Mikrofonaufnahmetechnik für Atmosphären als Sounddesign-Komponente immersiver Audioproduktionen

Masterarbeit im Studiengang Audiovisuelle Medien

vorgelegt von

Marcel Remy

Matr.-Nr.: 37445

am 08. Januar 2021 an der Hochschule der Medien Stuttgart zur Erlangung des akademischen Grades Master of Engineering (M. Eng.)

Erstprüfer: Prof. Dr. Frank Melchior Zweitprüfer: Prof. Oliver Curdt

Ehrenwörtliche Erklärung

Hiermit versichere ich, Marcel Remy ehrenwörtlich, dass ich die vorliegende Masterarbeit mit dem Titel: "Mikrofonaufnahmetechnik für Atmosphären als Sounddesign-Komponente immersiver Audioproduktionen" 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 ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§ 23 Abs. 2 Master-SPO (3 Semester) der HdM) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.

Marcel Remy Stuttgart, den 08.01.2021

Kurzfassung

Die vorliegende Arbeit befasst sich mit Atmosphären im Kontext immersiver Audioproduktionen. Neben relevanten Definitionen und theoretischen Grundlagen werden die Funktionen aufgezeigt, die Atmosphären im Rahmen des Sounddesigns übernehmen. Anhand der beschriebenen gestalterischen Anforderungen an Atmosphären werden die grundlegenden psychoakustischen Aspekte betrachtet, welche die Umsetzung der gestalterischen Anforderungen perzeptiv ermöglichen. Daran anknüpfend werden mehrkanalige Wiedergabesysteme untersucht und Anforderungen an Mikrofonsysteme zur Atmosphärenaufnahme abgeleitet. Hierbei liegt der Fokus auf den technischen Bedingungen, die erfüllt werden müssen, um eine möglichst umhüllende Wiedergabe zu realisieren. Es werden zudem technische Aspekte und Systeme der Mikrofontechnik für Atmosphärenaufnahmen vorgestellt. Zusätzlich erfolgt die Konzeption verschiedener Mikrofonsysteme ihre praktische Erprobung im und Rahmen von Atmosphärenaufnahmen im Freien. Darauf aufbauend werden mit Hilfe einer Kohärenzanalyse die erstellten Atmosphärenaufnahmen signaltechnisch untersucht. Dabei steht die Frage im Vordergrund, inwieweit die Vorhersage der Diffusfeldkohärenz bei der Konzeption von Mikrofonsystemen gewinnbringend eingesetzt werden kann und ob durch die Kohärenzanalyse von Ohrsignalen die Abbildungseigenschaften von Mikrofonsystemen bewertet werden können. Abschließend werden die Analyseergebnisse kritisch betrachtet und Empfehlungen zur praktischen Atmosphärenaufnahme für immersive Audioproduktionen ausgesprochen.

Abstract

This paper deals with ambiences in the context of immersive audio productions. Besides important definitions and theoretical basics, the creative functions that ambiences take over in the context of sound design are presented. Based on the sound design requirements, the basic psychoacoustic aspects are considered, which enable the implementation of the design requirements perceptually. Following on from this, multichannel playback systems are examined and requirements for microphone systems for ambience recording are derived. In this context, the focus is on the technical requirements that must be met in order to realize a highly enveloping reproduction. Relevant technical aspects and systems of microphone technology for ambience recording are presented. In addition, the conception of different microphone systems and their practical testing in outdoor ambience recording will be presented. Based on this, a coherence analysis is performed on the signals of the recorded ambiences. The main question is to what extent the prediction of the diffuse field coherence can be profitably used in the conception of microphone systems for the recording of ambiences and if the coherency analysis of binaural signals can be used to evaluate the imaging properties of microphone systems. Finally, the results of the analysis are critically reviewed and recommendations for practical ambience recording for immersive audio productions are given.

Danksagung

An dieser Stelle möchte ich all jenen danken, die durch ihre fachliche und persönliche Unterstützung zum Gelingen dieser Masterarbeit beigetragen haben. Besonderer Dank gilt Professor Dr. Frank Melchior für die Betreuung meiner Arbeit. Die zahlreichen Fachgespräche haben mir sehr geholfen, die Thematik aus verschiedenen Perspektiven zu betrachten und mich dazu motiviert, mich inneue Themengebiete und Programmiersprachen einzuarbeiten. Weiterer Dank gilt meinem Zweitprüfer Professor Oliver Curdt, der mich während des gesamten Studiums begleitet und mit hohem Engagement zahlreiche spannende Vorlesungen, Studioproduktionen, Workshops und Exkursionen organisiert hat.

Ich danke Christoph Niesar und Udo Wüstendörfer vom Hessischen Rundfunk sowie Arne Morgner von der HDMK Stuttgart für die Bereitstellung hochwertiger Mikrofone. Ohne sie wären die praktischen Versuche dieser Arbeit nicht möglich gewesen. Besonderer Dank gilt Bernd Rischner, der mich während der Aufnahmen tatkräftig unterstützte und Wind und Wetter trotzte.

