant System
mp3	MPEG-Layer 3
OS	Operating System
phon	Maßeinheit für den (psychoakustischen) Lautstärkepegel
PRIR	Personalised Room Impulse Response
SACD	Super Audio Compact Disc
SDII	Sound Designer II, Audioformat
SM	Smoothing of sound source movement
TD-Mode	Tracker Disable (Funktion des Realiser A8)
WAV	WAVE File Format, Audioformat

1. Einleitung

Die konsequente Weiterentwicklung elektronischer Medien führte in den letzten Jahren sowohl im visuellen wie auch auditiven Bereich zu neuen Technologien, die versuchen, das mediale Erlebnis immer realistischer zu gestalten. Raumbezogene Wiedergabestandards ermöglichen durch steigenden Hardwareeinsatz eine dreidimensionale und zunehmend immersive Darstellung der Inhalte.

Der Mehrkanalton bildet hierbei keine Ausnahme. Waren seine Vorteile noch vor einiger Zeit nur im Kino zu genießen, so ist er mittlerweile - häufig als Surroundinstallation in Kombination mit stereoskopischem Bildmaterial - auch in der Heimanwendung verbreitet. Neue Wiedergabeanordnungen wie Auro3D können zwar durch eine erhöhte Anzahl von Lautsprechern realisiert werden, doch bei diesem hardwarebasierten Wettrüsten bleibt der weit verbreitete mobile Musikgenuss aufgrund seiner begrenzten Möglichkeiten leider außen vor.

Der mobile Konsum von Musik setzt die Verwendung eines Kopfhörers voraus und ist damit auf zwei Kanäle beschränkt. Dieser Umstand macht sich indirekt auch in der Musikproduktion bemerkbar. Es werden nach wie vor sehr wenige Mehrkanalmischungen angefertigt, weil sie durch die zunehmende Ausrichtung der Branche auf den mobilen Markt bis auf wenige Ausnahmen nicht rentabel sind. Da ein Großteil der Studios den Lautsprecher als Arbeitswerkzeug bevorzugt, kommt es bei der Benutzung von Kopfhörern zu Abbildungsfehlern.

Eine Lösung dieser Probleme kann mit Hilfe der Binauraltechnik erreicht werden. Durch die Verbreitung leistungsstarker mobiler Geräte wie Smartphones und Tabletcomputer wird die binaurale Reproduktion des Originalschallfeldes zu einer vielversprechenden Möglichkeit, ein dreidimensionales Hörerlebnis ohne großen technischen Aufwand und an jedem beliebigen Ort zu realisieren. Gleichzeitig werden sämtliche Abbildungsfehler der Kopfhörerwiedergabe von Lautsprechermischungen behoben und der mobile Markt für Mehrspuraufnahmen geöffnet.

Diese Arbeit versteht sich unter Anderem als Leitfaden für die Implementierung einer Syntheseapplikation und versucht, die grundlegenden Theorien und Zusammenhänge verständlich darzustellen. Rein informatische Inhalte wie konkrete Programmierbeispiele sind nicht enthalten.

2. Grundlagen zum menschlichen Gehör

2.1 Aufbau des Hörorgans

Ein menschliches Ohr hat die Aufgabe, Luftdruckschwankungen eines Schallfeldes in Nervenreize umzuwandeln¹. Es wird allgemein in drei Bereiche gegliedert: das Außenohr, das Mittelohr und das Innenohr. Jedem dieser Teilbereiche kommen beim Vorgang des Hörens spezifische Aufgaben zu, die im Folgenden kurz angesprochen werden. Die Ohrmuschel und der Gehörgang bilden zusammen das Außenohr und sind für die Aufnahme von Schallwellen verantwortlich. Diese werden im Mittelohr vom Trommelfell, einer 0,1mm dicken Membran², in mechanische Schwingungen umgewandelt und auf die Gehörknöchelchenkette aus Hammer, Amboss und Steigbügel übertragen. Die Eustachische Röhre ist mit dem Rachenraum verbunden und sorgt für einen stetigen Druckausgleich im Mittelohr. Das Innenohr besteht aus der Gehörschnecke und dem Gleichgewichtsorgan, zusammen bilden sie das sogenannte Labyrinth. Schnecke und Gleichgewichtsorgan sind mit einer Flüssigkeit gefüllt. Die mechanische Auslenkung der beweglichen Gehörknöchelchen bringt die Flüssigkeit im Inneren der Gehörschnecke in Bewegung. Kleine Haarzellen registrieren diese Veränderungen und geben sie als elektrische Impulse über den Hörnerv an das Gehirn weiter. Das Gehör ist so in der Lage, Frequenzen von 20 Hz bis zu 20 kHz wahrzunehmen. Die Übertragungskette der Schallwellen und deren Verarbeitung von der

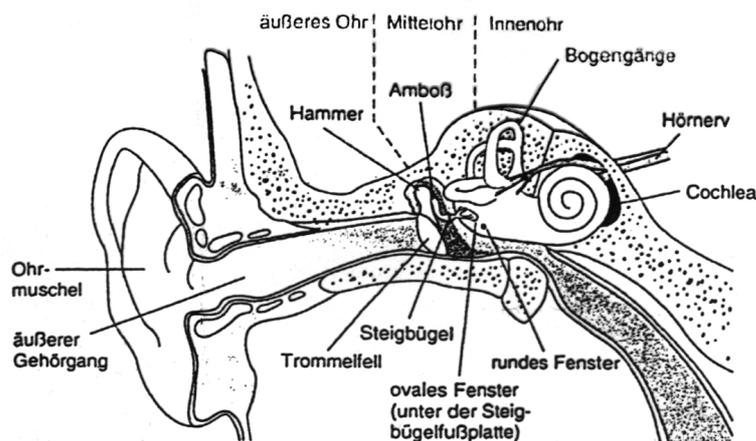


Abbildung 1: Aufbau des menschlichen Gehörs

¹ Vgl. Webers 2007, S. 93

² Vgl. Blauert 1974, S. 43

Quelle bis in unser Gehirn ist maßgeblich an der Wahrnehmung akustischer Ereignisse beteiligt. In der vorliegenden Arbeit steht allerdings die Funktion des Außenohres in Bezug auf seine richtungbestimmenden Eigenschaften beim räumlichen Hören im Vordergrund.

2.2 Räumliches Hören

Für alle folgenden Ausführungen wird zur eindeutigen Trennung von physikalischen Phänomenen und der menschlichen Wahrnehmung die Terminologie nach Blauert verwendet. Physikalische Begriffe werden stets in Kombination mit dem Wort „Schall-“ gebraucht, wahrnehmungsbedingte mit der Vorsilbe „Hör-“. Als Beispiel nennt Blauert die Begriffe Schallereignis und Hörereignis, und unterscheidet außerdem den Entstehungsort des Schalls und den Ort, an dem er wahrgenommen wird³. Diese beiden Ortsangaben stimmen zwar häufig, jedoch nicht immer überein.

Wie im letzten Abschnitt zum Aufbau des menschlichen Gehörs schon erwähnt, kommt dem Außenohr beim räumlichen Hören große Bedeutung zu. Die Ohrmuschel und der Gehörgang bilden zusammen ein komplexes Resonanzsystem, dessen Erregbarkeit von der Richtung und Entfernung der Schallquelle abhängt⁴. Der Schall wird von diesem System so in seiner spektralen Zusammensetzung verändert, dass er Informationen über seinen Entstehungsort enthält, die unser Gehirn auswerten kann. Die resultierenden Schallsignale werden als Ohrsignale bezeichnet. Sie können mit Hilfe von Mikrofonen im Gehörgang gemessen werden. Schon geringe Veränderungen dieser Ohrsignale können zu einer Diskrepanz zwischen Schallereignisort und Hörereignisort führen. So kommt es häufig vor, dass beim Hören mit fremden Ohrsignalen wie sie z.B. ein Kunstkopf⁵ erzeugt, Schallereignisse die sich ursprünglich vorne befanden vom Hörer plötzlich hinten wahrgenommen werden. Die Genauigkeit bei der Angabe von Ort und Dauer eines Schallereignisses hängt außerdem stark von dessen Gestalt und dem Raum ab, in dem es stattfindet. Beispielsweise ist die Herkunft eines tiefen, gehaltenen Orgeltons in einer großen, reflektierenden Kirche wesentlich schwieriger zu bestimmen, als die eines Pistolenschusses im Freien. In manchen Situationen kommt erschwerend hinzu, dass kein anderer Sinnesreiz mit einem Schallereignis verknüpft werden kann. Man denke an einen dunklen Raum oder Hindernisse wie eine

3 Vgl. Blauert 1974, S. 2

4 Ebd, S. 55

5 Siehe Kapitel 6.3

Wand, die uns die Sicht auf dahinter liegende Schallquellen versperrt⁶. Oder stellen Sie sich vor, in der Savanne schleiche sich ein Raubtier von hinten an einen ihrer Vorfahren heran. Glücklicherweise besitzt unser Gehör, im Gegensatz zu anderen Sinnen wie dem Sehen, einen entscheidenden Vorteil: Es nimmt Reize aus jeder Richtung wahr.

2.2.1 Lokalisation

Vorweg soll definiert werden, was man unter Lokalisation, und unter Lokalisationsunschärfe versteht:

Lokalisation: „Zuordnungsgesetz oder -regel [...] zwischen dem Ort eines Hörereignisses (z. B. bezüglich Richtung und/oder Entfernung) und einem bestimmten Merkmal oder bestimmten Merkmalen eines Schallereignisses oder eines anderen, mit dem Hörereignis korrelierten Ereignisses. [...]“⁷

Lokalisationsunschärfe: „Kleinste Änderung eines bestimmten Merkmals oder bestimmter Merkmale des Schallereignisses oder eines anderen, mit dem Hörereignis korrelierten Ereignisses, die gerade zu einer Ortsänderung des Hörereignisses (z. B. Bezüglich Richtung und/oder Entfernung) führt. Die Lokalisationsunschärfe ist eine Eigenschaft der Lokalisation.“⁸

Vereinfacht könnte man sagen, die Lokalisation meint die Zuordnung des Gehörten zu einer Schallquelle, die Lokalisationsunschärfe beschreibt die Genauigkeit, mit der uns diese Zuordnung möglich ist. Um die Ausführungen zur Lokalisation von Schallereignissen zu erleichtern, wird ein kopfbezogenes Koordinatensystem wie in Abbildung 2 eingeführt. Kopfbezogen bedeutet, dass es seine Lage den Bewegungen des Kopfes entsprechend verändert, wie es auch unsere Ohren tun.

⁶ Vgl. Blauert 1974, S. 3

⁷ Blauert 1974, S. 30

⁸ Ebd.

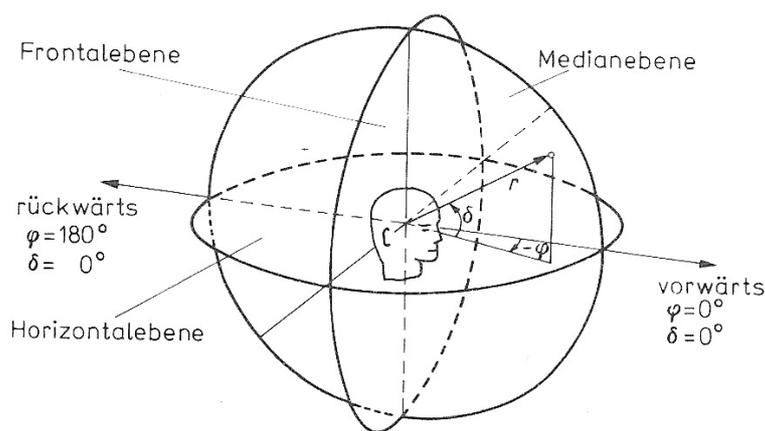


Abbildung 2: Kopfbezogenes Koordinatensystem nach Blauert

Wir gehen zunächst von nur einer Schallquelle aus. Die Horizontalebene befindet sich genau auf Ohrniveau und liefert im Gegensatz zu Frontal- und Medianebene keine Höheninformationen. Die Frontalebene deckt alle Schallereignisse in einem Winkel von 90° zur Blickachse ab. Auf der Blickachse liegt die Medianebene, sie teilt den Kopf in linke und rechte Seite und bildet die Symmetrieebene. Das Koordinatensystem soll die Ortsangabe von Schallquellen erleichtern.

Das Gehör verwendet beim Richtungshören verschiedene Mechanismen, die jeweils auf unterschiedlichen physikalischen Eigenschaften der Schallquelle beruhen.

2.2.1.1 Richtungshören in der Horizontalebene

Bei der ersten hier angewandten Methode, der **Lokalisation durch interaurale Laufzeitunterschiede** (ITD: „Interaural Time Difference“), wertet das Gehirn den Zeitunterschied zwischen linkem und rechtem Ohr aus; man spricht auch von binauralem („zweiohrigem“) Hören. Trifft ein Schall aus seitlichem Einfallswinkel auf den Kopf, so benötigt er für die Strecke von der Quelle zu dem ihm zugewandten Ohr weniger Zeit, als für die Strecke zum abgewandten Ohr. Dabei ist das Gehirn in der Lage, sowohl die zeitliche Verschiebung der Trägerfrequenz, als auch die der Hüllkurve beider Ohrsignale auszuwerten⁹. Bei Frequenzen unterhalb von 1,6 kHz ist eine Bewertung anhand der Hüllkurvenverschiebung allerdings nicht mehr möglich.

⁹ Vgl. Blauert 1974, S. 132

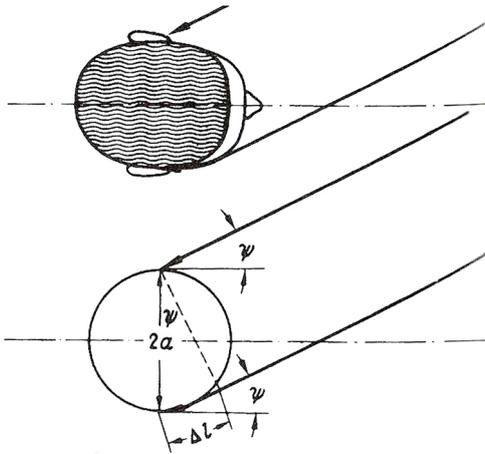


Abbildung 3: Einfallswinkel ψ einer ebenen Schallwelle beim Auftreffen auf den Kopf

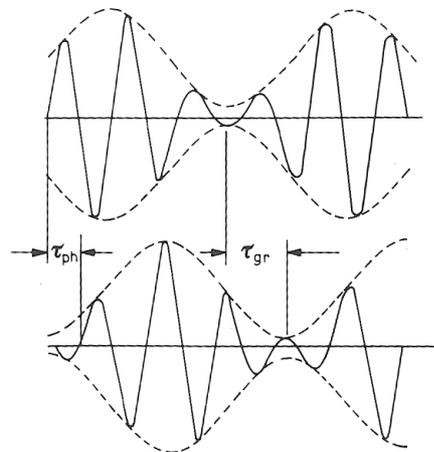


Abbildung 4: Interaurale Zeitdifferenzen von Trägerschwingung und Hüllkurve

Bei einem durchschnittlichen Kopfdurchmesser von 17cm entsteht ein maximaler Laufzeitunterschied von $0,63 \text{ ms}^{10}$. Die geringste erreichbare Lokalisationsunschärfe und damit das maximale räumliche Auflösungsvermögen unseres Gehörs ergibt sich aus der kleinsten wahrnehmbaren Laufzeitdifferenz. Diese liegt bei $0,03 \text{ ms}$, woraus eine Lokalisationsgenauigkeit von $3^\circ - 5^\circ$ nahe der Blickachse abgeleitet werden kann¹¹.

Bei der zweiten Methode, der **Lokalisation durch interaurale Intensitätsunterschiede** (ILD: „Interaural Level Difference“), wertet das Gehirn Unterschiede der Schallintensität am linken und rechten Ohr aus. Sie entstehen, da die Schallenergie auf dem Weg von einem Ohr zum anderen quadratisch abnimmt ($1/r^2$ -Gesetz)¹². Bei zunehmender Entfernung zur Schallquelle erreicht die Wegdifferenz zwischen den Ohren im Verhältnis zum Gesamtweg des Schalls immer kleinere Werte, die Abnahme der Schallenergie wird zu gering für die Lokalisation. Jetzt kommen verstärkt Beugungserscheinungen zum tragen, die bei Wellenlängen in der Größenordnung des Ohrabstandes auftreten. Die Beugung ist frequenzabhängig und damit umso größer, je größer

¹⁰ Vgl. Dickreiter 2008, S. 106

¹¹ Ebd

¹² Vgl. Webers 2007, S. 60

die Wellenlänge in Relation zu ihrem Hindernis ist¹³. Die Schallwellen beugen sich also je nach Frequenz mehr oder weniger stark um das Hindernis (den Kopf) herum, wodurch frequenzabhängige Intensitätsunterschiede entstehen, die das Gehirn auswerten kann. Sehr kleine Wellenlängen werden vom Kopf reflektiert, was zu großen Differenzen zwischen beiden Ohren führt. Bei sehr großen Wellenlängen von Frequenzen unter 300 Hz stellt er kein Hindernis mehr dar, und es können keine interauralen Intensitätsunterschiede mehr festgestellt werden. In sehr tiefen Frequenzbereichen nimmt die Lokalisationsunschärfe deutlich zu.

Beim Vorgang des natürlichen Hörens kann keine eindeutige Grenze zwischen beiden Methoden gezogen werden. Vielmehr ergibt oft erst die Kombination beider Verfahren eine gute räumliche Wahrnehmung. Das Gehirn ist sogar in der Lage, bestimmten Pegeldifferenzen entsprechende Laufzeitunterschiede zuzuordnen und umgekehrt. Welches Verfahren bevorzugt zum Einsatz kommt, hängt stark von der Beschaffenheit des Schallereignisses ab. Wie oben schon erwähnt, entstehen bei tiefen Frequenzen keine verwertbaren Intensitätsunterschiede mehr. Auf der anderen Seite kann das Gehör für reine Töne, deren Wellenlänge genau ein- oder mehrfach dem Ohrabstand entspricht keine Laufzeitunterschiede mehr feststellen, da in diesem Fall kein Phasenversatz auftritt¹⁴. Darüber hinaus besteht auch die Möglichkeit, dass zwei Schallereignisse aus unterschiedlichen Richtungen den gleichen Laufzeitunterschied zwischen den Ohren hervorrufen. Betrachtet man die Horizontalebene als Kreis und geht von 0° in Blickrichtung aus, so erzeugen zwei identische Schallereignisse aus den Richtungen 30° und 150° auch identische Laufzeiten. Folglich könnte man annehmen, dass der Mensch akustisch nicht in der Lage ist, zwischen Schalleinfall von vorne und hinten zu unterscheiden. Allerdings wissen wir aus eigener Erfahrung, dass diese Vermutung falsch ist.

2.2.1.2 Richtungshören in der Medianebene

Die Lokalisation von Schallereignissen in der Medianebene stellt im Vergleich zu den beiden anderen Ebenen einen Spezialfall dar. Der Abstand der Schallquelle zum linken und rechten Ohr ist hier immer gleich, weshalb weder interaurale Laufzeit- noch Pegeldifferenzen auftreten können. Das bedeutet, dass an beiden Ohren identische Signale anliegen und wirft die Frage auf, wie eine Richtungsbestimmung hier überhaupt möglich ist. Bei Hörversuchen mit einer Reihe von

¹³ Vgl. Webers 2007, S. 76

¹⁴ Ebd., S. 124

Testpersonen konnte Blauert nachweisen, dass die Ohrmuschel für die Unterscheidung der Höreignisrichtungen vorne und hinten eine wesentliche Rolle spielt. Den Testpersonen wurden aus beiden Richtungen Signale vorgespielt und mit einem Sondenmikrofon im Gehörgang aufgenommen. Bei anschließender Wiedergabe setzte man den Einfluss der Ohrmuschel durch eine Entzerrung der Wiedergabevorrichtung außer Kraft. Die Personen hörten also die exakt gleichen Signale wie bei der Aufnahme. Bei allen Testhörern stimmten die Höreignisrichtungen von Aufnahme und Wiedergabe überein¹⁵. Ein weiterer Faktor beim Hören in der Medianebene ist die Bandbreite des Signals. Wie im Abschnitt *Räumliches Hören* zu Beginn dieses Kapitels schon erwähnt wurde, bildet die Ohrmuschel zusammen mit dem Gehörgang ein Resonanzsystem, das bestimmte Frequenzen in Abhängigkeit vom Schalleinfallswinkel verzerrt und gewichtet. Daraus lässt sich schließen, dass breitbandige Signale deutlich genauer und mit größerer Sicherheit lokalisiert werden können, da hier dem Außenohr mehr Frequenzen für die Verzerrung zur Verfügung stehen als bei schmalbandigen Signalen. Die Lokalisationsunschärfe in der Medianebene kann dabei bis auf 4° sinken¹⁶. Diese Beobachtung wirft die Frage auf, welche Frequenzanteile vom Außenohr bei den verschiedenen Schalleinfallsrichtungen verändert werden. Eine Antwort geben die **Richtungsbestimmenden Bänder** nach Blauert.

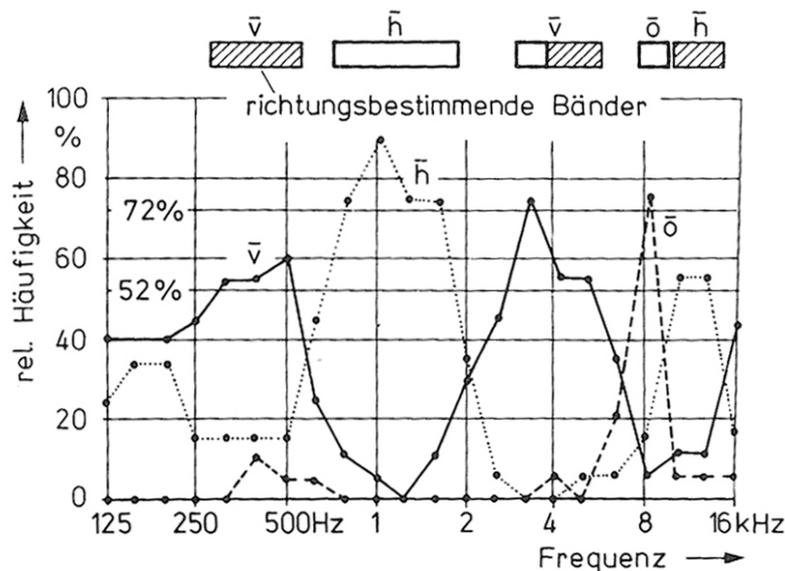


Abbildung 5: Relative Häufigkeit der Versuchspersonen, die mit 95% Sicherheit eine der drei Antworten „h“, „o“, „v“ häufiger angeben als die beiden anderen zusammen.

¹⁵ Vgl. Blauert 1974, S. 83

¹⁶ Ebd, S. 84

In Abbildung 5 sind die Richtungsbestimmenden Bänder eingezeichnet. Die Bereiche zeigen, bei welchen Frequenzen das Hörereignis der absoluten Mehrheit von Versuchspersonen in dieser Richtung auftritt, und zwar häufiger als in den beiden anderen Richtungen zusammen. In den weissen Bereichen ist dies mit einer Sicherheit von 90%, in den schraffierten am wahrscheinlichsten der Fall. Beispielsweise lässt sich für eine Beschallung von oben eine deutliche Anhebung im Bereich um 8 kHz herauslesen. Die Filterwirkung des Außenohres und des Kopfes sind demnach für eine korrekte Lokalisation sehr wichtig. Außerdem steigt mit zunehmender Erfahrung die Wahrscheinlichkeit eine Schallquelle richtig einzuordnen; das Gehirn lernt die entstehenden Klangfarben besser zu unterscheiden. Auch der Bekanntheitsgrad des Signals ist hilfreich. Je vertrauter eine Person mit einem Schallereignis ist, desto geringer wird die Lokalisationsunschärfe¹⁷.

Bei der Lokalisation von mehreren Schallquellen spielt der Präzedenz-Effekt (auch „Haas-Effekt“ oder „Gesetz der ersten Wellenfront“) eine tragende Rolle. Treffen die Schallwellen mehrerer Quellen aus unterschiedlichen Richtungen mit einer sehr geringen zeitlichen Verzögerung von nur wenigen Millisekunden auf das Trommelfell, so ist der früher auftreffende Reiz ausschlaggebend für die Lokalisation¹⁸. Nachfolgende Reize werden vernachlässigt.

2.2.2 Entfernungswahrnehmung

Der zweite große Bereich des räumlichen Hörens ist neben der Lokalisation die Entfernungswahrnehmung. Zwar ist sie weniger ausführlich erforscht, doch es besteht gegenüber der Richtungswahrnehmung ein Vorteil, denn Entfernungen kann der Mensch bei monauralem Hören („einohriges Hören“) nahezu genauso gut erkennen wie bei binauralem¹⁹. Dabei werden vom Gehör drei verschiedene Phänomene ausgewertet.

Bei der **Entfernungswahrnehmung durch Hallerscheinungen** steht das Verhältnis von direktem Schall, der das Ohr ohne Umwege erreicht, und indirektem, diffusem Schall im Vordergrund²⁰. In der freien Natur kommt dieser Diffusschall nur sehr selten als Echo vor, in geschlossenen Räumen jedoch reflektieren Wände und andere Flächen den Schall, sodass ein diffuses Schallfeld mit mehr oder weniger langen Nachhallzeiten entsteht. Je größer also der Diffus-

17 Vgl. Blauert 1974, S. 85

18 Ebd, S. 163

19 Vgl. Webers 2007, S. 125

20 Ebd.

schallanteil im Vergleich zum Direktschall, desto weiter entfernt erscheint uns das Schallereignis. Umgekehrt, bei überwiegendem Direktschall, entsteht der Eindruck von Nähe. Dieser Effekt wird in der Postproduktion und vor allem beim Mischen von Musik sehr häufig genutzt, da sich durch nachträgliches Hinzufügen von künstlichem Hall der Entfernungseindruck beeinflussen lässt.

Die **Entfernungswahrnehmung durch Klangfarbenunterschiede** nutzen wir vorwiegend im Freifeld, da sie eher bei großen Distanzen eine Rolle spielt. Auf dem langen Weg, den der Schall von der Quelle zum Hörer zurücklegen muss, gehen hohe Frequenzen stärker verloren als tiefe Frequenzen, die von der Luft weniger stark gedämpft werden. Die Klangfarbe des Schalls wird entfernungsabhängig verändert, das Gehirn wertet diese Veränderungen aus. Ein Beispiel: Wenn Sie sich aus großer Entfernung einem Konzert nähern, werden Sie feststellen, dass Sie zuallererst den Bass wahrnehmen können, bevor mit abnehmender Entfernung auch die anderen, höherfrequenten Instrumente dazukommen.

Eine dritte Möglichkeit stellt die **Entfernungswahrnehmung durch Lautstärkeänderungen** dar. Sie basiert auf dem $1/r$ -Gesetz, nach dem der Schalldruck um 6 dB bei doppelter Entfernung von der Schallquelle abnimmt. Sie wird vorwiegend bei mittleren Distanzen zwischen drei bis fünfzehn Metern relevant. Das Gehirn ist in der Lage, aufgrund des Schalldruckpegels der Ohrsignale, die Entfernung zur Schallquelle einzuschätzen; allerdings muss das Signal dafür bekannt sein. Erfahrungswerte spielen hier eine wichtige Rolle. Bei unbekanntem Schallereignissen kann es vorkommen, dass solche mit großer Schallleistung in großer Entfernung, gleich laut wahrgenommen werden, wie solche mit kleiner Schallleistung in geringer Entfernung²¹. Daraus ist abzuleiten, dass Schallquellenentfernung und Hörereignisentfernung voneinander unabhängig sind²². Die rechte Grafik zeigt das Verhältnis von Schallquellen- und Hörereignisentfernung, letztere steigt bei zunehmender Distanz langsamer an, was auf eine maximal wahrnehmbare Entfernung schließen lässt. Der linken Grafik ist zu entnehmen, dass erst eine Schalldruckpegelabnahme von 20 dB zu einer Verdopplung der Hörereignisentfernung führt, und nicht wie erwartet eine Abnahme von 6 dB.

21 Vgl. Webers 2007, S. 125

22 Vgl. Blauert 1974, S. 99

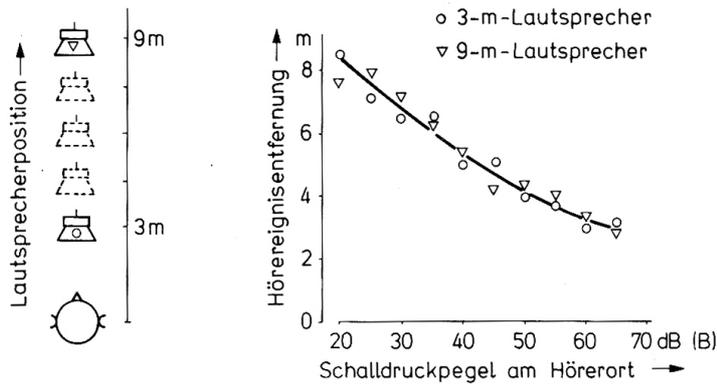


Abbildung 7: Abhängigkeit der Hörereignisentfernung vom Schalldruckpegel am Hörerort

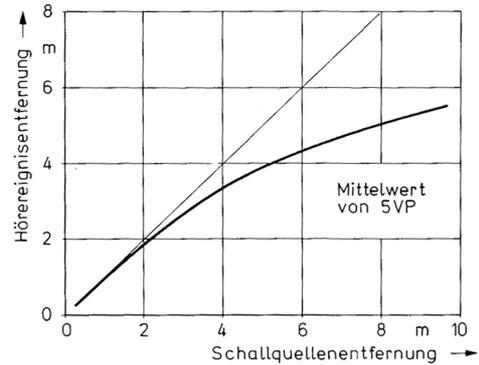


Abbildung 6: Lokalisation zw. Schallquellen- und Hörereignisentfernung

Als letzter Punkt ist zu beachten, dass durch das inkonsistente Lautstärkeempfinden des Menschen bei unterschiedlichen Frequenzen laute Schallereignisse in der Regel in größerer Entfernung wahrgenommen werden als leise, da bei lauten Schallereignissen eine Überbetonung der tiefen Frequenzen auftritt. Die in Abbildung 8 dargestellten Kurven gleicher Lautstärke von Fletcher und Munson aus dem Jahre 1936 zeigen das menschliche Lautstärkeempfinden in Abhängigkeit des Schalldruckpegels über das wahrnehmbare Frequenzspektrum von 20 Hz bis 20 kHz. Der Lautstärkepegel wird in der Einheit *phon* angegeben.

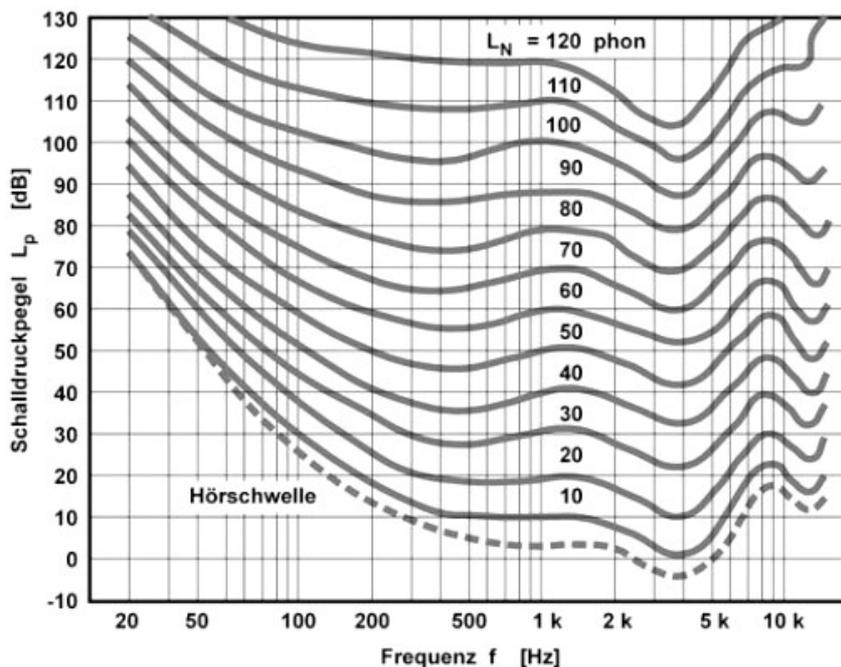


Abbildung 8: Kurven gleicher Lautstärke

3. Mobiles Wachstum und seine Auswirkungen auf den Musikkonsum

3.1 Technischer Fortschritt

Die Revolution des mobilen Musikgenusses markiert eine Erfindung des deutschen Fraunhofer-Institutes für integrierte Schaltungen IIS in Zusammenarbeit mit der Universität Erlangen-Nürnberg aus den 90er Jahren²³: Das Audioformat MPEG Layer 3, kurz mp3. Es komprimiert und speichert Audioinhalte unter Berücksichtigung psychoakustischer Effekte und kommt so mit einem Zehntel der ursprünglichen Dateigröße aus²⁴. Mp3 war „[...] der Schlüssel zur einfachen Speicherung und Übertragung von Musik [...]“²⁵ - Harald Popp, mp3-Entwickler. Ursprünglich zur Distribution von Audiosignalen über eine Telefonleitung entwickelt, ließen die ersten mp3-Player durch den Erfolg des Formates nicht lange auf sich warten. Den endgültigen Durchbruch schaffte dann der im Jahre 2001 von Apple vorgestellte iPod²⁶. Mit dem Werbeslogan „1000 songs in your pocket“ war er der erste tragbare Musikplayer mit einer Speicherkapazität von 5 GB²⁷ und damit trotz des horrenden Preises von rund 1000 DM seiner Konkurrenz weit voraus. Heute ist jedes aktuelle Handy in der Lage, über einen integrierten Musikplayer mp3-Dateien und diverse andere Audioformate wiederzugeben. Von großer Bedeutung für die vorliegende Arbeit ist jedoch ein anderer historischer Tag in der Geschichte der mobilen Elektronik. Am 09.01.2007 stellte Steve Jobs auf der Macworld Conference in San Francisco das Apple iPhone der ersten Generation vor. Es vereinte die Funktionen eines Telefons und eines iPods mit denen eines internetfähigen Gerätes und ließ sich via Touchscreen bedienen. Weniger ein Telefon als viel mehr ein Taschencomputer, diente es als Vorbild für unzählige Modelle. Heutige Mobilgeräte haben in ihrer technischen Ausstattung ein Niveau erreicht, das sie zur ernsthaften Konkurrenz von Laptops und Deskto-

23 MP3 Infobroschüre des Fraunhofer IIS,

http://www.iis.fraunhofer.de/content/dam/iis/de/dokumente/amm/broschueren/mp3_Broschuere_A4_16S2012.pdf, S. 10, Stand: 25.10.2013

24 Ebd, S. 2

25 Ebd, S. 6

26 Entwicklung des iPod, <http://www.apple.com/pr/products/ipodhistory/> Stand: 25.10.2013

27 Zehn Jahre MP3-Player: Als Digitalmusik tragbar wurde, <http://www.spiegel.de/netzwelt/spielzeug/zehn-jahre-mp3-player-als-digitalmusik-tragbar-wurde-a-540664.html>, Stand: 25.10.2013

prechnern werden lässt. Speziell aktuelle Tablet-PCs verringern mit ihrem Leistungsumfang und großen Bildschirmen den Absatz klassischer PCs merklich. Dabei ist im Hinblick auf die technische Weiterentwicklung der Geräte ein Ende vorerst nicht abzusehen. Die Verkaufszahlen stiegen in den letzten Jahren kontinuierlich; so wurden 2012 in Deutschland 673 544 Kinder geboren, aber ca. 27 mal so viele Smartphones gekauft, nämlich rund 18 Millionen Stück²⁸. Für das Jahr 2013 wird ein Absatz von rund 26 Millionen Geräten erwartet²⁹. Jeder dritte Deutsche über 14 Jahre ist mittlerweile im mobilen Internet unterwegs³⁰, was sich täglich in der Öffentlichkeit beobachten lässt. Mobile Geräte sind in unserer Gesellschaft omnipräsent geworden. In der U-Bahn, in Cafés und Restaurants, auf öffentlichen Plätzen, in der Fußgängerzone und im Park; überall trifft man Menschen mit einem Smartphone an. Einer Studie von Google zufolge, bei der eintausend deutsche Smartphonebesitzer befragt wurden, gehen 64 % von ihnen nicht mehr ohne ihr Gerät aus dem Haus. Interessanterweise gaben sogar 97 % der Befragten an, es auch zu Hause zu verwenden³¹. Im Rahmen dieser Arbeit stellt sich also die Frage, inwieweit die Verbreitung mobiler Endgeräte unsere Art Musik zu hören verändert hat.