Herzlich bedanken möchte ich mich auch bei meinen Kommilitonen Daniela Rieger, Tobias Kurzweg, Leon Hofmann, Jonas Kieser, Jona Eisele und Bastian Kilper für den freundschaftlichen Zusammenhalt während des gesamten Masterstudiums und die Realisierung unseres Masterprojektes. Ein besonderes Dankeschön möchte ich Daniela Rieger aussprechen für all die persönliche Unterstützung und Ermutigung, auch in schwierigen Zeiten.

Inhaltsverzeichnis

E	hrenwö	rtliche Erklärung	II	
K	KurzfassungIII			
A	AbstractIV			
D	anksagu	ung	V	
Ir	nhaltsve	erzeichnis	VI	
A	bbildun	ngsverzeichnis	IX	
F	ormelve	erzeichnis	XII	
Т	abellen	verzeichnis	XII	
L	istingve	erzeichnis	XII	
A	bkürzuı	ngsverzeichnis	XIII	
1	Einl	eitung	1	
	1.1	Motivation und Zielsetzung	1	
	1.2	Aufbau der Arbeit	3	
2	Defi	nition und Kategorisierung atmosphärischer Klänge	6	
	2.1	Begriffsdefinitionen	6	
	2.2	Kategorisierungsmodelle	9	
3	Bed	eutung und Funktion von Atmosphären im Sounddesign-Prozess	14	
4	Imm	nersion	20	
	4.1	Immersion in Medienproduktionen	20	
	4.2	Auditive Dimensionen von Immersion	22	
5	Psyc	choakustische Aspekte der Wahrnehmung von Atmosphären	26	
	5.1	Lokalisation einer Schallquelle	26	
	5.1.1	1 Außenohr-Übertragungsfunktion	27	
	5.1.2	2 Lokalisation in der Horizontalebene	27	
	5.1.3	3 Lokalisation in der Vertikalebene	30	
	5.1.4	4 Lokalisation mehrerer Schallquellen	32	
	5.2	Entfernungswahrnehmung	33	
	5.3	Interaurale Korrelation und Umhüllung	34	
	5.4	Hörszenenanalyse	38	

6	Wied	lergabesysteme	. 41
	6.1	Grundsätzliche Anforderungen	. 42
	6.2	Etablierte Lautsprechersetups	. 43
	6.3	Einsatz der Surroundkanäle	. 46
	6.4	Einsatz der Höhenkanäle	. 49
	6.5	Einsatz des Centerkanals	. 52
	6.6	Tieftonwiedergabe	. 54
	6.7	Entwicklung einer virtuellen Abhörumgebung in MAX	. 55
7	Mikr	ofon- und Aufnahmetechnik	62
	7.1	Anforderungen an die Mikrofontechnik	. 62
	7.1.1	Technische Anforderungen	. 62
	7.1.2	Anforderungen für Atmosphären als Sounddesignkomponente	. 64
	7.2	Spezielle Mikrofonsysteme zur Atmosphärenaufnahme	. 65
	7.2.1	ESMA-3D	. 65
	7.2.2	Ortf-3D	. 66
	7.2.3	Doppel U-Fix	. 67
8	Test	aufnahmen und Analyse	. 69
	8.1	Dokumentation der erstellten Testaufnahmen	. 69
	8.1.1	Konzept	. 70
	8.1.2	Aufnahmeorte	. 70
	8.1.3	Aufnahmetechnik	. 73
	8.1.4	Aufbau der verwendeten Mikrofonsysteme	. 75
	8.2	Signaltechnische Analyse der Abbildungseigenschaften	. 79
	8.2.1	Theoretische Grundlagen	. 80
	8.2.2	Vorhersage von Abbildungseigenschaften mit Hilfe des Schoeps Image	
		Assistant	. 83
	8.2.3	Signalanalyse zur Betrachtung der Abbildungseigenschaften realer	
		Atmosphären–Aufnahmen	. 89
	8.2.4	Vergleich von Vorhersage und Praxis	. 92
	8.3	Analyse der interauralen Kohärenz	103

8.3.	1 Funktionstest mit Rauschsignalen und verschiedenen
	Lautsprecherkonfigurationen104
8.3.2	2 Vergleich zwischen Kunstkopfaufnahmen und binauralisierten
	Mikrofonarrays107
8.4	Diskussion der Analyseergebnisse111
9 Faz	it und Ausblick
Literatur	verzeichnisX
Anhang.	- 1 -
A1 MA	ATLAB Live-Skripte 1 -
A1.	1 SOFA-Analyse 1 -
A1.5	2 Kohärenz–Auswertung 8 -
A1.:	3 DFI-P Berechnung 15 -
A1.4	4 Binaural-Analyse Wiedergabesystem 4+5+0 19 -
A1.	5 Binaural-Analyse Wiedergabesystem 4+7+0 23 -
A1.0	6 Binaural-Analyse zum Vergleich der Wiedergabesysteme 28 -
A2 Ko	härenzfunktionen der Mikrofonsysteme 32 -
A2.	l "Hamasaki+Height" 32 -
A2.2	2 "Kugel-Array" 33 -
A2.:	3 "Grenzflächen-Array" 35 -
A2.4	4 "Nieren-Array I" 36 -
A2.	5 "Nieren-Array II" 38 -
A2.0	3 "Doppel-IRT" 39 -
A3 Ko	härenzfunktionen der Kunstkopfaufnahmen und der binauralisierten
Mikrot	önsysteme im Vergleich 41 -
A3.	l "Hamasaki+Height" 41 -
A3.2	2 "Kugel-Array" 41 -
A3.:	3 "Grenzflächen-Array" 42 -
A3.4	4 "Nieren-Array I" 42 -
A3.	5 "Nieren-Array II" 43 -
A3.0	5 "Doppel-IRT" 43 -