3.2 Marktentwicklung

Noch vor wenigen Jahren war der mp3-Player unangefochtener Spitzenreiter bei der mobilen Musikwiedergabe. Mittlerweile hat sich das Blatt gewendet, das Smartphone nimmt jetzt seinen Platz ein. 77 % der Smartphone-Besitzer nutzen es als Musikplayer. Jeder zweite Deutsche hat damit schon einmal unterwegs Musik gehört, immerhin 17 % mit dem Tablet³². Doch auch zu Hause dient es vielen Menschen als zentrale Musikbibliothek in Verbindung mit ihrer Hifi-Anlage.

28 Infografik: Nicht ohne mein Handy – Handynutzung in Deutschland und der Welt,

<http://mobilbranche.de/2013/10/infografik-nicht-handy/39586>, Stand 26.10.2013

29 Consumer Electronics 2013: Marktentwicklung und Trends,

http://www.bitkom.org/files/documents/BITKOM_Praesentation_CE-PK_02_09_2013.pdf, Stand: 26.10.2013

30 „Mobile“ knackt 20-Millionen-Grenze: Mobiles Wachstum ungebremst,

http://www.iqm.de/fileadmin/user_upload/Marketing_Service/Downloads/AGOF/Pressemitteilung_AGOF_moble_facts_2012-II.pdf, Stand: 26.10.2013

31 Google Studie - Unser mobiler Planet: Deutschland,

http://services.google.com/fh/files/blogs/our_mobile_planet_germany_de.pdf, S. 8 + 9, Stand 26.10.2013

32 Musiknutzung, http://www.musikindustrie.de/jahrbuch-musiknutzung-2012/?no_cache=1&type=1, Stand: 26.10.2013

Dabei war der Erwerb von Musik noch nie so einfach. Das mp3-Format schuf in Kombination mit dem Internet die Grundlage für neue Distributionsplattformen. Mit einem Klick bezahlt der Kunde über ein Onlinezahlverfahren und bekommt direkt im Anschluss seinen Einkauf als Download zur Verfügung gestellt. Bekannte Beispiele sind der iTunes Store, Amazon MP3 oder Musicload. Laut dem Jahreswirtschaftsbericht 2012 des Bundesverbandes Musikindustrie ist seit Jahren ein konstanter zweistelliger Zuwachs der digital generierten Umsätze zu beobachten³³. Der durch Downloads erbrachte Anteil stieg demnach um 24 %. Besonders junge Menschen nutzen die neu entstandenen Möglichkeiten der digitalen Vertriebswege.

Auf der anderen Seite schrumpft der Markt mit physischen Tonträgern kontinuierlich, wenn auch nur in geringem Maße. Nach wie vor bleibt die CD mit einem Anteil von 71 % am Gesamtumsatz das wichtigste Medium³⁴. Daraus lässt sich schließen, dass das Album eines Künstlers immer noch als Gesamtkunstwerk geschätzt wird. Es hat sich außerdem gezeigt, dass der Erwerb und das Hören einer CD für viele Menschen ein bewussteres Erlebnis darstellt als die Anschaffung immaterieller Dateien über das Internet³⁵. Dies mag auch maßgeblich mit der Möglichkeit zur unendlichen Vervielfältigung digitaler Dateien zusammenhängen. Außerdem ist für bekennende Musikliebhaber die Frage der Qualität des Mediums oft von zentraler Bedeutung. Hier ist die klassische CD gegenüber dem komprimierten mp3-Format immer noch klar im Vorteil, was als Alleinstellungsmerkmal sicherlich ein weiterer Kaufgrund sein dürfte.

Als neue Form des Musikkonsums wird gerade in den letzten Monaten das Streamen von Musik immer populärer. Hier reicht das Angebot von diversen Webradios, über werbefinanzierte Freemium-Modelle mit Premiumoptionen gegen Bezahlung wie Spotify oder iTunes Radio³⁶, bis zu Anbietern von kostenpflichtigen Abonnements (z.B. Napster, simfy). Für Musikschafter ist das Streaming darüber hinaus als weiterer Teil der Wertschöpfungskette einer Produktion von Interesse. Da für jeden Abruf eines Titels auch eine Vergütung gezahlt werden muss, ist er im Gegensatz zu einer einmaligen Zahlung auch längerfristig noch rentabel. Für den Konsumenten ist das

33 Jahreswirtschaftsbericht 2012 Bundesverband Musikindustrie,

http://www.musikindustrie.de/fileadmin/piclib/statistik/branchendaten/jahreswirtschaftsbericht-2012/download/Jahrbuch_BVMI_2012.pdf, S. 9, Stand: 29.10.2013

34 Ebd., S.10

35 Vgl. Zöllner Oliver (Hrsg.) HdM 2013, S. 225

36 In Deutschland noch nicht verfügbar, Stand: 29.10.2013

Streaming allerdings kein Ersatz für den tatsächlichen Besitz von Musik³⁷. Vielmehr werden kostenfreie Angebote gerne genutzt, um neue Musik zu entdecken und sich vor einem eventuellen Kauf z.B. das entsprechende Album schon einmal anzuhören. Nicht nur für das Streaming, sondern für nahezu jede erdenkliche Form des Musikhörens gibt es heute eine Applikation. Einige von ihnen stellen den Songtext synchron zur Musik auf dem Display dar, andere erkennen auf Wunsch auch unbekannte Titel und manche kombinieren kostenfreie Musik mit sozialen Netzwerken. Die Musikapps gehören zu den am häufigsten genutzten Programmen auf den Smartphones der Deutschen. Neun von Zehn Jugendlichen im Alter zwischen 16 und 24 Jahren³⁸ nutzen sie zur Musikwiedergabe auf ihrem Mobiltelefon. Interessanterweise ist der Anteil derjenigen, die auch Geld für eine Musikapp ausgeben würden, nur sehr gering.

3.3 Auswirkungen

Durch die gesteigerte Mobilität des Menschen und die voranschreitende Globalisierung sind wir zunehmend längere Zeit unterwegs, etwa auf dem täglich Arbeitsweg, zu Fuß, im Bus, der Bahn oder mit dem Flugzeug auf Geschäftsreise. Es bieten sich uns demnach genügend Situationen, in denen wir gerne auf unsere mobile Musiksammlung zurückgreifen. Die nachfolgende Grafik verdeutlicht, dass für viele Menschen der mobile Konsum von Audioinhalten zu einem festen Teil des Alltags geworden ist.

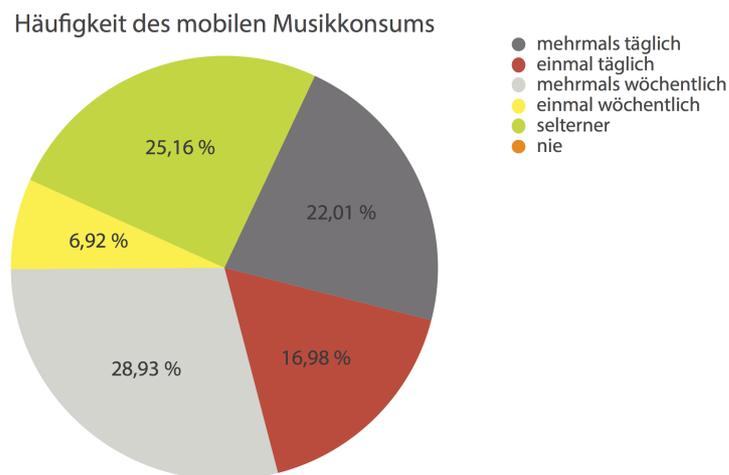


Abbildung 9: Häufigkeit des mobilen Musikkonsums, 159 befragte Personen

³⁷ Jahreswirtschaftsbericht 2012 Bundesverband Musikindustrie, S. 26

³⁸ Ebd, S. 29

Ebenfalls interessant ist, wie bewusst die Menschen unterwegs Musik konsumieren. Ein Großteil wählt aufgrund von Titeln und/oder Interpreten aus, was gespielt werden soll. Einige investieren sogar zu Hause Zeit in die Erstellung persönlicher Wiedergabelisten. Dies lässt auf ein konzentriertes und bewusstes Hören schließen. Oft werden die Stücke nach der aktuellen Stimmung des Hörers ausgewählt, etwa um sich nach einem stressigen Arbeitstag zu entspannen oder beim Sport motivieren zu lassen. Natürlich geht es auch anders. Knapp die Hälfte überlässt die Auswahl der Wiedergabetitel dem Zufallsmodus, wobei die Musik oft neben anderen Tätigkeiten, wie etwa der Arbeit oder dem Autofahren gehört wird und meist der Zerstreuung dient. Bei der gezielten Wahl von Titeln und Interpreten kann eher auf ein bewusstes Hören geschlossen werden als bei zufälliger Wiedergabe.

3.4 Zusammenfassung

Alle genannten Neuerungen und technischen Entwicklungen beeinflussen die Art, wie wir Audioinhalte konsumieren maßgeblich. Mit dem mp3-Format und den ersten tragbaren Playern wurde digitale Musik mobil. Durch das Internet entstanden neue Vertriebsplattformen, die heute mit dem Smartphone oder Tablet über das mobile Netz praktisch überall zugänglich sind. Mit der rasanten Verbreitung dieser Geräte hat sich auch unser Konsumverhalten verändert. Sie bieten uns die Möglichkeit, an jedem beliebigen Ort auf ein nahezu endloses Angebot von auditiven Inhalten zugreifen zu können. Dadurch ist die Vielfalt und Verfügbarkeit von digitaler Musik fast schon zu einem Selbstverständnis, und der mobile Konsum von Musik ein wichtiger Teil unseres Alltags geworden. Es ist also kaum verwunderlich, dass heute in nahezu allen Situationen die Möglichkeiten der Technik genutzt werden. Die Motive hierfür sind sehr unterschiedlich und von den Vorlieben und Verhaltensweisen jedes Einzelnen abhängig. So wird mobile Musik zu Informationszwecken, als Hobby, zur Zerstreuung, Isolation von der Umwelt, Überbrückung von Langweile oder Hervorrufen einer bestimmten Stimmung beim Hörer genutzt³⁹. Je nach Situation wird ihr dabei mehr oder weniger Aufmerksamkeit geschenkt. Egal auf welche Weise oder aus welchen Gründen wir unterwegs Musik hören, der mobile Musikkonsum ist zu einem Teil unserer Gesellschaft geworden und aus dem Alltag nicht mehr wegzudenken.

Bedingt durch die Mobilität gestaltet sich, im Gegensatz zu stationären Wiedergabeanordnun-

³⁹ Vgl. Zöllner Oliver (Hrsg.) HdM 2013, S. 293

gen, eine räumliche Erweiterung des Klangbildes sehr schwierig. Im folgenden Kapitel werden die wichtigsten dieser Mehrkanalanordnungen vorgestellt. Sie sollen später mit Hilfe der Binauraltechnik für eine mobile Verwendung synthetisiert werden.

4. Mehrkanalstereophonie

Bis in die 1950er Jahre war die Monophonie (von griechisch *mono* = „eins“ und *phonos* = „Klang“) der Standard zur Wiedergabe akustischer Signale über Lautsprecher. Ihr größter Nachteil besteht in der nicht vorhandenen räumlichen Abbildung von Schallquellen. Da nur mit einem Lautsprecher gearbeitet wird, ist es mit diesem Verfahren nicht möglich eine räumliche Aufteilung einzelner Signale zu erreichen. Ein undifferenziertes Klangbild ist die Folge; es kann keine getrennte Lokalisation einzelner Instrumente innerhalb einer Mischung erfolgen. Lediglich eine Tiefenstaffelung durch natürliche oder künstliche Hallerscheinungen kann erreicht werden. Berechtigterweise wurde sie später von der Zweikanalstereophonie abgelöst.

Genau wie der Begriff Monophonie stammt auch das Wort Stereophonie aus dem griechischen und setzt sich aus *stereo* = „körperlich, plastisch“ und *phonos* = „Klang“ zusammen⁴⁰. Man könnte, frei übersetzt, auch von „Raumklang“ sprechen. Dabei enthält der Begriff entgegen dem üblichen Gebrauch keine Informationen über die Anzahl der verwendeten Kanäle, weshalb es eigentlich falsch ist, ihn ausschließlich im Kontext des zweikanaligen Wiedergabeverfahrens zu verwenden. Stereophonie ist also ein Oberbegriff für alle mehrkanaligen Verfahren.

Es soll an dieser Stelle außerdem kurz erwähnt werden, dass unbedingt eine Unterscheidung zwischen Lautsprecherstereophonie und Kopfhörerstereophonie erfolgen muss (siehe Kapitel 5.1). Folgende Ausführungen behandeln ausschließlich die Lautsprecherstereophonie.

4.1 Zweikanalstereophonie

Der Grundgedanke dieses Wiedergabeverfahrens basiert auf der Abbildung von Schallquellen mit Hilfe von zwei Lautsprechern. Diese werden in einem bestimmten Abstand zueinander, der sogenannten Basisbreite aufgestellt. Die Basisbreite kann innerhalb einer gewissen Toleranz variiert werden und hängt primär vom Abstand des Hörers zur Lautsprecherachse ab⁴¹. Bei korrekter Aufstellung bildet der Hörer mit den beiden Lautsprechern ein gleichschenkliges Dreieck mit einem Öffnungswinkel von 60°, also jeweils 30° von der Mittelachse zum linken und rechten Lautsprecher.

⁴⁰ Vgl. Birkner 2002, S.13

⁴¹ Vgl. Weinzierl (Hrsg.) 2008, S. 611

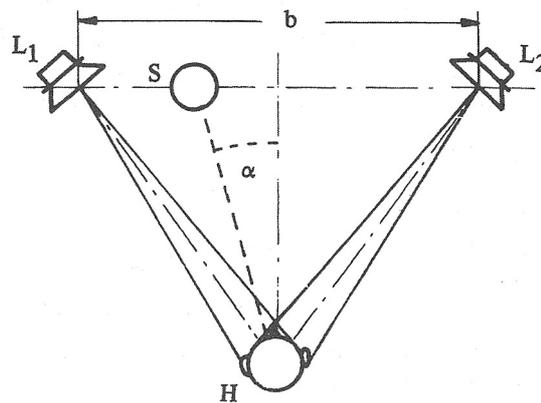


Abbildung 10: Stereodreieck mit Lautsprechern L1 und L2, Hörer H, Phantomschallquelle S, Basisbreite b und Lokalisationswinkel α

Diese Anordnung wird als Stereodreieck bezeichnet, an dessen Spitze sich der „Sweet-Spot“ befindet. Nur an dieser Position herrschen optimale Abhörbedingungen. Die Verteilung der einzelnen Signale auf der Lautsprecherachse basiert auf der Summenlokalisierung und kann bei der Aufnahme analog zur Funktionsweise des Gehörs sowohl über Laufzeitdifferenzen, Pegeldifferenzen oder deren Kombination (Äquivalenzstereophonie) erreicht werden. Die Wahl des Verfahrens hat allerdings Auswirkungen auf das spätere Klangbild. So liefert die Laufzeitstereophonie eine bessere räumliche Tiefenstaffelung bei geringerer Lokalisationschärfe, umgekehrt gelingt mit der Intensitätsstereophonie eine schärfere Trennung einzelner Schallquellen bei weniger ausgeprägtem Raumeindruck.

Die Positionierung von Schallquellen zwischen den beiden Lautsprechern wird über einen psychoakustischen Effekt ermöglicht. Strahlen beide Lautsprecher das gleiche Signal ohne zeitliche Verzögerung oder Pegelunterschiede ab, so erreicht die gleiche Information (bei gerader Kopfhaltung) beide Ohren, was beim natürlichen Hören einer frontalen Schallquelle in Richtung der Blickachse gleicht. Das Gehirn fehlinterpretiert die Situation und es entsteht eine sogenannte Phantomschallquelle in der Mitte der Lautsprecherachse. Der Entstehungsort der Phantomschallquelle kann durch zeitliche Verzögerungen oder Pegeldifferenzen zwischen linkem und rechtem Kanal entlang der Lautsprecherachse verschoben werden. Durch eine Automation der entsprechenden Parameter werden Bewegungen innerhalb des Stereopanoramas möglich, was speziell im Bildbereich von großer Bedeutung ist. Leider sind bewegte wie unbewegte Phantomschallquellen nicht ortsstabil, d.h. beim verlassen des Sweet-Spot verschiebt sich die Phantom-

schallquelle nach den Regeln des Präzedenz-Effektes⁴².

Die Lautsprecherstereophonie ist auch heute noch das meist verbreitete Mehrkanalverfahren. Durch unterschiedliche Aufnahmetechniken, auf die in dieser Arbeit wegen des begrenzten Umfangs nicht näher eingegangen werden kann⁴³, wird eine zweidimensionale Darstellung von Schallquellen im Raum erreicht. Dieser Vorteil gegenüber der Monophonie stellt sich jedoch im weiteren Verlauf der technischen Entwicklungen als Nachteil im Vergleich zu neueren Verfahren heraus. Das Bestreben nach größerer Räumlichkeit mündete in der Erweiterung der Zweikanalstereophonie um eine weitere Dimension.

4.2 Surround

Das Surroundverfahren wurde nach einer seiner wichtigsten Eigenschaften benannt, dem „umgebenden“ Klang. Sein Vorgänger war die Quadrophonie, welche das etablierte zweikanalige Verfahren um zwei zusätzliche, rückwärtige Lautsprecher und damit um eine räumliche Dimension erweiterte, sich aber beim Endanwender nie wirklich durchsetzen konnte. Auch beim Surroundstandard wird das Schallfeld durch eine steigende Anzahl von Lautsprechern erweitert, um räumliche Tiefe darstellen zu können. Die neu gewonnene Räumlichkeit bringt zum Einen viele gestalterische Möglichkeiten mit sich, sie ist zum Anderen aber auch immer mit zusätzlicher Hardware und steigenden Kosten verbunden. Seinen Ursprung hat das Surroundverfahren im Kinobereich. Um die Kompatibilität zu mehrkanaligen Heiminstallationen gewährleisten zu können, wurde der 3/2-Standard 1994 von der International Telecommunications Union in der Empfehlung ITU-R BS.775-1 festgelegt⁴⁴. Mit drei Frontkanälen und zwei rückwärtigen Surroundkanälen sollte ein Kompromiss zwischen räumlichem Klang und Praxis-tauglichkeit erreicht werden. Das Ergebnis war der ITU-Abhörkreis, an dem sich die meisten mehrkanaligen Wiedergabeverfahren orientieren.

42 Siehe Kapitel 2.2.1.2

43 Informationen finden Sie u.A. bei Dickreiter 2008 oder Webers 2007

44 Vgl. Weinzierl (Hrsg.) 2008, S. 642

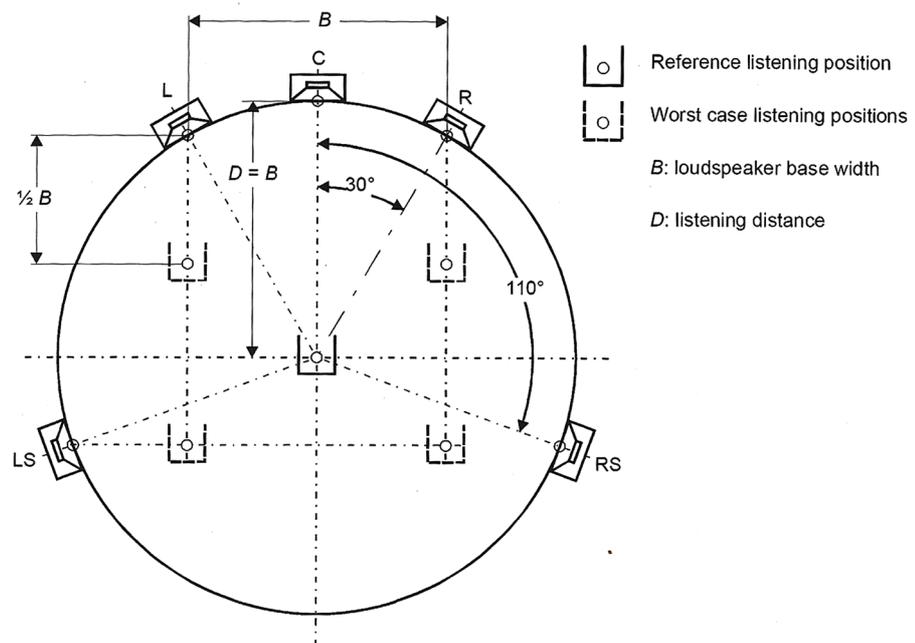


Abbildung 11: Lautsprecheranordnung nach ITU-R BS.775-1

Die Standardanordnung baut auf dem Stereodreieck der Zweikanalwiedergabe auf. Es wurde um einen zusätzlichen Center-Lautsprecher in der Mitte erweitert, der anstelle einer Phantomschallquelle für bessere Lokalisation und weniger Klangfärbung sorgt. Die beiden rückwärtigen Lautsprecher Left-Surround und Right-Surround werden in einem Winkel von $\pm 110^\circ$ zur Mittelachse aufgestellt und geben dem Klangbild die gewünschte Räumlichkeit. Alle Lautsprecher sind dabei auf den Hörer ausgerichtet. Bei der am weitesten verbreiteten 5.1-Anordnung wird der Abhörkreis mit einem auf 120 Hz bandbegrenzten LFE-Kanal (Low Frequency Effect) ergänzt⁴⁵, der für zusätzliche Tieftonwiedergabe sorgt. Die Bezeichnung 5.1 steht für fünf diskrete, und einen bandbegrenzten Kanal. Sie ist auf alle Wiedergabeverfahren anwendbar; so spricht man bei der Zweikanalstereophonie auch von 2.0 Stereophonie oder bei der Quadrophonie von 4.0 Stereophonie. Die Bezeichnung 3/2 hingegen beschreibt die räumliche Anordnung der Lautsprecher. Ein großer Vorteil dieser Surroundanordnung ist ihre Abwärtskompatibilität zu Wiedergabeverfahren mit weniger Kanälen. Wie oben schon kurz angedeutet, gibt es diverse Erweiterungen des Formates wie 7.1- oder 9.1-Anordnungen. Jedoch sind alle weiterhin auf die Wiedergabe in der Horizontalebene beschränkt. Das folgende Verfahren versucht diese Grenze zu durchbrechen.

⁴⁵ Vgl. Weinzierl (Hrsg.) 2008, S. 644

4.3 Auro 3D

Bei dem 2006 von Wilfried van Baelen in Paris vorgestellten Verfahren handelt es sich um eine Erweiterung des 3/2 ITU-Standards, das durch vier zusätzliche Lautsprecher über dem Hörer innerhalb bestimmter Grenzen ein dreidimensionales Klangbild erzeugt⁴⁶. Damit steht laut Theile „[...] eine weit größere Palette an räumlichen Effekten und eine natürlichere räumliche Wiedergabe in Bezug auf Direktschall, frühe und späte Reflexionen, Hall, Diffusschall und Atmo⁴⁷“ zur Verfügung. Beispielsweise kann durch die Wiedergabe früher Reflexionen in der Elevationsebene der räumliche Eindruck im Gegensatz zu horizontal beschränkten Verfahren stark verbessert werden. Die Anordnung der vier Höhenlautsprecher entspricht, wie im folgenden Schaubild zu sehen, der Quadrophonie:

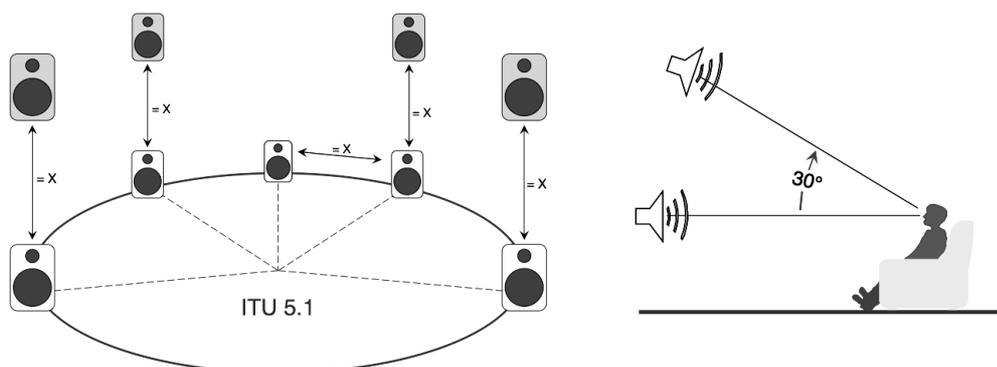


Abbildung 12: Basiskonfiguration "Auro 3D 9.1", kompatibel mit ITU-R BS.775-1. Rechts: Elevationswinkel der Höhenlautsprecher

Zur Abbildung der Schallquellen nutzt Auro 3D genau wie alle anderen Mehrkanalverfahren Phantomschallquellen. Allerdings bestehen erhebliche Einschränkungen zwischen den Lautsprechern der beiden Ebenen. Hier können praktisch keine Phantomschallquellen gebildet werden⁴⁸. Die Verbreitung von Auro 3D beschränkt sich sehr stark auf die Filmbranche, was durch den immensen Hardware- und damit verbundenen Kostenaufwand begründet werden kann. Eine Investition lohnt sich allerdings oft in Kombination mit dreidimensionalem Bildmaterial; hier

46 Fachartikel: Die dritte Dimension für Lautsprecher-Stereophonie, Theile und Wittek 2011,

http://www.hauptmikrofon.de/doc/Auro3D_D_0311.pdf, S. 2 ff, Stand: 07.11.2013

47 Ebd, S. 1

48 Ebd, S. 3

können beeindruckende Ergebnisse erzielt werden. Verständlicherweise ist das Verfahren in der Heimanwendung praktisch nicht existent.

4.4 Bedeutung für die Musikindustrie

Nachdem die Quadrophonie mangels geeigneter Medien niemals richtig Fuß fassen konnte, kamen die ersten Surroundproduktionen 1992 auf den Markt⁴⁹. Heutige Trägermedien wie die Super Audio CD oder die DVD-Audio sind für die Wiedergabe von mehreren diskreten Kanälen konzipiert worden. Leider führen sie nach wie vor ein Nischendasein und spielen in den Absatzstatistiken der Branche fast keine Rolle⁵⁰. Das mag nicht zuletzt daran liegen, dass für die Wiedergabe spezielle Abspielgeräte benötigt werden, die im Vergleich zu herkömmlichen Modellen recht teuer sind. Dennoch stünde einer Verbreitung von Mehrkanalproduktionen mit mehr als zwei Kanälen technisch nichts im Wege, doch noch immer ist die Zweikanalstereophonie das dominierende Format in der Musikindustrie. Die Hauptgründe dafür sind aus Sicht der Konsumenten die hohen Kosten für zusätzliche Hardware und Platzmangel im heimischen Wohnzimmer. So ist es kaum verwunderlich, dass laut einer Umfrage lediglich 15 % der deutschen Haushalte über eine Surroundinstallation verfügen. Anordnungen mit noch mehr Kanälen, wie 7.1- oder gar 9.1-Systeme, sind so gut wie nicht vorzufinden. Meist werden Surroundsysteme in Kombination mit Videogeräten als Heimkinoanlage verwendet. Dies zeigt auch, dass Mehrkanalproduktionen in der Filmindustrie schon komplett etabliert sind. Sie nimmt hier eine Vorreiterrolle ein, da gerade in Kombination mit stereoskopischem Bildmaterial ein räumlicher Klang zu einer größeren Immersion des Zuschauers beiträgt. In der Musikindustrie hingegen wird nach wie vor hauptsächlich für die zweikanalige Wiedergabe produziert. Mehrkanalaufnahmen findet man vorwiegend im Klassikbereich, da hier die Konsumenten als ausgesprochene Musikliebhaber auch vor höheren Kosten und größerem Aufwand nicht zurückschrecken. Viele Studios wären technisch durchaus zu Surroundproduktionen in der Lage, oftmals sind sie aber aus wirtschaftlichen Gründen nicht lohnenswert. Ein weiteres Argument gegen Mehrkanalproduktionen ist die zunehmende Orientierung des Musikmarktes in Richtung Mobilität. Das mobile Wachstum generiert inzwi-

49 Fachartikel: Surround – Das Format der Zukunft? Seidl 2010, http://www.hicktown-records.de/High_End_Tonstudio_Bayern/Know_How_files/Hicktown%20Records_Fachartikel_Surround%20Sound_Seidl.pdf, Stand: 07.11.2013

50 Jahreswirtschaftsbericht 2012 Bundesverband Musikindustrie, S. 17

schen einen beachtlichen Teil der Einnahmen der Musikindustrie⁵¹, doch leider ist die mobile Wiedergabe von mehr als zwei Kanälen nach wie vor problematisch. Hier versucht die vorliegende Arbeit anzuknüpfen und eine Möglichkeit aufzuzeigen, wie die Vorteile des Mehrkanalton zukünftig auch mobil genutzt werden können. Eine wesentliche Voraussetzung ist dabei die Wiedergabe über tragbare Kopfhörer.

51 Siehe Kapitel 3

5. Der Kopfhörer als Schallwandler

Kopfhörer gibt es in den verschiedensten Ausführungen. Die gängigste Bauform ist der Doppelkopfhörer, bestehend aus jeweils einem elektroakustischen Wandler für jedes Ohr⁵². Ein Großteil der Kopfhörer arbeitet wie die meisten Lautsprecher nach dem elektrodynamischen Wandlerprinzip, man findet aber auch elektromagnetische oder elektrostatische Typen. Sie sind allerdings recht selten, da aufgrund des verbauten Kondensators eine hohe Vorspannung bereitgestellt werden muss. Weiterhin unterscheidet man, abhängig von der Dämpfung des Umgebungsschalls, zwischen offenen, halboffenen und geschlossenen Bauformen. Der Frequenzgang eines Kopfhörers kann für einen spezifischen Einsatzzweck entzerrt sein; so wird beispielsweise für die Wiedergabe von kopfbezogener Stereophonie ein diffusfeldentzerrter Kopfhörer empfohlen⁵³. Im professionellen Bereich wird der Kopfhörer aufgrund einiger Nachteile nur sehr selten eingesetzt. Gerade in Situationen, die durch laute Umgebungsgeräusche oder Platzmangel ein korrektes Abhören über Lautsprecher nicht erlauben, kann er aber ein unverzichtbares Werkzeug sein. Im Consumerbereich genießt der Kopfhörer mittlerweile einen hohen Stellenwert und erfreut sich, bedingt durch die zunehmende Verbreitung mobiler Endgeräte, immer größerer Beliebtheit.

5.1 Abgrenzung von Kopfhörer- und Lautsprecherstereophonie

An dieser Stelle soll eine grundlegende Unterscheidung von Kopfhörerstereophonie und Lautsprecherstereophonie vorgenommen werden. Die eindeutige Trennung beider Systeme ist für das Verständnis dieser Arbeit essentiell und wird leider oft nicht beachtet. Der wesentliche Unterschied liegt in den für das jeweilige System benötigten Signalen. Für die Wiedergabe mit Kopfhörern muss das Signal über spektrale Differenzen, also frequenzabhängige Pegelunterschiede verfügen, die unser Gehirn wie beim räumlichen Hören auswerten kann⁵⁴. Nur dann entsteht ein natürlicher, räumlicher Höreindruck. Da bei der Kopfhörerwiedergabe die Funktion unserer Ohrmuschel ausgeschaltet ist, können die Signale nur durch eine geeignete Aufnahme-

⁵² Vgl. Dickreiter 2008, S. 178

⁵³ Ebd.

⁵⁴ Vgl. Sengpiel: Stereophonie für Lautsprecher und Stereophonie für Kopfhörer 2,

<http://www.sengpielaudio.com/StereoFuerLautsprUndKopfhoerer2.pdf>, Stand: 08.11.2013

technik mit den spektralen Richtungsinformationen versehen werden. Hier ist der in Kapitel 6.3 beschriebene Kunstkopf mit seinen Ohrmuschelnachbildungen aus Silikon prädestiniert. Bei der Wiedergabe solcher Kopfhörersignale über Lautsprecher treten starke Verfälschungen der Klangfarbe durch eine Überlagerung der spektralen Informationen auf, die aber bei der Kopfhörerwiedergabe durch die akustische Trennung der Ohren ausgeschlossen werden kann. Deshalb sollten in Signalen für die Lautsprecherwiedergabe möglichst keine spektralen Differenzen enthalten sein. Hier werden frequenzneutrale Pegelunterschiede benötigt⁵⁵. Im überlagerten Schallfeld erreichen die einzelnen Signale der Lautsprecher stets beide Ohren, wodurch eine Lokalisation entsprechend den natürlichen Hörgewohnheiten möglich ist und den bei der Aufnahme beabsichtigten Höreindruck entstehen lässt.

5.2 Stärken und Schwächen des Kopfhörers

Einer der großen Vorteile des Kopfhörers sind seine sehr guten elektroakustischen Eigenschaften bei vergleichsweise geringem Aufwand. Wegen des kurzen Abstands zum Ohr kann er als Einwegesystem mit nur einer Membran gebaut werden, wodurch Phasenverschiebungen, wie sie bei Lautsprechern auftreten, vermieden werden. Außerdem erreicht man mit einer kleinen Membran eine sehr gute Impulstreue, kann aber wegen des kurzen Abstandes noch hohe Lautstärken bei geringer elektrischer Leistung erzielen⁵⁶. Diese Eigenschaften äußern sich in einem großen darstellbaren Dynamikbereich. Eine weitere Stärke des Kopfhörers ist seine bauartbedingte Abschottung von Umgebungsgeräuschen, speziell bei geschlossenen Modellen ist diese sehr ausgeprägt und reicht von 5 dB bei tiefen Frequenzen bis zu 40 dB bei 10 kHz⁵⁷. Manche Modelle verwenden zur völligen Abschottung ein eingebautes Noise-Cancelling-System, bei dem die durch ein äußeres Mikrofon gemessenen Umgebungsgeräusche invertiert wiedergegeben, und so unterdrückt werden. Durch die Unabhängigkeit von den Eigenschaften des Abhörortes hängt die Wiedergabequalität allein vom verwendeten Kopfhörer ab und ist keinen räumlichen Einflüssen unterworfen. So kann z.B. für mehrere Hörer das exakt gleiche Hörerlebnis bereitgestellt werden, was bei einer Lautsprecheranordnung durch den begrenzten Abhörbereich sehr schwierig zu verwirklichen ist. Umgekehrt dringt durch einen Kopfhörer relativ wenig Schall

55 Vgl. Sengpiel: Stereophonie für Lautsprecher und Stereophonie für Kopfhörer 2

56 Vgl. Dickreiter 2008, S. 177

57 Ebd.

nach außen, der in der Umgebung als störend empfunden werden könnte.

Die oben erwähnte Unabhängigkeit des Kopfhörers von der Abhörsituation entpuppt sich bei genauerer Betrachtung allerdings nicht nur als Vorteil, sondern wird bei der Wiedergabe von Lautsprecher-mischungen, wie sie heute hauptsächlich produziert werden, zum Hauptproblem. Die Abschottung der Ohrmuschel durch das Kopfhörergehäuse bewirkt eine akustische Trennung von linkem und rechtem Ohr; es findet kein interaurales Übersprechen mehr statt. Die Folge ist ein unnatürlicher Höreindruck durch das Fehlen eines überlagerten Schallfeldes, wie wir es für das räumliche Hören benötigen. Die in Kapitel 2.2 beschriebenen Filtereigenschaften der Ohrmuschel werden beim Aufsetzen des Kopfhörers außer Kraft gesetzt. Dadurch kommt es, ähnlich wie bei der Phantomschallquellenbildung in der Lautsprecherstereophonie, zu einer „Im-Kopf-Lokalisation“ von Schallquellen auf der Achse zwischen den Ohren⁵⁸. Werden beide Ohren über Kopfhörer zeitgleich mit dem gleichen Signal versorgt, nehmen wir die Schallquelle in der Mitte unseres Kopfes wahr. Die Signale enthalten keine spektralen Richtungsinformationen, die unser Gehirn auswerten könnte (es sei denn es handelt sich um kopfbezogene Signale, wie sie z.B. ein Kunstkopf erzeugt). Eine korrekte Abbildung des Stereopanoramas ist nicht mehr möglich und jegliche Räumlichkeit geht verloren. Interessanterweise empfinden jüngere Menschen diese Effekte als weniger störend, was vermutlich maßgeblich mit dem gestiegenen Musikkonsum über Kopfhörer gerade in dieser Altersklasse zusammenhängt. Weitere Probleme tragen zusätzlich zur Inkompatibilität von Lautsprecherstereophonie und Kopfhörerwiedergabe bei. Hört man eine auf Lautsprechern angefertigte Mischung über Kopfhörer ab, wird der Höreindruck erheblich verfälscht. Neben der gerade angesprochenen Im-Kopf-Lokalisation sind eine überbreite Stereobasis und ein statisches Klangbild häufige Störeffekte. Durch das Tragen des Kopfhörers wandert das Klangbild im Gegensatz zur Lautsprecherwiedergabe entsprechend unserer Kopfbewegung mit. Dieses Mitwandern des Klangbildes ausgenommen, treten alle genannten Nachteile bei der Verwendung von kopfbezogenen Signalen **nicht** auf.