Abbildungsverzeichnis

Abbildung 1:	Soundscape Kategorisierungsmodelle	13
Abbildung 2:	Definitionskonzepte von Immersion (Agrawal et al., 2019)	21
Abbildung 3:	Lokalisationsschärfe elevierter Sprachsignale (Pulkki & Karjalainen, 2015)	31
Abbildung 4:	Lage der Hörereignisorte bei Kopfhörerwiedergabe von Rauschsignalen mit unterschiedlichen Korrelationsgraden k (Chernyak & Dubrowsky, 1968 zitiert nach Blauert, 1974)	36
Abbildung 5:	Auswirkung des Korrelationsgrades bei Lautsprecherwiedergabe (Links: Korrelationsgrad 0,72. Mitte: Korrelationsgrad 0,35. Rechts: Korrelationsgrad 0,15) (Theile, 2001) schematisiert nach (Damaske, 1967).	37
Abbildung 6:	Hörereignislage von Sprach- und Rauschsignalen bei unterschiedlichen Korrelationsgraden k (Licklider, 1948 zitiert nach Blauert, 1974)	40
Abbildung 7:	Wiedergabesystem D (4+5+0) nach ITU-R BS.2051-2. Links: Upper-Layer. Rechts: Middle-Layer (ITU, 2018)	44
Abbildung 8:	Wiedergabesystem J (4+7+0) nach ITU-R BS.2051-2. Links: Upper-Layer. Rechts: Middle-Layer (ITU, 2018)	45
Abbildung 9:	Dipol-Lautsprecher im 5.1 Wiedergabesystem (Holman, 2008)	48
Abbildung 10:	Beispielhafte Anordnung eines 5.1 Wiedergabesystems im Kino (Peter, 2020)	49
Abbildung 11:	Schalldruck-Isobaren der frequenzabhängigen Abstrahlwinkel eines 3-Wege Lautsprechers. Oben: Abstrahlcharakteristik horizontal. Unten: Abstrahlcharakteristik vertikal. Modell: Neumann KH 310 (NEUMANN, o. J.)	52
Abbildung 12:	Virtuelle Abhörumgebung in MAX (Präsentationsansicht)	57
Abbildung 13:	Messpunkte des HRTF-Datensatzes der virtuellen Abhörumgebung	59
Abbildung 14:	HRTF-Messpunkte der Lautsprecherkonfiguration (4+5+0)	59
Abbildung 15:	HRTF-Messpunkte der Lautsprecherkonfiguration (4+7+0)	60
Abbildung 16:	Virtuelle Lautsprecherpositionen (4+5+0) in MAX	60
Abbildung 17:	Virtuelle Lautsprecherpositionen (4+7+0) in MAX	61
Abbildung 18:	Koinzidente Mikrofonanordnung in der Vertikalebene beim Mikrofonsystem "ESMA 3D" (Lee, 2018).	66
Abbildung 19:	Mikrofonsystem "ORTF 3D" im Windkorb (Wittek, o. J.)	67

Abbildung 20:	Kapselabstände und Versatzwinkel des Mikrofonsystems "Doppel U-Fix", links: Bottom-Layer, rechts: Height-Layer (Camerer, 2020)	68
Abbildung 21:	Das Mikrofonsystem "Doppel U-Fix" im Einsatz (Camerer, 2020)	68
Abbildung 22:	Atmosphärenaufnahme im Reinhardswald	69
Abbildung 23:	Aufnahmeort 1 – Habichtswald (Quelle: Google Maps)	71
Abbildung 24:	Aufnahmeort 2 – Reinhardswald (Quelle: Google Maps)	72
Abbildung 25:	Aufnahmeort 3 - Kasseler Innenstadt, Parkhausdeck (Quelle: Google Maps)	73
Abbildung 26:	$Mikrofonsystem \ Hamasaki \ + \ Height. \ \dots \ \dots$	75
Abbildung 27:	Mikrofonsystem Kugel-Array	76
Abbildung 28:	Mikrofonsystem Grenzflächen-Array	77
Abbildung 29:	Mikrofonsystem Nieren-Array I.	77
Abbildung 30:	Mikrofonsystem Nieren-Array II.	78
Abbildung 31:	Mikrofonsystem Doppel IRT-Kreuz	79
Abbildung 32:	Kohärenzfunktionen des Schoeps Image Assistant für das Mikrofonsystem "Hamasaki + Height"	84
Abbildung 33:	Kohärenzfunktionen des Schoeps Image Assistant für die Mikrofonsysteme "Kugel–Array" und "Grenzflächen–Array"	85
Abbildung 34:	Kohärenzfunktionen des Schoeps Image Assistant für die Mikrofonsysteme "Nieren–Array I", "Nieren–Array II" und "Doppel–IRT".	86
Abbildung 35:	Wichtungsfunktion zur Berechnung des DFI-P in MATLAB	91
Abbildung 36:	Über zwei Minuten gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Habichtswald mit dem "Kugel-Array"	93
Abbildung 37:	Über zwei Minuten gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Reinhardswald mit dem "Kugel-Array"	94
Abbildung 38:	Über zwei Minuten gemittelte Kohärenzverläufe der Atmosphärenaufnahme in der Kasseler Innenstadt mit dem "Kugel-Array"	94
Abbildung 39:	Über fünf Sekunden gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Reinhardswald mit dem "Kugel-Array", Ausschnitt 00:15 – 00:20.	96
Abbildung 40:	Über fünf Sekunden gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Reinhardswald mit dem "Kugel-Array", Ausschnitt 00:30 – 00:35.	96
Abbildung 41:	Über zwei Minuten gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Habichtswald mit dem "Doppel-IRT"	97