58 Vgl. Blauert 1974, S. 97

5.3 Aktuelle Marktsituation

Mit der steigenden Beliebtheit mobiler Geräte, wie mp3-Player, Smartphones und Tablets, steigt auch der Absatz von Kopfhörern deutlich. Da sich die intern verbauten Lautsprecher der Geräte (falls überhaupt vorhanden) nur sehr bedingt für das Hören von Musik eignen und unterwegs keine andere Möglichkeit besteht, ist der Kopfhörer das ideale Werkzeug für den mobilen Musikgenuss. Er gibt die Musik für den Hörer individuell wieder, ohne dabei das Umfeld zu stören. Umgekehrt schützen die meisten Kopfhörermodelle durch eine geschlossene Bauform auch vor Störgeräuschen von außen, einige verwenden darüber hinaus die erwähnte Geräuschunterdrückung. So ist es kaum verwunderlich, dass parallel zur Entwicklung des Mobilmarktes auch ein deutlicher Anstieg beim Absatz von Kopfhörern zu beobachten ist. Wurden im Jahr 2011 laut dem Consumer Electronics Marktindex der Gesellschaft für Unterhaltungs- und Kommunikationselektronik deutschlandweit noch knapp zehn Millionen Kopfhörer verkauft, so waren es 2012 schon über elf Millionen Stück. Der Umsatz steigerte sich dabei um 30%⁵⁹. Für das Jahr 2013 wird noch einmal mit einer deutlichen Steigerung des Absatzes gerechnet. Der starke Umsatzzuwachs lässt darauf schließen, dass mittlerweile immer mehr Menschen zu qualitativ hochwertigeren Kopfhörern greifen. Diese bestechen durch ein sehr gutes Preis-Leistungs-Verhältnis. Darüber hinaus ist der Kopfhörer heute schon fast zu einem modischen Accessoire geworden. Viele Jugendliche drücken sich stark über ihre äußere Erscheinung aus, was sich einige Hersteller zu Nutze machen und Kopfhörer in allen Farbvarianten und Formen auf den Markt bringen. Leider steht dabei oftmals die Optik und nicht die Klangqualität im Vordergrund.

⁵⁹ Gfu CEMIX 2012,

http://www.gfu.de/srv/easyedit/ts_1361800251000/page:home/download/marktstudien/sl_1361547008349/args.link01/de_CEMIX%20Q1-Q4%202012.pdf, Stand: 28.10.2013

6. Binauraltechnik

6.1 Grundlagen

Die Binauraltechnik ist neben der in Kapitel 4 vorgestellten Lautsprecherstereophonie und der Wellenfeldsynthese (auf die in dieser Arbeit nicht näher eingegangen werden soll) ein weiteres Prinzip zur Aufnahme und Übertragung von Schallereignissen. Sie macht sich dabei die Erkenntnisse des räumlichen Hörens zunutze. Der Name Binauraltechnik mag auf den ersten Blick verwirrend erscheinen, denn binaural bedeutet nichts anderes als „zweiohrig“, was dem natürlichen Hören des Menschen entspricht (evtl. vorhandene Hörschäden ausgeschlossen). Der wesentliche Unterschied zu anderen Prinzipien liegt in der Verwendung von Ohrsignalen, wie sie dem Leser aus Kapitel 2 bekannt sind. Folgende Annahme bildet die Grundlage der Binauraltechnik: Gelingt es, das Schallsignal auf dem Weg von der Quelle bis in unser Gehirn an einem Punkt abzugreifen, an dem es alle Richtungsinformationen enthält und nicht mehr durch äußere Einflüsse verändert werden kann, und es später an genau diesem Punkt originalgetreu wiederzugeben, so entsteht beide Male der exakt gleiche Höreindruck. Der gesuchte Punkt befindet sich in der Praxis direkt vor dem Trommelfell; das Schallsignal wird hier als Ohrsignal bezeichnet. Es wird mit Hilfe von kleinen Mikrofonen im Ohr oder mit einem Kunstkopf aufgenommen und später über Kopfhörer wiedergegeben. Die Schallübertragung auf dem beschriebenen Weg gleicht einem zeitinvarianten, linearen System und kann im Zeitbereich durch eine Impulsantwort und im Frequenzbereich durch eine Übertragungsfunktion beschrieben werden⁶⁰.

Im Hinblick auf eine originalgetreue Übertragung liefert die Binauraltechnik die besten Ergebnisse. Nicht nur Schalleinfallrichtungen und Entfernungen werden sehr gut abgebildet, auch ein hohes Maß an Umhüllung des Hörers und ein originalgetreuer Raumeindruck können erreicht werden. Tabelle 1 stellt die verschiedenen Verfahren und ihre Möglichkeiten einander gegenüber. Ein sehr entscheidender Vorteil des binauralen Prinzips ist außerdem, dass es sich trotz der guten Abbildungseigenschaften mit zwei Kanälen begnügt. So können bei großem Mehrwert die Infrastruktur und der bekannte Umgang mit zweikanaligem Audiomaterial beibehalten werden.

⁶⁰ Vgl. Weinzierl (Hrsg.) 2008, S. 671

Attribute der Richtungs- und Raumabbildung	2/0-Stereophonie	3/3-Stereophonie	Kunstkopfstereophonie ohne Headtracking
horizontale Richtungen	+30° bis -30°	+30° bis -30°, Einschränkungen für andere Richtungen	alle Richtungen, Einschränkungen für andere Richtungen
Elevation	nicht möglich	eingeschränkt möglich	möglich
kopfnahе Entfernungen	nicht möglich	nicht möglich	möglich
Entfernungen, Tiefe	simuliert	möglich	möglich
interaktive räumliche Perspektive	nicht möglich	nicht möglich	nicht möglich
Raumeindruck	simuliert	möglich	möglich
Umhüllung	nicht möglich	eingeschränkt möglich	möglich

Tabelle 1: Attribute der räumlichen Abbildung und prinzipielle Möglichkeiten der Übertragungsverfahren

Binaurale Aufnahmen sind heute hauptsächlich im Hörspielbereich und bei akustischen Messvorgängen von Bedeutung. Das erste binaural produzierte Hörspiel „Demolition“ wurde im Rahmen der Funkausstellung 1973 in Berlin ausgestrahlt⁶¹. Leider konnte sich die Technik aufgrund einiger Probleme, die später näher besprochen werden, in der Musikproduktion nie richtig etablieren.

6.2 HRTF

Die HRTF (engl.: „Head Related Transfer Function“), zu deutsch Außenohrübertragungsfunktion oder auch kopfbezogene Übertragungsfunktion beschreibt die linearen Verzerrungen des Schallsignals am Eingang des Gehörgangs, die durch das Zusammenwirken der individuellen anatomischen Merkmale beim Hören entstehen⁶². Sie „[...] ist das Ergebnis von akustischer Abschattung, Beugung, Verzögerung, Resonanzen und Reflexionen durch Torso, Schulter, Kopf, Ohrmuschel (pinnae), den Eingang in den Ohrkanal (cavum conchae) und den Ohrkanal selbst.“⁶³ Dabei gibt es zum Teil große Unterschiede zwischen den Funktionen verschiedener Personen,

⁶¹ Vgl. Dickreiter 2008, S. 298

⁶² Vgl. Blauert 1974, S. 62

⁶³ Weinzierl (Hrsg.) 2008, S. 586

die auf die natürliche Varianz der beteiligten Körperteile zurückzuführen sind. Die HRTF ist dadurch von Mensch zu Mensch verschieden und gleicht in ihrer Einzigartigkeit einem Fingerabdruck. Spektrale Veränderungen, die von unseren eigenen Ohren zur Richtungsbestimmung verursacht werden, nehmen wir nicht als unnatürliche Klangfarbe wahr. Beim Hören mit einer fremden HRTF äußern sich die Differenzen aber deutlich in einer schlechten Lokalisation und Klangverfärbungen. Auf dieses und weitere Probleme wird später genauer eingegangen.

Wir wissen bereits, dass die entstehenden linearen Verzerrungen für das räumliche Hören von großer Bedeutung sind. Da der Schall eine für jeden Einfallswinkel spezifische Filterung erfährt, entsteht auch für jeden Schalleinfallswinkel eine andere Übertragungsfunktion⁶⁴. Abbildung 13 zeigt die HRTF-Kurven des linken Ohres verschiedener Personen bei unterschiedlichen Schalleinfallswinkeln und ihre bewerteten Mittelwerte (weiße Kurven).

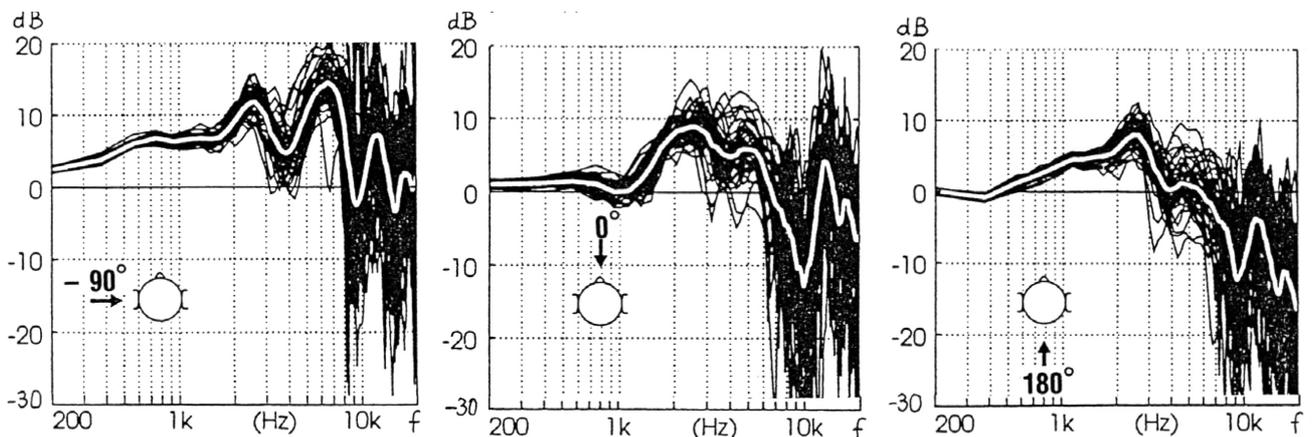


Abbildung 13: Vergleich der HRTF-Kurven des linken Ohres. Die weißen Kurven stellen die bewerteten Mittelwerte dar.

6.3 Kunstkopfstereophonie

Auch die Kunstkopfstereophonie macht sich die Außenohrübertragungsfunktion zunutze. Ein Kunstkopf ist ein dem menschlichen Kopf nachempfundenen Stereomikrofon. Dabei wird neben der meist grob abstrahierten Form des Kopfes besonderen Wert auf die Modellierung der Ohrmuscheln gelegt. Aus Silikon gefertigt, entsprechen sie der Durchschnittsform eines menschlichen Ohres. Anstatt des Trommelfells werden im künstlichen Gehörgang Druckempfangermikrofone mit Kugelcharakteristik verwendet. Die Richtcharakteristik sowie Pegel- und Laufzeitdiffe-

⁶⁴ Vgl. Sengpiel: Kopfbezogene Übertragungsfunktion HRTF,

<http://www.sengpielaudio.com/KopfbezogeneUebertragungsfunktionHRTF.pdf>, Stand: 12.11.2013

renzen entsprechen denen des menschlichen Ohres⁶⁵. Der Kunstkopf versucht, die Vorgänge des räumlichen Hörens so gut wie möglich nachzubilden und dadurch ein authentisches Hörerlebnis zu schaffen. Das mit ihm aufgenommene Schallfeld, welches im Aufnahmeraum an den Ohren des Hörers herrschen würde, wird später an dieser Stelle über einen Kopfhörer reproduziert und gibt dem Hörer das Gefühl, sich selbst im Aufnahmeraum zu befinden⁶⁶. Durch die Verallgemeinerung der anatomischen Merkmale des Kunstkopfes erzeugt er natürlich auch entsprechende Ohrsignale mit einer gemittelten HRTF. Die individuelle Wiedergabequalität hängt deshalb stark vom Grad der Übereinstimmung der eigenen Übertragungsfunktion mit der des Kunstkopfes ab. Mittlerweile gibt es viele unterschiedliche Modelle, von der ursprünglichen Form wie der KU100 der Firma Neumann, über solche mit Schultern und Torso, oder spezielle Kunstköpfe mit angepassten Ohrmuschelmodellierungen für den asiatischen Raum.



Abbildung 14: KU100 der Firma
Neumann

6.3.1 Probleme

Da der Kunstkopf wie schon erwähnt auf anatomischen Durchschnittswerten beruht, kommt es bei Personen mit stark abweichenden Kopf- und Ohrmuschelformen oft zu einer unzureichenden Ortung bei frontaler Schalleinfallrichtung. Die Schallquelle wird stattdessen hinten oder über dem Kopf wahrgenommen. Beim natürlichen Hören werden frontale Schallereignisse häufig durch unbewusste Peilbewegungen des Kopfes analysiert. Die dadurch entstehenden minimalen Änderungen der Ohrsignale werden vom Gehirn zur Lokalisation ausgewertet und stabili-

⁶⁵ Vgl. Dickreiter 2008, S. 296

⁶⁶ Ebd.

sieren das Hörereignis⁶⁷. Leider kann die statische Aufstellung des Kunstkopfes keine Änderungen der Ohrsignale bei unbewegten Schallquellen liefern. Außerdem entsteht durch diesen Umstand ein kopfbezogenes Klangbild.

Wie im Kapitel zur Abgrenzung von Lautsprecher- und Kopfhörerstereophonie schon deutlich wurde, ist das Kunstkopffverfahren durch die Verwendung von Ohrsignalen nicht für die Lautsprecherwiedergabe geeignet. Aus diesem Grund findet das Verfahren auch keine Anwendung im Rundfunk. Es existieren zwar Methoden, die versuchen, die Aufnahmen durch eine Übersprechkompensation zwischen den Lautsprechern kompatibel zu machen, allerdings ist die Hörzone dabei auf einen sehr kleinen Bereich beschränkt und der Kopf muss absolut still gehalten werden⁶⁸. Für die praktische Anwendung sind sie damit untauglich.

Weiterhin ist schon bei der Aufnahme mit einem Kunstkopfsystem darauf zu achten, dass im Gegensatz zu herkömmlichen Mikrofonaufnahmen in der Postproduktion einige Einschränkungen bestehen. Wie bei selbst durchgeführten Aufnahmen festgestellt wurde, ist ein Schneiden nur dann möglich, wenn die Schallquellen ihre Position dauerhaft beibehalten. Andernfalls können später ungewollte Sprünge der Hörereignisse im Klangbild die Folge sein. Außerdem muss bei Lautstärkeanpassungen darauf geachtet werden, dass diese eine Veränderung der Entfernungswahrnehmung hervorrufen können (siehe Kapitel 2.2.2).

Ein interaktives Hörspiel, das mit einem Kunstkopf produziert wurde, finden sie unter www.verdacht.de⁶⁹.

67 Vgl. Weinzierl (Hrsg.) 2008, S. 676

68 Vgl. Sengpiel: Transaural Stereo – Kunstkopf-Stereophonie über Lautsprecher,
<http://www.sengpielaudio.com/TransauralStereo.pdf>, Stand: 13.11.2013

69 Stand: 14.02.2014

6.4 Akustische Raumsimulation

6.4.1 Exkurs: Das LTI-System

Ein LTI-System ist ein lineares, zeitinvariantes System (engl.: linear time-invariant system; LTI) mit dem sich viele Vorgänge in der Tonstudioteknik eindeutig beschreiben lassen. Es ist nach einer Messung der entscheidenden Parameter digital reproduzierbar und bildet die Grundlage für eine akustische Raumsimulation⁷⁰. Unter LTI-Systemen versteht man „[...] diejenigen Übertragungssysteme, die bei der Überführung eines gegebenen Zeitsignals $s(t)$ in eine Systemantwort $g(t)$ lediglich lineare Verzerrungen hervorrufen. Als linear gelten Verzerrungen, bei denen nur die bereits vorhandenen Komponenten des Spektrums in ihrer Amplitude oder in ihrer Phasenlage verändert werden.“⁷¹ Es werden vom System $h(t)$ also nur vorhandene Frequenzen beeinflusst und keine neuen Anteile z.B. in Form von Obertönen hinzugefügt.



Abbildung 15: LTI-System im Zeitbereich

6.4.2 Die binaurale Raumimpulsantwort BRIR

Die Schallübertragung kann von der Quelle bis an unser Trommelfell mit dem LTI-System im Frequenzbereich durch eine Übertragungsfunktion und im Zeitbereich durch eine Impulsantwort beschrieben werden⁷². Aufgrund der Komplexität einiger LTI-Systeme wie z.B. der HRTF ist eine mathematische Darstellung sehr schwierig bis unmöglich. Die einzige Möglichkeit das System zu reproduzieren ist eine Messung des Übertragungsverhaltens. Auf diese Weise erhält man eine sogenannte Impulsantwort, die das zeitliche Verhalten vollständig beschreibt. Mit ihr lassen sich

⁷⁰ Vgl. Tarnow und Keinath: Entwicklung und Erprobung einer Methode zur Integration binauraler Impulsantworten bei der Mischung von Mehrspurproduktionen 2009, http://www.eti.hfm-detmold.de/lehraktiv/diplomarbeiten/diplomarbeitenordner/DA_Keinath_Tarnow_2009.pdf/view, S. 12, Stand: 14.11.2013

⁷¹ Ebd.

⁷² Vgl. Weinzierl (Hrsg.) 2008, S. 671

die Übertragungseigenschaften des Systems auf jedes beliebige Signal anwenden⁷³.

Im Folgenden soll beispielhaft die Messung des Übertragungsverhaltens eines Raumes betrachtet werden, da diese in der Tonstudioteknik häufig zur Anwendung kommt. Bei der Messung wird einem oder auch mehreren Lautsprechern als Eingangssignal entweder ein kurzer Impuls oder ein sogenannter Sinus-Sweep zugeführt. Ein solcher Sweep durchläuft in einer wählbaren Zeitspanne bei gleichbleibender Amplitude den gesamten hörbaren Frequenzbereich von 20 Hz bis 20 kHz. Die Reaktion des Raumes auf den Impuls wird mit einem Messmikrofon eingefangen. Nach einer Trennung des Eingangssignals von der Raumimpulsantwort liegt diese in Reinform vor.

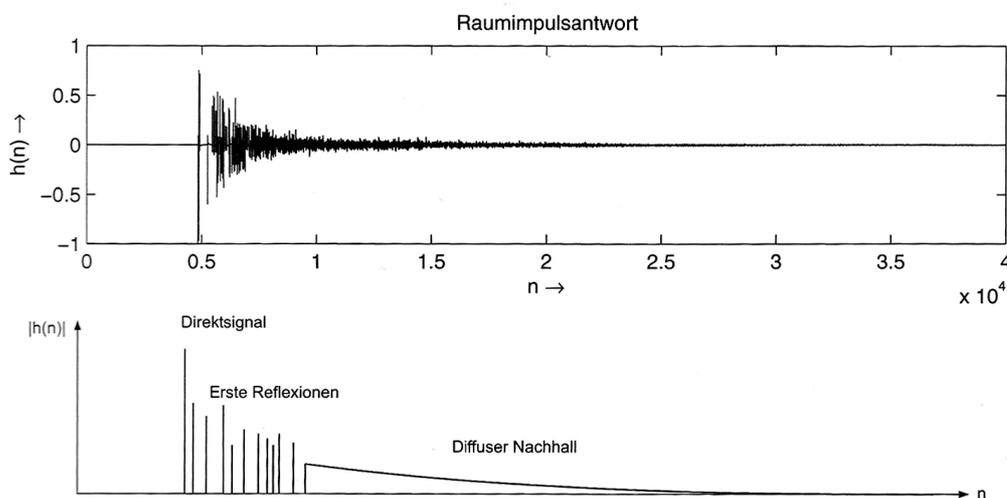


Abbildung 16: Raumimpulsantwort $h(n)$ und vereinfachte Aufteilung der Raumimpulsantwort in Direktsignal, erste Reflexionen und diffusen Nachhall (Darstellung als $|h(n)|$)

Zur Erstellung einer binauralen Raumimpulsantwort oder BRIR (engl.: binaural room impulse response) kann statt eines einfachen Mikrofons auch ein Kunstkopf zur Messung verwendet werden. Dabei verändert sich das Übertragungsverhalten des LTI-Systems durch die linearen Verzerrungen am Kunstkopf; es wird zusätzlich die Übertragungsfunktion des Kopfes mit aufgenommen. Auf diese Weise lassen sich später aus trockenen, weitgehend hallfreien Signalen auch kopfbezogene Ohrsignale für eine Kopfhörerwiedergabe generieren.

⁷³ Vgl. Tarnow und Keinath 2009, S. 13

6.4.3 Die personalisierte Raumimpulsantwort PRIR

Um bei der Verwendung von Impulsantworten eine optimale und fehlerfreie Wiedergabe zu gewährleisten, empfiehlt es sich, die Messung der Antwort nicht mit einem genormten Kunstkopf, sondern mit dem individuellen Kopf durchzuführen. Es werden hierfür kleine Messmikrofone zur Aufzeichnung der Impulsantwort in den Gehörgang der Messperson eingeführt. Dabei entsteht eine sogenannte personalisierte Raumimpulsantwort oder PRIR (engl.: personalised room impulse response), mit der sich die Probleme der Kunstkopfstereophonie weitgehend vermeiden lassen. Eine detailliertere Beschreibung zur PRIR-Messung folgt in Kapitel 7.

Zur Vermeidung einer Fehllokalisierung von Schallquellen im Frontalbereich setzen einige Systeme auf einen Head-Tracker, der während der Wiedergabe mittels Triangulation die Kopfposition des Hörers bestimmt und die passende Impulsantwort zur Faltung aus einer Datenbank auswählt. So werden auch unterbewusste Peilbewegungen des Kopfes registriert und die Ohrsignale entsprechend angepasst, wodurch keine Vertauschung der Hörereignisrichtungen vorne und hinten mehr stattfindet. Allerdings müssen beim Einsatz des Head-Trackings auch mehrere Impulsantworten für die gleiche Hörposition aufgenommen werden, zwischen denen das System entsprechend interpolieren kann. Die hierfür benötigte Rechenleistung muss ebenfalls zur Verfügung stehen.

Es gibt allerdings auch einen entscheidenden Nachteil einer PRIR. Durch die Personalisierung der Impulsantwort mit der individuellen HRTF kann nur für diese eine Person eine optimale Wiedergabe erreicht werden. Für andere Personen hängt die Wiedergabequalität, wie auch beim Kunstkopfverfahren, von der Übereinstimmung der jeweils eigenen HRTF mit der gemessenen Übertragungsfunktion ab. Um ein Audiosignal schließlich mit der Impulsantwort zu verrechnen, bedient man sich der sogenannten Faltung.

6.4.4 Faltung

Bei dieser Rechenoperation werden die beiden Komponenten miteinander multipliziert. Jedoch ist die Faltung im Zeitbereich nur mit erheblichem rechnerischen Aufwand zu bewerkstelligen. Dabei wird jedes einzelne Sample des ersten Signals mit jedem Sample des zweiten Signals verrechnet, was eine immense Anzahl erforderlicher Rechenschritte nach sich zieht⁷⁴. Deshalb nutzt man die Eigenschaften des LTI-Systems und transformiert beide Zeitsignale mit Hilfe der schnellen Fouriertransformation (engl.: Fast Fourier Transformation, FFT) in die Spektraldarstellung. Der komplexe Übertragungsfaktor $H(f)$ repräsentiert das Übertragungsverhalten des Systems und lässt sich aus der gewonnen Impulsantwort $h(t)$ wie folgt berechnen⁷⁵:

$$H(f) = \int_{-\infty}^{\infty} h(t) e^{-j2\pi ft} dt$$

Durch die Transformation der Impulsantwort in den Frequenzbereich ist es möglich, jedes beliebige Signal schrittweise über die Multiplikation seines Eingangsspektrums mit dem gewonnenen Übertragungsfrequenzgang zu verrechnen. Man erhält das Ausgangssignal $G(f)$ des LTI-Systems in Frequenzdarstellung⁷⁶.

$$S(f) \cdot H(f) = G(f)$$

S(f): Eingangssignal in Frequenzdarstellung

H(f): Übertragungsfrequenzgang

G(f): Ausgangssignal in Frequenzdarstellung

Jetzt muss es noch durch eine inverse Fouriertransformation wieder in den Zeitbereich gewandelt werden.

$$g(t) = \int_{-\infty}^{\infty} G(f) e^{j2\pi ft} dt$$

Eine Faltung im Zeitbereich entspricht demnach einer Multiplikation im Frequenzbereich. Trotz des entstandenen Mehraufwands durch die Überführung der Signale in eine geeignete Spekt-

⁷⁴ Vgl. Tarnow und Keinath 2009, S. 12

⁷⁵ Ebd, S. 9

⁷⁶ Ebd, S. 11

raldarstellung lässt sich das Ausgangssignal auf diese Weise wesentlich schneller berechnen als mit der diskreten Faltung, weshalb auch von **schneller Faltung** gesprochen wird⁷⁷. Sie ist Voraussetzung für eine Berechnung des Ausgangssignals in Echtzeit.

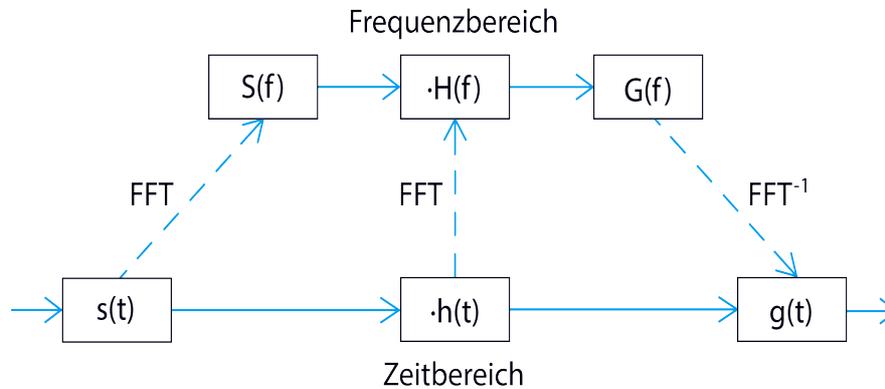


Abbildung 17: Schnelle Faltung

6.4.5 Zusammenfassung

Durch die Messung einer binauralen Impulsantwort und der anschließenden Faltung mit einem halbfreien Eingangssignal lässt sich also eine akustische Raumsimulation verwirklichen. Damit wäre ein entscheidender Schritt für die Wiedergabe von Mehrspurmischungen auf mobilen Endgeräten getan. Durch die Verwendung einer BRIR/PRIR können solche Mischungen, die ursprünglich für die Lautsprecherwiedergabe produziert wurden, in zwei diskrete Signale mit den für die Kopfhörerwiedergabe benötigten spektralen Informationen umgewandelt werden. Die Impulsantwort repräsentiert sämtliche Übertragungsparameter der Abhörumgebung, einschließlich der Wiedergabeanordnung zum Zeitpunkt der Messung, dem Verhalten des Raumes und den Wiedergabeeigenschaften der verwendeten Lautsprecher. Ein optionales Head-Tracking-System stabilisiert Hörereignisse in frontaler Richtung. Theoretisch ist es auf diese Weise möglich, Mehrkanalsysteme mit beliebig vielen Kanälen im Kopfhörer zu simulieren. Einzig die Rechenleistung des Computersystems auf dem die Faltung stattfindet, begrenzt die Anzahl der Kanäle.

⁷⁷ Vgl. Tarnow und Keinath 2009, S. 12

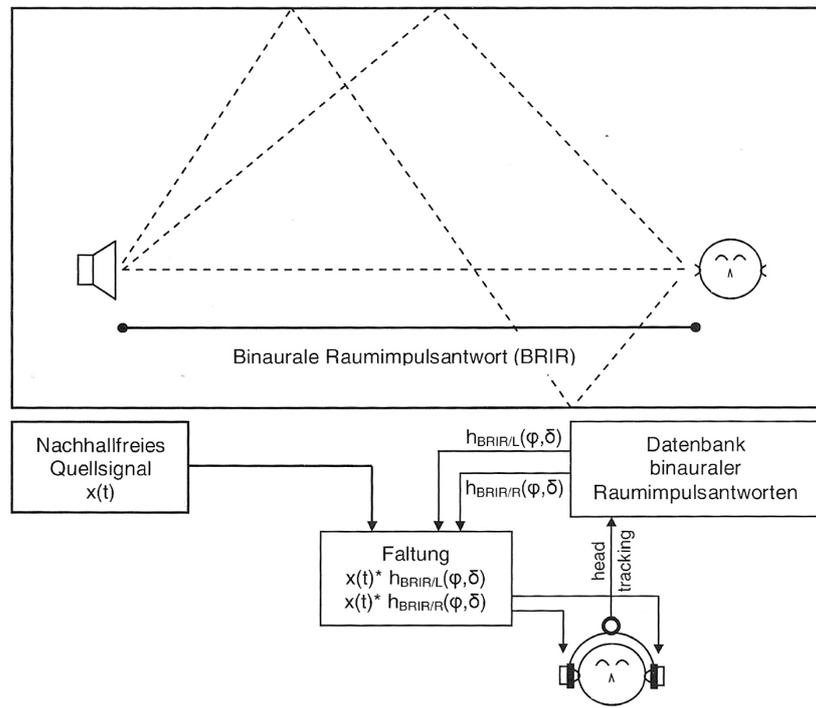


Abbildung 18: Zusammenfassung Binauraltechnik.

7. Messung einer PRIR mit dem Smyth Realiser A8

Der Realiser A8 der Firma Smyth Research ist ein Gerät zur Messung und Simulation einer Abhörumgebung via Kopfhörer. Er arbeitet dabei mit den im vorigen Kapitel beschriebenen individuellen Impulsantworten, was ihn von ähnlichen Geräten zur virtuellen Raumsimulation wie beispielsweise dem Headzone-System der Firma Beyerdynamic unterscheidet. Die Personalisierung der Parameter garantiert die beste erreichbare Abbildungsqualität des zu simulierenden Raumes, weshalb der Realiser A8 zur Erstellung der für den praktischen Teil dieser Arbeit benötigten binauralen Impulsantworten ausgewählt wurde. Im Folgenden soll der Realiser A8 kurz vorgestellt und der Vorgang zur Messung der Impulsantworten beschrieben werden. Detaillierte Informationen zum Gerät finden Sie bei Phillip Reineboth⁷⁸.

7.1 Das Gerät im Überblick



Der Realiser A8 erscheint auf den ersten Blick recht unspektakulär. Auf der Vorderseite befinden sich ein kleines Display, zehn LEDs zur Anzeige der Kanalkonfiguration, vier weitere Status-LEDs, ein SD-Karten Slot zur externen Speicherung der erstellten PRIR, Anschlüsse für die Messmikrofone und das Headtrackingsystem, ein Ausgang für 3,5 mm Miniklinke und ein USB-Anschluss zum Laden des Headtrackers. Die Rückseite beherbergt neben dem obligatorischen Netzteilanschluss zur Stromversorgung jeweils 8 unsymmetrische, analoge Cincheingänge und -ausgänge, jeweils

⁷⁸ Reineboth: „Virtuelle Realität – Kann eine simulierte Abhörumgebung eine Reale Tonregie ersetzen?“

<http://www.hdm-stuttgart.de/~curdt/Reineboth.pdf>, Stand: 20.11.2013

einen digitalen HDMI Ein- und Ausgang sowie einen Anschluss für externe Kopfhörerverstärker und einen sogenannten „Seatshaker“, der am Sitz angebracht den fehlenden Körperschall simulieren soll. Bedient wird der Realiser ausschließlich über eine Fernbedienung. Mitgeliefert werden neben diversen Adaptern zur Anbindung an eine Studioperipherie das Headtrackingsystem, ein hochwertiger Kopfhörer mit elektrostatischem Wandlerprinzip sowie die für die Messungen benötigten Ohrmikrofone. Leider standen für diese Arbeit der elektrostatische Kopfhörer und der Seatshaker nicht zur Verfügung. Aufgrund der Hardwarekonfiguration des Realiser A8 kann man davon ausgehen, dass er sowohl für den Consumerbereich (Cinch Ein- und Ausgänge, HDMI, Miniklinke) als auch für den professionellen Einsatz (u.A. XLR-Adapter) konzipiert wurde.

7.2 Vorbereitungen

Bevor der Messvorgang gestartet werden kann, sind einige Vorbereitungen zu treffen. Nach dem Anschluss des Gerätes an die Abhörumgebung müssen die Messmikrofone im Gehörgang der zu messenden Person platziert werden. Hierbei helfen kleine Schaumstofftunnel, die sowohl als Halterung der Mikrofone dienen als auch den Gehörgang verschließen. Es ist sehr genau darauf zu achten, dass die Mikrofone durch nichts verdeckt werden, da sonst die Messung stark beeinträchtigt und das Ergebnis in der Regel unbrauchbar wird. Außerdem sollten die Mikrofone mit dem Eingang des Gehörgangs plan abschließen.

Als nächsten Schritt stellt man im Realiser A8 über die Fernbedienung die verwendete Abhörordnung ein. Es wird hierbei nach ITU-R BS.775-1 von einer kreisförmigen Aufstellung der Lautsprecher ausgegangen, die Positionen der einzelnen Kanäle werden durch die Angabe des Winkels in Relation zum Centerkanal bei 0° angegeben. Dabei sind auch Angaben von Elevationswinkeln und damit die Verwendung von Höhenlautsprechern möglich.

Sind diese Einstellungen vorgenommen, kann mit der Kalibrierung der Lautsprecher begonnen werden. Hierfür ist ein Testprogramm im Gerät integriert, das allen Kanälen einzeln ein Testsignal zuspielt und über die Ohrmikrofone deren Pegel misst. Das Ergebnis wird im kleinen Display ausgegeben. Auch dabei ist strengstens auf einen ausgeglichenen Pegel zwischen den Kanälen zu achten, da es sonst bei der späteren Simulation zu Balanceproblemen kommen kann. Gleichzeitig kann bei der Kalibrierung die vorher eingestellte Lautsprecherzuordnung überprüft werden.

7.3 Die Messung

Für die eigentliche Messung stehen ebenfalls verschiedene Parameter zur Verfügung. Zur Anregung des Systems wird anstatt eines Knallimpulses ein sogenannter Sweep verwendet. Er durchläuft den gesamten hörbaren Frequenzbereich von 20 Hz bis 20 kHz und wird jedem Kanal einzeln zugeführt. Es kann zwischen Sweeps von drei und zwölf Sekunden Länge und der Anzahl der Durchläufe pro Lautsprecher (bis zu 8) gewählt werden. Durch die getrennte Ansteuerung der einzelnen Lautsprecher wird die Abwärtskompatibilität der erstellten PRIR zu kleineren Wiedergabeanordnungen wie z.B. der Zweikanalstereophonie sichergestellt. Die Messung wird jeweils für die Blickrichtungen geradeaus, links und rechts für alle Kanäle vorgenommen. Dabei stehen drei Verfahren zur Ermittlung des Blickwinkels zur Verfügung. Bei der Standardmessung verwendet der Realiser A8 die zuvor angegebenen Winkeleinstellungen. Alternativ lässt sich der korrekte Blickwinkel über das Headtrackingsystem oder mit Hilfe von Pilottönen ermitteln, die abhängig von der Kopfhaltung in drei verschiedenen Frequenzen eingespielt werden. Der mittelfrequente Ton repräsentiert dabei eine korrekte Kopfhaltung, ertönt eine andere Frequenz muss der Blickwinkel entsprechend angepasst werden.