Abbildung 42:	Über zwei Minuten gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Reihnardswald mit dem "Doppel-IRT" 98
Abbildung 43:	Über zwei Minuten gemittelte Kohärenzverläufe der Atmosphärenaufnahme in der Kasseler Innenstadt mit dem "Doppel-IRT"
Abbildung 44:	Über fünf Sekunden gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Habichtswald mit dem "Doppel-IRT", Ausschnitt 00:45 – 00:50100
Abbildung 45:	Über fünf Sekunden gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Habichtswald mit dem "Doppel-IRT", Ausschnitt 01:20 – 01:25100
Abbildung 46:	Kohärenzvergleich synthetischer Binauralsignale, erstellt durch Wiedergabe von dekorreliertem rosa Rauschen über verschiedene Lautsprecher des (4+5+0)-Wiedergabesystems; gemittelt über 60 Sekunden
Abbildung 47:	Kohärenzvergleich synthetischer Binauralsignale, erstellt durch Wiedergabe von dekorreliertem rosa Rauschen über verschiedene Lautsprecher des (4+7+0)-Wiedergabesystems; gemittelt über 60 Sekunden
Abbildung 48:	Kohärenzvergleich synthetischer Binauralsignale, erstellt durch Wiedergabe von dekorreliertem Rauschen über die Wiedergabesysteme (4+5+0) und (4+7+0); gemittelt über 60 Sekunden
Abbildung 49:	Kohärenzverläufe von Kunstkopf und binauralisiertem "Kugel- Array", gemittelt über zwei Minuten109
Abbildung 50:	Kohärenzverläufe von Kunstkopf und binauralisiertem "Doppel- IRT", gemittelt über zwei Minuten109
Abbildung 51:	Kohärenzverläufe von Kunstkopf und binauralisiertem "Kugel- Array", gemittelt über fünf Sekunden, Ausschnitt 00:30 – 00:35110
Abbildung 52:	Kohärenzverläufe von Kunstkopf und binauralisiertem "Kugel- Array", gemittelt über fünf Sekunden, Ausschnitt 00:45 – 00:50110

Formelverzeichnis

1:	$\label{eq:Kreuzkorrelationsfunktion} Kreuzkorrelations funktion$	34
2:	Normierte Kreuzkorrelationsfunktion	34
3:	Normierte Kreuzkorrelationsfunktion für zeitdiskrete Signale	80
4:	Kohärenzgrad	81
5:	Kohärenzfunktion	81
6:	Nicht-quadrierte Kohärenzfunktion	82
7:	Diffusefield Image Predictor	83

Tabellenverzeichnis

Tabelle 1:	Zusammenhang von Hörereignismerkmalen und Schallfeldkomponenten (Theile et al., 2014).	24
Tabelle 2:	Lautsprecherkonfiguration D (4+5+0) nach ITU-R BS.2051-2 (ITU, 2018)	44
Tabelle 3:	Lautsprecherkonfiguration J (4+7+0) nach ITU-R BS.2051-2 (ITU, 2018)	46
Tabelle 4:	Mikrofonsysteme der Testaufnahmen	74
Tabelle 5:	Vorhersage des DFI-P mit Hilfe des Schoeps Image Assistant.	88
Tabelle 6:	DFI-P-Werte des Schoeps Image Assistant und der Berechnung für das Mikrofonsystem "Kugel-Array".	102
Tabelle 7:	DFI-P-Werte des Schoeps Image Assistant und der Berechnung für das Mikrofonsystem "Doppel-IRT".	103

Listingverzeichnis

Listing 1:	Funktionsaufruf der magnitude-squared-coherence in MATLAB	39
Listing 2:	Funktionsaufrufe zur Berechnung von Leistungsdichtespektren in MATLAB)0
Listing 3:	Berechnung der nicht-quadrierten Kohärenz in MATLAB)1
Listing 4:	Zählschleife und abschließende Berechnung des DFI-P in MATLAB)2