Die Messung der für die vorliegende Arbeit verwendeten individuellen Impulsantwort wurde in der Regie A des Tonstudios der Hochschule der Medien Stuttgart durchgeführt. Die dortige Abhörordnung entspricht dem 5.1 Standard nach ITU-Norm. Es wurde bewusst auf den Einsatz des Headtrackingsystems verzichtet, da dieses bei einer späteren externen Wiedergabe nicht zur Verfügung stehen wird. Um dennoch eine korrekte Lokalisation in frontaler Richtung zu erreichen, wurde die Anzahl der Durchläufe für diese Blickrichtung auf acht erhöht, für den linken und rechten Kanal wurden vier Durchläufe verwendet. Bei der Ermittlung des korrekten Blickwinkels halfen die beschriebenen Pilottöne. Des Weiteren kam zur Qualitätssteigerung der Impulsantwort der längere Sweep von zwölf Sekunden zum Einsatz. Mit einer Unterbrechung von vier Sekunden zwischen den einzelnen Sweeps (hier sollte unbedingt die Nachhallzeit des Raumes bedacht werden!) ergab sich für die 5.1 Wiedergabeanordnung eine Dauer von 25,6 min für die Messung, in der der Kopf absolut ruhig gehalten werden muss. Natürlich lässt sich die Messung durch die Wahl der entsprechenden Parameter verkürzen, allerdings sollte hier die bestmögliche Qualität erreicht werden.

7.4 Ergebnis

Die erstellte PRIR lässt sich im Anschluss anhand von zwei Diagnoseseiten auf dem Display überprüfen. Die erste Seite zeigt das Signal-Rausch-Verhältnis der Impulsantworten, die zweite Seite die Laufzeitunterschiede zwischen den beiden Messmikrofonen in Samples, jeweils für die einzelnen Kanäle und Blickrichtungen⁷⁹. Die Messung in der Regie A lieferte hier sehr gute Werte.

Nach Abschluss der Messung kann die erstellte PRIR im Realiser geladen werden. Über den Input zugeführte Audiosignale werden dann vom Gerät in Echtzeit gefaltet und für bis zu zwei angeschlossene Kopfhörer gleichzeitig ausgegeben. Die Raumsimulation konnte beim anschließenden Hörversuch mit einer Surroundmischung überzeugen, das Ergebnis war beeindruckend und sehr realitätsnah. Besonders beim direkten Vergleich zwischen Lautsprecherwiedergabe und Kopfhörersimulation zeigten sich die Stärken des Realisers. Durch die Individualisierung der Impulsantwort ließen sich alle Kanäle einwandfrei lokalisieren. Lediglich eine leichte Balanceverschiebung des Center-Kanals nach rechts lies sich auch durch weitere Messungen leider nicht verhindern. Am Ende wurde die gemessene PRIR zur externen Verarbeitung über eine SD-Karte exportiert.

7.5 SVS-Dateiformat

Der anschließende Import der PRIR in die DAW stellte sich leider als großes Problem heraus. Voraussetzung für die Faltung mit den einzelnen Kanälen sind kanalbezogene Impulsantworten. Es wird für jeden Kanal eine separate Impulsantwort benötigt, da diese später die Position des simulierten Lautsprechers bestimmt. Leider verwendet die Firma Smyth Research ein eigenes Dateiformat mit der Endung „.SVS“, in dem die Impulsantworten der einzelnen Kanäle zusammengefasst werden. Das Format wird von Smyth Research nicht offengelegt, weshalb es nicht dekodiert werden kann. Es musste also eine Möglichkeit gefunden werden, die benötigten Impulsantworten gesondert aufzunehmen.

⁷⁹ Smyth Research: Realiser A8 User Manual, <http://www.smyth-research.com/downloads/A8manual.pdf>,

7.6 Workaround

Bei folgendem Lösungsansatz hilft das Verständnis des in Kapitel 6.4.1 vorgestellten LTI-Systems. Nicht nur die Schallübertragung in einem Raum, sondern auch signalverarbeitende Geräte können als LTI-System betrachtet werden. So ist es z.B. durchaus üblich, das Verhalten eines Hallgerätes durch eine Impulsantwort zu simulieren. Voraussetzung für diese Methode bleibt aber weiterhin die Zeitinvarianz des Systems, die glücklicherweise auch beim Realiser A8 gegeben ist. Das Verhalten des Realisers kann demnach ebenfalls mit einer Impulsantwort beschrieben werden.

Um die benötigten kanalbezogenen Impulsantworten zu erhalten, geht man wie folgt vor: Anstelle eines Sweeps kann zur Vereinfachung der späteren Signalverarbeitung ein sogenannter Diracstoß⁸⁰ verwendet werden. In der Mathematik ist ein Diracstoß ein unendlich schmaler Impuls mit unendlicher Amplitude⁸¹. Da ein solcher Impuls in der Praxis nicht existiert, wird meist ein Pistolenschuss oder Ähnliches verwendet. Alternativ lässt sich ein schmaler Impuls auch im Editor der DAW „zeichnen“. Der Vorteil ist, dass der Impuls nicht manuell aufwändig durch inverse Filterung von der Impulsantwort getrennt werden muss. Diese Arbeit übernimmt das in Logic Pro 9 enthaltene Hallplugin „Space Designer“, mit dem auch die spätere Faltung durchgeführt werden kann.

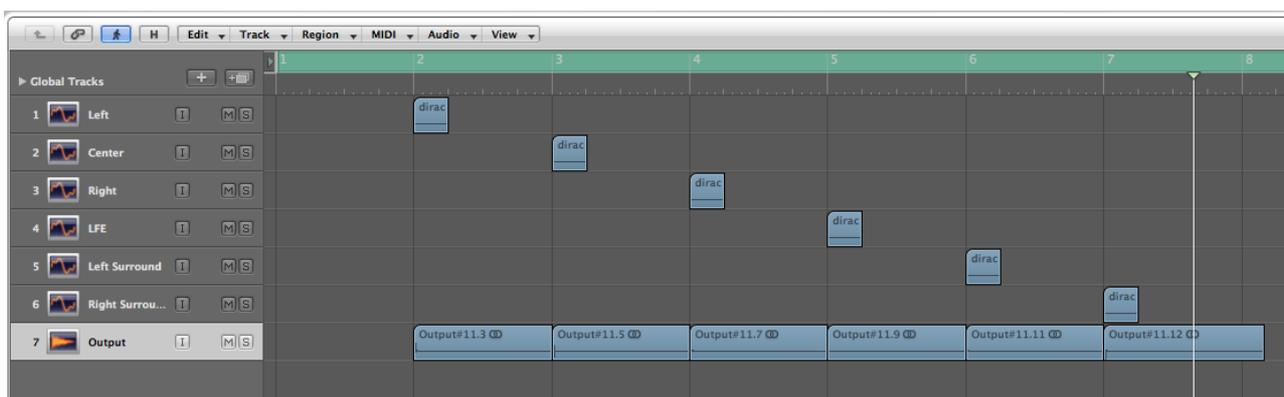


Abbildung 19: Screenshot des Workaround-Projektes in Logic Pro 9

Dem Realiser wird über die DAW ein Mono-Diracstoß sequenziell nacheinander auf allen Kanälen zugeführt und das jeweilige Ausgangssignal als Stereoimpulsantwort aufgenommen. Da die Im-

⁸⁰ Benannt nach dem englischen Physiker Paul Dirac

⁸¹ Vgl. Dickreiter 2008, S. 606

pulsantworten zur Synthese von Musik verwendet werden sollen, wurde bei den praktischen Versuchen zu dieser Arbeit eine Abtastrate von 44,1 kHz und eine Quantisierungstiefe von 24bit verwendet. Dabei muss unbedingt das Headtracking über die Fernbedienung deaktiviert werden (TD-Mode), damit der Realiser nicht zwischen den kanalbezogenen Impulsantworten interpoliert. Anschließend müssen alle Impulsantworten samplegenau geschnitten werden, da sonst bei der Raumsimulation interaurale Laufzeitunterschiede entstehen können, auf die das Gehör sehr empfindlich reagiert. So erhält man für jeden Kanal der Abhörumgebung (Left, Center, Right, LFE, Surround-Left, Surround-Right) die zugehörige Impulsantwort als WAV-Datei, die anschließend in der DAW zur Faltung verwendet werden kann.

8. Konzeption einer Applikation zur Faltung auf Mobilgeräten

Im folgenden Kapitel soll zunächst untersucht werden, ob und wie die aufgenommenen Impulsantworten auf dem Rechner zur Binauralsynthese verwendet werden können. Die gewonnenen Erkenntnisse dienen als Grundlage für die späteren Anforderungsdefinitionen zur Synthese auf Mobilgeräten.

8.1 Externe Faltung in einer DAW

Bei den praktischen Versuchen zu dieser Arbeit wurde die DAW Logic Pro 9 der Firma Apple Inc. verwendet. Das darin enthaltene Hallplugin „Space Designer“ (deutsch: „Raumgestalter“) ist in der Lage, die erstellten Impulsantworten zu importieren und die Faltung auszuführen.



Abbildung 20: Das Hallplugin "Space Designer"

Hierfür müssen die Impulsantworten der einzelnen Kanäle, als mono- oder stereo-Datei entweder im AIFF, SDII oder WAV-Format vorliegen⁸². Nachdem das Space Designer Plugin in alle Kanäle

⁸² Logic Pro 9 Hilfe Bibliothek,

<http://help.apple.com/logicpro/mac/9.1.6/de/logicpro/effects/index.html#chapter=12%26section=0>

Stand: 14.02.2014

le eingefügt wurde, können über das IR Sample Dropdown-Menü auf der linken Seite des Hauptfensters die Impulsantworten ausgewählt und geladen werden. Es stehen dabei verschiedene Optionen wie z.B. eine Hüllkurvenanalyse zur Verfügung. Für eine korrekte Synthese der Surroundanordnung ist unbedingt darauf zu achten, dass die passende IR für den jeweiligen Kanal selektiert wird, da diese die Position der Schallquelle innerhalb der synthetisierten Abhörumgebung bestimmt. Andernfalls kommt es zu einer Vertauschung der Kanäle und die Synthese entspricht nicht mehr dem Originalmix. Aus dem gleichen Grund sollte **kein** Panning vorgenommen werden, alle räumlichen Informationen sind bereits in der IR enthalten. Rechts des Hauptfensters befinden sich die Regler zur Kontrolle des Outputs. Der Schieberegler „Dry“ legt den Pegel des unbearbeiteten, der Regler „Reverb“ den Output-Pegel des Effektsignals fest. Eine sehr nützliche Funktion bei der Verwendung von selbst erstellten IRs ist die 'reverb volume compensation'. Sie gleicht den Pegel der einzelnen IRs untereinander automatisch an, wodurch Lautstärkeunterschiede und damit verbundene Änderungen der Schallquellenentfernungen in der Synthese vermieden werden⁸³. Darüber hinaus hält der Space Designer viele verschiedene Möglichkeiten zur kreativen Gestaltung der Impulsantworten bereit. Sie finden jedoch im Rahmen dieser Arbeit keine Verwendung, da das Schallfeld möglichst originalgetreu synthetisiert werden soll.

Der anschließende Hörtest ergab eine weitgehend zufriedenstellende Synthese des gemessenen Systems. Lediglich geringe Pegelunterschiede zwischen den einzelnen Kanälen verfälschten das Ergebnis; dieses Problem konnte aber mit Hilfe der IR-Pegelkompensation des Space Designers ausgeglichen werden. Dennoch zeigte sich eine leicht verminderte Lokalisationsschärfe der einzelnen Schallquellen gegenüber dem Originalsystem. Außerdem wurde festgestellt, dass das Fehlen einer optischen Information in Form realer Lautsprecher bei den ersten Hörversuchen verwirrend war. Der Effekt verringerte sich jedoch mit zunehmender Hörerfahrung. Eine Anpassung des Kopfhörerfrequenzganges war leider nicht möglich, was die Qualität der Synthese zusätzlich beeinträchtigte. Keiner der üblichen Abbildungsfehler bei der Kopfhörerwiedergabe von Lautsprechermischungen (siehe Kap. 5.2) konnte beobachtet werden. Es stellte sich heraus, dass die Simulation der gemessenen Regie dem Originalraum durchaus ähnlich, aber nicht ganz ebenbürtig war. Dieser Umstand entsprach weitgehend den Erwartungen, da die Messung der

83 Logic Pro 9 Hilfe Bibliothek,

<http://help.apple.com/logicpro/mac/9.1.6/de/logicpro/effects/index.html#chapter=12%26section=0> Stand:
14.02.2014

PRIR schon durch kleinste Veränderungen empfindlich beeinflusst und dadurch praktisch keine „perfekte“ Messung erstellt werden kann. Dennoch konnte die Raumsimulation überzeugen und lieferte ein sehr gutes, räumliches Hörerlebnis.

8.2 Anforderungen

Die behandelten Inhalte dieser Arbeit sollen nun in einer Anforderungsdefinition für die mobile Synthese festgehalten werden. Die einzelnen Anforderungen ergeben sich sowohl aus technischen als auch klanglichen Voraussetzungen, die an die Applikation gestellt werden müssen. So muss neben einer stabilen Schallquellenlokalisierung und einer korrekten räumlichen Darstellung die Vertauschung der Hörereignisrichtungen entlang der Medianebene vermieden werden. Außerdem muss ein ökonomischer Umgang mit der Rechenkapazität des Endgerätes gewährleistet werden, da sie in der Regel auch von anderen Prozessen mehr oder weniger stark in Anspruch genommen wird.

Die Basis einer Applikation zur Binauralsynthese bildet ein auf der Grundlage der in Kapitel 6.4.4 besprochenen Fouriertransformation implementierter Faltungsalgorithmus. Dieser soll in der Lage sein, sowohl die Impulsantworten als auch die zu synthetisierenden Signale vom Zeitbereich in den Frequenzbereich und umgekehrt zu transformieren. Da es sich um Mehrspurproduktionen handelt, deren Kanäle einzeln verarbeitet werden, muss die Applikation mehrkanalfähige Formate wie beispielsweise WAV oder FLAC lesen und dekodieren können. Darüber hinaus ist die Anbindung an eine Datenbank, in der die Impulsantworten gespeichert werden können, zwingend notwendig. Ein Head-Tracking könnte vernachlässigt werden, da es mobil nur sehr schwer umzusetzen⁸⁴, und ein statisches Klangbild in Situationen mit häufigen Kopfbewegungen durchaus erwünscht ist, da es bei einer dynamischen Nachführung eher zu Verwirrungen käme, die das Hörerlebnis stören würden. Als zuschaltbare Funktion würde es aber zu einer Stabilisierung der Hörereignisse entlang der Medianebene führen. Um der variablen Beanspruchung der Rechenkapazität des Gerätes gerecht zu werden, muss eine qualitative und quantitative Skalierbarkeit der Synthese verwirklicht werden.

84 Für nähere Informationen siehe Pörschmann, Christoph: 3D Audio in Mobile Communication Devices – Methods for Mobile Head-Tracking 2007, *Journal of Virtual Reality and Broadcasting*, Volume 4, no.13

Zusammenfassung der genannten Anforderungen:

- Faltungsalgorithmus
- Umgang mit Mehrkanalformaten
- PRIR-Datenbank
- Skalierbarkeit

8.2.1 Skalierbarkeit

Um den verschiedenen Anforderungen an die Applikation im speziellen und die Rechenleistung des Gerätes im Allgemeinen gerecht zu werden, schlagen Sander et al. vor dem Hintergrund weiterer Anwendungsmöglichkeiten der mobilen Schallfeldsynthese die Verwendung einiger Qualitätsparameter vor. In auswählbare Stufen von 1 - 12 unterteilt, sollen sie je nach gewünschter Qualität, verfügbarer oder benötigter Rechenleistung und Schallfeldkomplexität, die Skalierbarkeit der Synthese gewährleisten. Zur anschließenden Leistungsmessung implementierten sie diese in einer Testapplikation.

8.2.1.1 Qualitätsparameter

Im Folgenden werden die von Sander et al. verwendeten Qualitätsparameter vorgestellt und ihre Bedeutung für die Synthese einer Mehrkanalanordnung beleuchtet⁸⁵.

- **HEQ** (Headphone Equalization) – Ausgleich des Kopfhörerfrequenzganges:
Steigert die Qualität der Schallfeldsynthese durch Vermeidung von Klangfarbenverfälschungen durch den KH.
- **Individuelle HRTFs** – Verwendung eigener Außenohrübertragungsfunktionen:
Erhöht die Lokalisationsschärfe und reduziert die Wahrscheinlichkeit einer Vertauschung der Hörereignisrichtungen entlang der Medianebene
- **Head-Tracking** – Nachverfolgung der Kopfbewegungen des Hörers:
Verhindert eine Vertauschung der Hörereignisrichtungen entlang der Medianebene und ermöglicht ein in Relation zur Kopfposition dynamisch anpassbares Klangbild.

85 Vgl. Sander et al.: Scalable binaural synthesis on mobile devices 2012, AES Convention Paper 8783, S. 3

Alle weiteren Parameter sind nur bei der Verwendung von dynamischen Schallquellen von Bedeutung, sollen aber der Vollständigkeit halber kurz erwähnt werden:

- **FIF** (filter interpolation in the frequency domain) – Filterinterpolation im Frequenzbereich: Erhöht die räumliche Auflösung der Synthese
- **FIT** (filter interpolation in the time domain) – Filterinterpolation im Zeitbereich: Erhöht die räumliche Auflösung der Synthese
- **DS** (dynamic sources) - Unterscheidung von statischen und dynamischen Schallquellen
- **ER** (early reflections) – Erstellung von Raumimpulsantworten auf Basis der Polarkoordinaten des Eingangssignals
- **FUR** (filter update rate) – Anpassbares Aktualisierungsintervall der Filter zur Einsparung von Rechenkapazität
- **LA & LE** (limited angular range in azimuth and elevation) – horizontale und vertikale Winkelbegrenzung zur Vermeidung des Ladens von unnötigen Filtern
- **SM** (smoothing of sound source movement) - Glättung von Schallquellenbewegungen
- **FC** (filter crossfading) – Filter Crossfading

8.3 Entwurf der Rendering Engine

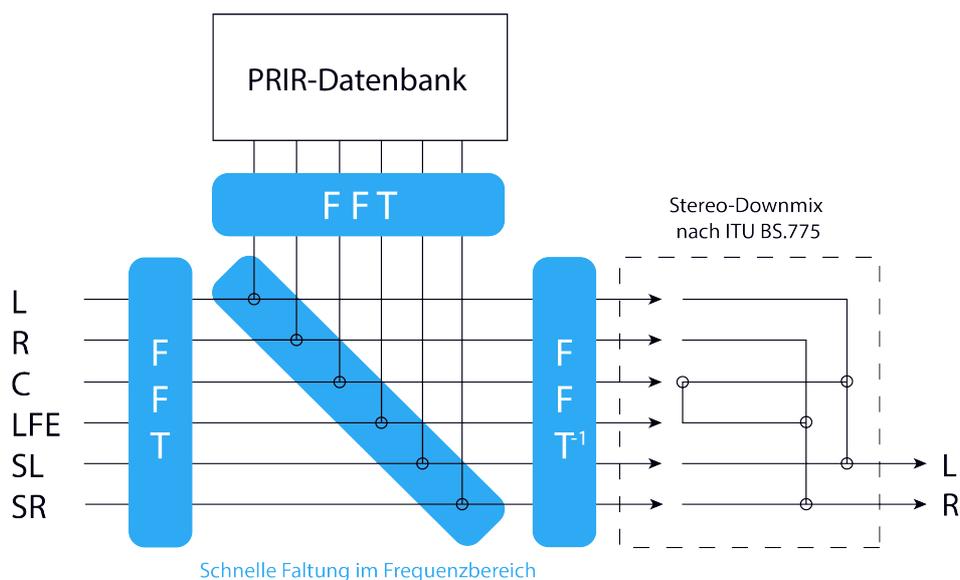


Abbildung 21: Rendering Engine zur Synthese einer 5.1 Aufnahme

Sowohl alle Eingangssignale, als auch die entsprechenden PRIR's werden blockweise durch eine Fouriertransformation vom Zeit- in den Frequenzbereich gewandelt, multipliziert und anschließend durch die inverse Fouriertransformation wieder in den Zeitbereich überführt. Es folgt ein automatischer Stereo-Downmix nach ITU-Empfehlung BS.775. Der optionale Ausgleich des Kopfhörerfrequenzgangs durch inverses Filtern einer zuvor gemessenen Kopfhörerübertragungsfunktion findet nach dem Downmix statt.

8.4 Leistungsmessung

Sander et al. verwendeten für die Messung der Leistungsfähigkeit ihrer Testapplikation ein Apple iPhone 4s mit zwei ARM Cortex-A9 Prozessorkernen bei einer Taktfrequenz von 800 MHz, 512MB Arbeitsspeicher und dem Betriebssystem iOS 5.1⁸⁶. Es wurden 10 Kanäle bei einer Blocklänge von 1024 Samples und 44.1 kHz Abtastrate in den Qualitätsstufen 1 – 12 synthetisiert.

Qualitätsstufe	Aktive Qualitätsparameter	Rechenzeit ∅	Speichernutzung	Maximale Kanalanzahl
1	statisch, HEQ	6001 µs	2,26 MB	28
2	statisch, HEQ, ER	6021 µs	2,28 MB	27
3	DS, FUR=2, LA=180°, LE=0°	7001 µs	2,04 MB	26
4	DS, FUR=2, LA=360°, LE=0°, SM,ER	7186 µs	2,08 MB	25
5	DS, FUR=2, LA=180°, LE=90°, SM,ER, HEQ	7453 µs	2,92 MB	24
6	DS, FUR=2, LA=180°, LE=180°, FC,ER,HEQ	8428 µs	4,87 MB	21
7	DS, FUR=1, full-sphere, SM, ER, HEQ	8661 µs	4,48 MB	20
8	DS, FUR=1, full-sphere, SM, ER, HEQ, FIT	9149 µs	5,09 MB	19
9	DS, FUR=1, full-sphere, FC, ER, HEQ	10617 µs	6,11 MB	17
10	DS, FUR=1, full-sphere, FC, ER, HEQ, FIT	11093 µs	6,43 MB	16
11	DS, FUR=1, full-sphere, SM, ER, HEQ, FIF	33089 µs	5,87 MB	7
12	DS, FUR=1, full-sphere, FC, ER, HEQ, FIF	35449 µs	8,39 MB	6

Tabelle 2: Die einzelnen Qualitätsstufen und ihre Ressourcennutzung in der Übersicht

Aus den in Tabelle 2 dargestellten Ergebnissen ist abzulesen, dass auf der höchsten Qualitätsstufe insgesamt sechs Kanäle synthetisiert werden können. Allerdings übersteigt die Rechenzeit der

⁸⁶ Vgl. Sander et al., S. 5

Stufen elf und zwölf den maximal verfügbaren Zeitraum von 23220 μs ⁸⁷. Eine Synthese der in dieser Arbeit exemplarisch besprochenen 5.1 Mehrkanalanordnung mit statischen Schallquellen (Lautsprecher) kann schon auf Stufe zwei problemlos durchgeführt werden. Hier sind bis zu 27 Kanäle darstellbar. Damit wäre theoretisch die Synthese einer 22.2 Mehrkanalanordnung bei einer Ausnutzung von 65% der verfügbaren Rechenleistung möglich⁸⁸. Bei halber Blocklänge sinkt die Anzahl der maximal synthetisierbaren Kanäle (auf Stufe eins von 28 auf 13) aufgrund einer erhöhten Anzahl von Rechenoperationen⁸⁹. Vor dem Hintergrund der stetig steigenden Leistungsfähigkeit aktueller mobiler Geräte kann davon ausgegangen werden, dass die Synthese aller in der Musikproduktion gängigen Mehrkanalanordnungen mit diesen problemlos möglich ist. In Abbildung 22 sind die Ergebnisse der Leistungsmessung nochmals grafisch dargestellt.

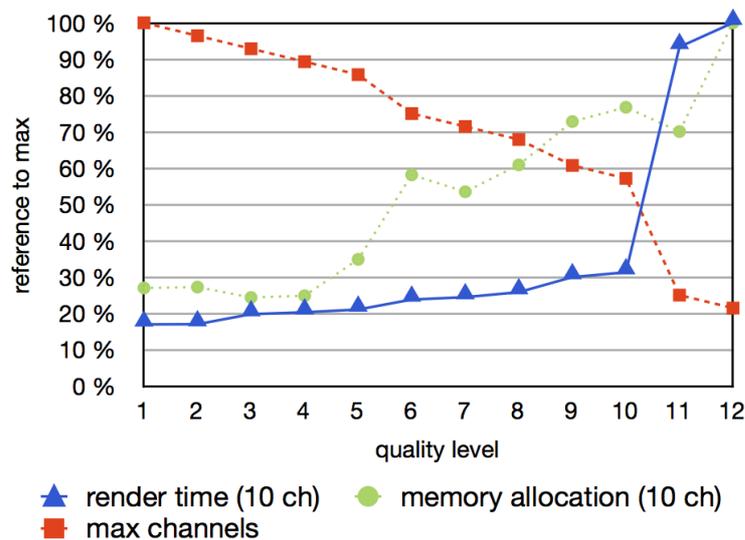


Abbildung 22: Ergebnisse für die einzelnen Qualitätsstufen der Synthese: maximale Rechenzeit = 35449 μs , maximale Arbeitsspeicherbelastung = 8,38 MB, maximale Kanalanzahl = 28

87 Vgl. Sander et al., S. 6

88 Ebd, S. 7

89 Ebd, S. 6

9. Fazit

Zusammenfassend lässt sich sagen, dass die Binauralsynthese auf mobilen Geräten mit den gegebenen technischen Möglichkeiten sehr gut zu verwirklichen ist. Die Achillesverse des vorgestellten Workflow ist sicherlich die Messung der individuellen Außenohrübertragungsfunktion. Zwar liefert sie bei der späteren Synthese eine sehr gute räumliche Darstellung der Inhalte, ist aber kompliziert, fehleranfällig und nicht ohne großen technischen Aufwand durchzuführen. Für den durchschnittlichen Musikliebhaber, der keinen Realiser A8 und keine Mehrkanalanlage sein Eigen nennt, bleibt momentan nur das Angebot einiger weniger Tonstudios, in ihren Räumlichkeiten eine Messung anfertigen zu lassen. An dieser Stelle wäre eine Vereinfachung wünschenswert und nötig, um das Verfahren für die breite Masse zugänglich zu machen. Die Verwendung bereits existierender Messungen in einer Datenbank, aus der sich der Anwender die für seine anatomischen Merkmale passende auswählen kann, geht mit einem deutlichen Qualitätsverlust einher. Diese Option käme dem Hören mit fremden Ohren gleich und ist deshalb nur als Notlösung in Betracht zu ziehen.

Das in der vorliegenden Arbeit besprochene Verfahren bringt für die mobile Musikwiedergabe, im Vergleich zu gängigen Methoden, einen deutlichen Mehrwert mit sich. So konnte gezeigt werden, dass mit Hilfe der Binauraltechnik eine Mehrkanalanlage mit allen räumlichen und klanglichen Eigenschaften über zwei Kanäle auf Mobilgeräten synthetisiert werden kann. Dabei traten keine der üblichen Abbildungsfehler bei der Verwendung von Kopfhörern auf (mit Ausnahme eines durchaus erwünschten, statischen Klangbildes), wodurch bestehende sowie zukünftige Lautsprechermischungen für die Kopfhörerwiedergabe kompatibel werden. Dieser Umstand öffnet nicht nur dem Konsumenten, sondern auch Musikschaffenden neue Wege und bildet die Grundlage, um den Mehrkanalton auch im mobilen Bereich fest zu etablieren. Wer hätte nicht gerne eine „Surroundanlage für die Hosentasche“?

Die Performance aktueller Geräte stellt dabei keinesfalls ein Hindernis dar. Es konnte gezeigt werden, dass deren Leistung für die reine Musikwiedergabe mehr als ausreichend ist und auch anspruchsvollere Anwendungen mit dynamischen Schallquellen umgesetzt werden können. Im Hinblick auf die rasanten technischen Entwicklungen der letzten Jahre, darf aufgrund einer stetig verbesserten Leistungsfähigkeit und effektiveren Algorithmen, ein Blick über den Tellerrand

gewagt werden. Neben der reinen Musikwiedergabe ergeben sich weitere Anwendungsmöglichkeiten im mobilen Bereich. Beispielsweise ließe sich eine Verbesserung der Sprachverständlichkeit durch räumliche Trennung der Teilnehmer von Telefonkonferenzen erreichen. Ferner wären erweiterte Navigationsanwendungen, Steigerung der Realitätsnähe bei Videospielen oder auditive Interfaces denkbar. Die zur Umsetzung solcher Anwendungen benötigten Grundlagen und Anforderungen wurden in dieser Arbeit untersucht. Es bleibt abzuwarten, ob und in welcher Form die vielfältigen Möglichkeiten der Binauraltechnik dafür genutzt werden, das mobile Hörerlebnis um eine Dimension zu erweitern.

Literaturverzeichnis

Bücher

Birkner, Christian: *Surround: Einführung in die Mehrkanalton-Technik ; das Audioformat der Zukunft ; Standards, Aufnahmepraxis, Weiterverarbeitung*, Presse-Project-Verlag GmbH, Bergkirchen 2002

Blauert, Jens: *Räumliches Hören*, S. Hirzel Verlag, Stuttgart 1974

Dickreiter, Micheal et al.: *Handbuch der Tonstudioteknik*, Band 1 & 2, 7. Auflage, K.G. Saur Verlag, München 2008

Friesecke, Andreas: *Die Audio-Enzyklopädie. Ein Nachschlagewerk für Tontechniker*, K.G. Saur Verlag, München 2007

Webers, Johannes: *Handbuch der Tonstudioteknik für Film, Funk und Fernsehen*, 9. Auflage, Franzis Verlag, 2007

Stefan Weinzierl (Hrsg.): *Handbuch der Audiotechnik*, Springer Verlag, Berlin 2008

Zölzer, Udo: *Digitale Audiosignalverarbeitung*, Teubner Verlag, Stuttgart 2005

Papers

Heinrichs, Heiner: *Web-basierter Konsum und Vertrieb von Musik, Film und Videospielen: Chancen veränderter Beziehungen zum Medienprodukt*, Masterarbeit, Bielefeld 2012

Keinath, Daniel & Tarnow, Christopher: *Entwicklung und Erprobung einer Methode zur Integration binauraler Raumimpulsantworten bei der Mischung von Mehrspurproduktionen*, Diplomarbeit, Detmold 2009

Pörschmann, Christian: *3-D Audio in Mobile Communication Devices: Methods for Mobile Head-Tracking*, 2007

Reineboth, Phillipp: *Virtuelle Professionalität – kann eine simulierte Abhörumgebung eine reale Tonregie ersetzen?*, Bachelorarbeit, Stuttgart 2012

Runow, Bernfried: *Automatischer Stereo-Downmix von 5.1-Mehrkanalproduktionen*, Diplomarbeit, Stuttgart 2008

Sander, Christian et al.: *Scalable binaural synthesis on mobile devices*, AES Convention Paper 8783, 2012

Seidl, Michael: *Surround – Das Format der Zukunft?* Hicktown-Records, 2010

Theile, Günther: *Über die Lokalisation im überlagerten Schallfeld*, Berlin, 1980

Theile, Günther & Wittek, Helmut: *Die dritte Dimension für Lautsprecher-Stereophonie*, Mai 2011

Zöllner, Oliver (Hrsg.): *Nutzung mobiler Medien, Struktur und Ordnung – drei Projektberichte*, Stuttgart 2013

Internet

Aguilera, Emanuel et al.: An Immersive Multi-Party Conferencing System for Mobile Devices Using 3D Binaural Audio, http://iteamserver.iteam.upv.es/revista/2012/1 ITEAM_2012.pdf, Stand: 10.02.2014

Consumer Electronics 2013: Marktentwicklung und Trends
http://www.bitkom.org/files/documents/BITKOM_Praesentation_CE-PK_02_09_2013.pdf,
Stand: 26.10.2013

Entwicklung des iPod, <http://www.apple.com/pr/products/ipodhistory/>, Stand: 25.10.2013

Gfu CEMIX 2012
http://www.gfu.de/srv/easyedit/ts_1361800251000/page:home/download/marktstudien/sl_1361547008349/args.link01/de_CEMIX%20Q1-Q4%202012.pdf, Stand: 28.10.2013

Google Studie - Unser mobiler Planet: Deutschland
http://services.google.com/fh/files/blogs/our_mobile_planet_germany_de.pdf,
Stand 26.10.2013

Gurski, Rainer: *Wahrnehmen – ein Lehrbuch*, 2006,
<http://eco.psy.ruhr-uni-bochum.de/ecopsy/download/Gurski-Lehrbuch/Inhaltsverzeichnis.html>,
Stand: 02.10.2013

Infografik: Nicht ohne mein Handy – Handynutzung in Deutschland und der Welt,
<http://mobilbranche.de/2013/10/infografik-nicht-handy/39586>, Stand 26.10.2013

Jahreswirtschaftsbericht 2012 Bundesverband Musikindustrie,
http://www.musikindustrie.de/fileadmin/piclib/statistik/branchendaten/jahreswirtschaftsbericht-2012/download/Jahrbuch_BVMI_2012.pdf, S. 9, Stand: 29.10.2013

Logic Pro 9 Hilfe Bibliothek,
<http://help.apple.com/logicpro/mac/9.1.6/de/logicpro/effects/index.html#chapter=12%26section=0>, Stand: 14.02.2014

„Mobile“ knackt 20-Millionen-Grenze: Mobiles Wachstum ungebremst,
http://www.iqm.de/fileadmin/user_upload/Marketing_Service/Downloads/AGOF/Pressemitteilung_AGOF_mobile_facts_2012-II.pdf, Stand: 26.10.2013

MP3 Infobroschüre des Fraunhofer IIS,
http://www.iis.fraunhofer.de/content/dam/iis/de/dokumente/amm/broschueren/mp3_Broschue_re_A4_16S2012.pdf, Stand: 25.10.2013

Musiknutzung,
http://www.musikindustrie.de/jahrbuch-musiknutzung-2012/?no_cache=1&type=1,
Stand: 26.10.2013

Sengpiel, Eberhard: Kopfbezogene Übertragungsfunktion HRTF,
<http://www.sengpielaudio.com/KopfbezogeneUebertragungsfunktionHRTF.pdf>,
Stand: 12.11.2013

Sengpiel, Eberhard: Stereophonie für Lautsprecher und Stereophonie für Kopfhörer 2,
<http://www.sengpielaudio.com/StereoFuerLautsprUndKopfhoerer2.pdf>, Stand: 08.11.2013

Sengpiel, Eberhard: Transaural Stereo – Kunstkopf-Stereophonie über Lautsprecher,
<http://www.sengpielaudio.com/TransauralStereo.pdf>, Stand: 13.11.2013

Smyth Research: Realiser A8 User Manual,
<http://www.smyth-research.com/downloads/A8manual.pdf>, Stand: 23.11.2013

Zehn Jahre MP3-Player: Als Digitalmusik tragbar wurde,
<http://www.spiegel.de/netzwelt/spielzeug/zehn-jahre-mp3-player-als-digitalmusik-tragbar-wurde-a-540664.html>, Stand: 25.10.2013

Anhang

Auf dem beigelegten Datenträger befindet sich diese Arbeit als PDF-Datei, sowie zwei Hörproben der gleichen Mischung für einen Hörvergleich via Kopfhörer. Die erste Hörprobe ist eine nicht synthetisierte Zweikanalmischung, die zweite eine mit den gemessenen Impulsantworten beaufschlagte Synthese der Surroundanlage der Regie A, HdM. Es sei an dieser Stelle nochmals darauf hingewiesen, dass die Synthese lediglich für den Autor der Arbeit volle Gültigkeit aufweist. Bei anderen Personen wird erwartungsgemäß nicht die volle Klangqualität und Lokalisationsschärfe erreicht werden. Es soll nach der Lektüre aber trotzdem die Möglichkeit bestehen, sich beide Varianten einmal anzuhören.