Abkürzungsverzeichnis

<i>Atmo</i>	
С	
DFI-P	Diffuse Field Image Predictor
EAS	Elements Of Auditory Setting
ESMA	Equal Segment Microphone Array
HRIR	
HRTF	
<i>IAC</i>	Interaural Cross-Correlation
IACC	Interaural Cross-Correlation Coefficient
<i>ICCC</i>	Interchannel Cross-Correlation
<i>ILD</i>	Interaural Level Difference
<i>ITD</i>	Interaural Time Delay
ITU	. International Telecommunication Union
L	
LEV	Listener Envelopment
LFE	Low Frequency Enhancement
Ls	Left Surround
Lss	Left Side Surround
Ltb	Left Top Back
Ltf	Left Top Front
Ltr	Left Top Rear
R	
Rs	Right Surround
Rss	Right Side Surround
Rtb	Right Top Back
Rtf	Right Top Front
Rtr	

1 Einleitung

Eine Atmo bspw. kann ein unglaublich dichtes, weil vielschichtiges Gebilde sein und sie offeriert – bei entsprechender Aufnahmetechnik – [...] einen, weiteren elementaren Baustein der Radioarbeit: den bereits angesprochenen Raum des Geschehens, der in seiner Eigenschaft alle in ihm stattfindenden Ereignisse zu reflektieren regelrecht geschwätzig erscheint und darüber hinaus dezidiert von seiner Beschaffenheit (als Größe, Einrichtung und akustische Bedingung der Möglichkeit) spricht (Hagelüken, 2009, S. 145).

1.1 Motivation und Zielsetzung

Immersive Audioproduktionen und 3D-Audio stellen derzeit ein polarisierendes Thema im Kontext audiovisueller Medien dar. Dabei ist die Thematik längst nicht mehr bloßer Diskussionsgegenstand eines Fachpublikums auf Messen und Kongressen, sondern beim Endkonsumenten angekommen. Dabei verspricht 3D-Audio das auditive Erlebnis einer Medienproduktion zu steigern, und den Rezipienten mit einem beeindruckenden räumlichen Klangerlebnis und imposanten Soundeffekten in die virtuelle Klangwelt der Produktion zu ziehen. 3D-Audio beschränkt sich dabei nicht auf eine Mediengattung, sondern kann im Kontext von Filmen, Hörspielen, Games und Virtual Reality zum Einsatz kommen.

Im Filmbereich und speziell im Kino hat sich 3D-Audio bereits seit längerem etabliert und es sind verschiedene Formate kommerziell vertreten (Dolby Atmos, Auro 3D, DTS:X). Die Entwicklung der Wiedergabetechnik und Distributionswege ist inzwischen so weit fortgeschritten, dass immersive Audioproduktionen direkt zum Endkonsumenten übertragen werden können. Neuere Entwicklungen in der Soundbar-Technologie ermöglichen es nun, 3D-Audio ins heimische Wohnzimmer zu bringen, ohne auf eine Installation einer Vielzahl von Lautsprechern angewiesen zu sein. Zudem ist es möglich, 3D-Audio-Inhalte für die Kopfhörerwiedergabe aufzubereiten und so dem Endkonsumenten verfügbar zu machen, ohne dass dieser besondere technische Voraussetzungen benötigt. Doch neben all den technischen Neuerungen sind es die Produktionen an sich, die dem Rezipienten den Mehrwert immersiver Audioproduktionen vermitteln können. Der Großteil der Endkonsumenten interessiert sich nicht für die technologische Komplexität, die sich hinter 3D-Audio verbirgt, sondern möchte möglichst unkompliziert eine immersive Medienproduktion erleben. Nur durch überzeugende Inhalte kann die Nachfrage nach immersiven Audioproduktionen bei einer breiten Masse initiiert werden und so der Thematik 3D-Audio langfristig zum Erfolg verhelfen.

Für den Tonschaffenden stellt sich die Frage, mit welchen technischen und gestalterischen Mitteln eine immersive Audioproduktion umgesetzt werden kann, um den Anforderungen des Rezipienten gerecht zu werden. Eine bedeutsame, wenn auch oftmals unterschätzte Rolle kommt in diesem Kontext den Atmosphären zu. Es hat sich etabliert, dass atmosphärische Klänge genutzt werden, um den Hörer auditiv an den Ort des medialen Geschehens zu versetzen. Dabei entfalten Atmosphären ihre Wirkung in der Regel subtil und agieren im Hintergrund. Ungeachtet dessen stellen sie einen elementaren Baustein der kreativen Tongestaltung dar. Unzählige Sound-Libraries enthalten eine Vielzahl anAtmosphärenaufnahmen, welche bei verschiedensten Produktionen zum Einsatz kommen und kommerziell vertrieben werden. Die Aufnahme von Atmosphären hat sich daher zu einem Spezialgebiet entwickelt. Der hohe Aufwand bei der Anfertigung hochwertiger Atmosphärenaufnahmen wird durch die mehrfache Nutzung in verschiedenen Produktionen amortisiert. Spezielle Atmosphären, die im Hinblick auf immersive Audioproduktionen erstellt werden, sind bislang bei den großen Anbietern kommerzieller Sound-Libraries jedoch kaum zu finden. Dabei bieten die aktuellen 3D-Audio-Technologien beste Voraussetzungen, um mit Hilfe spezifischer immersiver Atmosphären ein beeindruckendes räumliches Klangerlebnis zu schaffen. Die Herausforderung für den Tonschaffenden lässt wie folgt zusammenfassen:

From a field recording and mixing standpoint, there are some clear changes implied by these new immersive sound formats. For one thing most digital field recorders have no more than 6–8 channels, so some new technology is required there, including ways of capturing and storing the various metadata that may accompany the source material. New microphone designs also seem likely to be needed to handle all-round capture. All this leads to a lot of new creative options for sound design and editing, and tools are emerging to assist that process. There'll be more tracks to edit, keep track of, and deliver to the mix stage, though. One of the particular sound design challenges with any system is to create the impression of actually being in a particular space. It's one thing simply to pan sources around all these new loudspeakers [...] but another entirely to be able to create believable acoustic environments (Rumsey, 2013, S. 344).

Das Ziel dieser Thesis ist es, sowohl die technischen als auch gestalterischen Grundlagen von Atmosphären im Kontext immersiver Audioproduktionen zu beleuchten und die im Hinblick auf die praktische Anwendung benötigten Gesetzmäßigkeiten herauszuarbeiten. Darauf aufbauend werden Mikrofon- und Aufnahmetechniken betrachtet, mit denen Atmosphärenaufnahmen in 3D-Audio praktisch umgesetzt werden können. Der Fokus dieser Thesis liegt dabei auf Atmosphärenaufnahmen, die im Freien erstellt werden und als Bestandteil einer größeren Audioproduktion zu verstehen sind. Im Rahmen eines Feldversuchs werden verschiedene Mikrofonsysteme bei Atmosphärenaufnahmen im Freien erprobt. Durch die signaltechnische Untersuchung der Aufnahmen in Form einer Kohärenzanalyse soll überprüft werden, welche Möglichkeiten bestehen, um die gewünschten Abbildungseigenschaften eines Mikrofonsystems vorherzusagen. Ferner soll untersucht werden, ob es möglich ist, verschiedene Mikrofonsysteme anhand einer Kohärenzanalyse der aufgezeichneten Signale miteinander zu vergleichen. So können dem Anwender wertvolle Hinweise für die Planung von Mikrofonsystemen und die praktische Umsetzung von Atmosphärenaufnahmen vermittelt werden.

1.2 Aufbau der Arbeit

In **Kapitel 2** werden zunächst die wichtigsten Begriffe im Zusammenhang mit Atmosphären in der Audioproduktion erklärt und definiert. Dabei wird auch auf besondere Unterschiede zwischen dem deutschen und englischen Sprachraum eingegangen. Darüber hinaus werden verschiedene Kategorisierungsmodelle vorgestellt, die es ermöglichen, Atmosphären systematisch strukturieren und abgrenzen zu können. **Kapitel 3** untersucht die Bedeutung von Atmosphären im Sounddesign-Prozess. Es wird herausgestellt, welche Funktionen von Atmosphären als integrativer Bestandteil einer komplexen Tongestaltung übernommen werden. Wichtige Merkmale, welche eine Atmosphäre zur Erfüllung ihrer Funktion aufweisen muss, werden aufgezeigt.

Kapitel 4 behandelt das Thema Immersion. Anhand aktueller Forschungsarbeiten wird eine Begriffsdefinition vorgeschlagen und es erfolgt eine Einordnung des Begriffs in den Kontext audiovisueller Medienproduktionen. Ferner werden die auditiven Dimensionen der Immersion herausgearbeitet und es wird dargelegt, welche Rolle Atmosphären im Kontext immersiver Medienproduktionen übernehmen.

In **Kapitel 5** werden diejenigen Aspekte der Psychoakustik näher beleuchtet, die im Zusammenhang mit der Aufnahme, dem Einsatz in der Postproduktion und der Wiedergabe von Atmosphären von besonderer Bedeutung sind.

Kapitel 6 behandelt das Thema Wiedergabesysteme. Ausgewählte Systeme zur Wiedergabe immersiver Audioproduktionen werden im Detail vorgestellt. Darüber hinaus wird die Bedeutung der einzelnen Wiedergabekanäle im Hinblick auf die Darstellung möglichst umhüllender Schallfelder analysiert. Den Abschluss des Kapitels bildet die Vorstellung einer virtuellen Abhörumgebung, die für weitere Untersuchungen im Rahmen dieser Arbeit entwickelt wurde.

Kapitel 7 behandelt das Thema Mikrofon- und Aufnahmetechnik. Es werden grundsätzliche Anforderungen an die Aufnahmetechnik für Atmosphären vorgestellt und die wichtigsten technischen Merkmale hervorgehoben. Zusätzlich werden exemplarisch spezifische Mikrofonsysteme präsentiert, die für den Einsatz bei Atmosphärenaufnahmen entwickelt wurden.

Kapitel 8 stellt die praktischen Untersuchungen dieser Arbeit vor. Die Durchführung verschiedener Atmosphärenaufnahmen wird dokumentiert und die eingesetzte Mikrofonund Aufnahmetechnik vorgestellt. In einem weiteren Schritt werden die Aufnahmen signaltechnisch analysiert. Dazu wird zunächst die Analysemethodik und deren Umsetzung in MATLAB vorgestellt. Die Auswertung und Diskussion der Analyseergebnisse schließen das Kapitel ab.

Kapitel 9 dient als Fazit und zeigt in einem Ausblick zukünftige Forschungsthemen auf, die im Zusammenhang mit Atmosphären in immersiven Medienproduktionen relevant sind.

Aus Gründen der besseren Lesbarkeit wird in dieser Arbeit an einigen Stellen die Sprachform des generischen Maskulinums verwendet. Es wird darauf hingewiesen, dass sämtliche Personenbezeichnungen gleichermaßen für alle Geschlechter gelten.

2 Definition und Kategorisierung atmosphärischer Klänge

[...] it is exhilarating and revealing to learn to listen to the voices of whole environments as structured and informative narratives and to discover what they can tell us as we begin to explore the natural world with more attentive and open ears (Krause, 2016, S. 38).

In diesem Kapitel werden die relevanten Begriffe im Zusammenhang mit Atmosphären in der Audioproduktion definiert und erklärt. Abgrenzungen gängiger Termini sollen dem besseren Verständnis dieser Arbeit dienen. Zusätzlich wird eine Kategorisierung verschiedener Atmosphären-Klänge vorgenommen, um eine systematische Grundlage zur Auswahl geeigneter Mikrofonaufnahmetechniken zu etablieren.

2.1 Begriffsdefinitionen

Im Zusammenhang mit Atmosphären in Audioproduktionen gibt es eine Vielzahl an Ausdrücken, die mitunter synonym verwendet werden. Darüber hinaus werden in Deutschland, Großbritannien und den USA teilweise verschiedene Bezeichnungen genutzt oder ähnliche Begriffe haben unterschiedliche Bedeutungen. Friesecke (2014) definiert Atmosphäre, kurz Atmo, allgemein als akustische Beschreibung einer Szene. Diese Begriffsbestimmung ist jedoch zu unspezifisch, um notwendige Charakteristika zur Ableitung einer Mikrofonaufnahmetechnik zu enthalten. Görne (2017) definiert Atmosphären spezifischer als klanglichen Hintergrund. Diese Erklärung weist somit auf die hintergründige Funktion einer Atmosphäre im Sounddesign-Prozess hin. Eine Atmosphäre stellt also einen akustischen Hinter- bzw. Untergrund dar, auf dem vordergründig weitere Klänge bestehen können. Butzman & Martin (2012) sprechen in diesem Zusammenhang von der Kulissenhaftigkeit einer Atmosphäre. Sie beziehen sich damit auf den, vornehmlich im Hörspiel gängigen, Begriff der Geräuschkulisse, um die Klänge des Hintergrunds zu beschreiben. Atmosphären sind weiterhin Klänge eines Ortes. Sie sollen dem Rezipienten einen Eindruck oder ein Gefühl von diesem Ort über den Hörsinn vermitteln (Viers, 2008).

Während in Großbritannien der Begriff atmosphere sounds, kurz atmos, genutzt wird, sind im US-amerikanischen Raum die Begriffe ambience und background üblich (Wyatt & Amyes, 2005). Die Bezeichnung background unterstreicht ebenfalls den Hintergrundcharakter von Atmosphären. Dieser Hintergrundcharakter ergibt sich durch Diffusität des Schalls einer Atmosphäre. Zusätzlich stellen Görne (2017) und Holman (2010) Stereophonie² als charakteristisches Merkmal einer Atmosphäre heraus. Dies bekräftigt den Bezug der Atmosphäre zu einem Ort oder Raum, weshalb ein räumliches Tonverfahren zum Einsatz kommt. Darüber hinaus zeichnet sich eine Atmosphäre durch ihre Konsistenz aus und stellt eine eigene, künstliche Tonspur dar, die unabhängig vom Set-Ton etabliert wird (Holman, 2010).

Eine weite Auffassung des Begriffs Atmosphäre findet sich bei Lensing (2017). Nach ihm sind Atmosphären "vordergründig allgemeine akustische Soundscape-Beschreibungen eines bestimmten Ortes (...)" (Lensing, 2017, S. 50). Weiterhin finden bei ihm die Begriffe Grundatmo und Set-Atmo Anwendung. Um eine deutliche Unterscheidbarkeit zur obigen Definition des Atmosphären-Begriffs zu ermöglichen, ist es sinnvoll genauer zu betrachten, was unter Grund- bzw. Set-Atmo zu verstehen ist und welche Funktion sie erfüllen sollen. Gemeint ist das originäre Schallfeld, welches am Ort und zur jeweiligen Zeit einer Aufnahme präsent war. Dies ist beispielsweise die Gesamtheit aller Hintergrundklänge, die während einer Aufzeichnung am Set zu hören sind. Folglich ist dieser Hintergrund bei allen Dialogaufnahmen vorhanden. Um beim Dialogschnitt auftretende Lücken füllen zu können, wird eine separate Aufnahme dieses Hintergrundklangs benötigt (Görne, 2017). Diese Aufnahme muss mit dem gleichen Mikrofon-Setup erfolgen, welches auch zur Dialog-Aufnahme genutzt wurde, um unterschiedliche Klangfarben zu vermeiden und so einen unmerklichen Dialogschnitt zu ermöglichen. In den meisten Fällen handelt es sich daher um eine monophone Aufnahme. Görne (2017) nutzt in diesem Zusammenhang den weniger geläufigen Begriff raumstatischer Ton und orientiert sich damit an der US-amerikanischen Bezeichnung room tone. Des Weiteren sind die Bezeichnungen (production) fill oder presence üblich (Holman, 2010), mit denen sich die akustischen Gegebenheiten am Drehoder Aufnahmeort genau von den Atmosphären, die im Rahmen des Sounddesign-Prozesses

 $^{^2}$ Stereophonie ist hier allgemein als Mehrkanaligkeit zu verstehen.

eingesetzt werden und unabhängig vom Drehort sind, unterscheiden lassen. Ebenso werden die Begriffe *ambient noise* oder *background noise* genutzt, um die – meist unerwünschten – Hintergrundgeräusche am Aufnahmeort zu beschreiben (Holman, 2010).

Der von Lensing (2017) verwendete Begriff Soundscape soll ebenfalls detaillierter betrachtet werden, da dieser eine besondere Stellung im Zusammenhang mit Atmosphären im Sounddesign-Prozess einnimmt. Der Begriff wurde maßgeblich von Murray Schafer (1993) geprägt und findet breite Anwendung in der akustischen Ökologie. Unter einem soundscape wird allgemein die akustische Umwelt verstanden. Dies kann sich auf einen natürlichen Ort beziehen, schließt aber ebenso künstlich erzeugte, komponierte Klangwelten mit ein. Daher bietet sich die Verwendung des Begriffs Soundscape auch im Rahmen audiovisueller Medienproduktionen an. Der Soundscape-Begriff bietet die Möglichkeit einer feineren Differenzierung der verschiedenen Gattungen und Funktionen atmosphärischer Klänge. Nach Schafer (1993) besteht ein soundscape aus verschiedenen Typen von Klängen. Man unterscheidet zwischen keynote sounds, signal sounds, soundmarks und archetypical sounds. Ein Keynote sound ist als Anker- oder Referenzpunkt zu allen anderen Klängen innerhalb eines soundscape zu betrachten; er bildet den Grundton. Er ist allgegenwärtig und relativ konstant, wodurch er meist nur unterbewusst wahrgenommen wird, jedoch große Auswirkungen auf die Stimmung einer Szene haben kann (Sonnenschein, 2013). Dabei ist er unmittelbar mit der Umwelt der Szene oder Handlung verbunden. Übertragen auf die auditive Ebene bildet der keynote sound im gestaltpsychologischen Sinn den Hintergrund für eine Figur im Rahmen der Figur-Grund-Beziehung. Signal sounds werden dagegen ganz bewusst wahrgenommen und stehen für die Figur (Schafer, 1993). Soundmarks können als spezielle Form eines signal sounds betrachtet werden. Sie weisen auf einen bestimmten Ort hin und bilden ein akustisches Wahrzeichen, daher die Ähnlichkeit zum Wort landmark. Dies kann beispielweise eine einzigartige Kirchenglocke sein, welche nur in einem bestimmten Ort zu hören ist und deshalb besondere Aufmerksamkeit von den Bewohnern dieses Ortes erfährt (Schafer & Breitsameter, 2010). Allgemeiner betrachtet können im Sounddesign viele Klänge im Sinne einer soundmark eingesetzt werden. Verantwortlich dafür ist das affektive Gedächtnis. Der Mensch lernt, durch Erfahrung bestimmte Geräusche mit Orten zu verknüpfen (Friesecke, 2014). Der Ruf einer Möwe weist zum Beispiel sofort auf eine Küstenregion hin. Archtypical Sounds beziehen sich auf Klänge, die Ur-menschliche

Reaktionen und Erinnerungen im Rezipienten hervorrufen (Schafer, 1993) und somit universell verstanden werden. Hierzu zählen beispielsweise typische Naturklänge wie Blitz und Donner, welche die Empfindungen Gefahr und das Verlangen nach Schutz hervorrufen. Durch den gezielten Einsatz solcher Klänge können Stimmungen bewusst erzeugt werden.

Zusammengefasst wird im Rahmen dieser Arbeit unter dem Begriff Atmosphäre ein relativ diffuser, stereophoner Klang verstanden, der sich durch einen hintergründigen Charakter auszeichnet. Er bildet die Kulisse oder das Fundament einer akustischen Szene und stellt einen Bezug zu Ort und Raum her. Eine Atmosphäre wird demnach als elementarer Baustein im Rahmen des Sounddesign-Prozesses verstanden. Eine vom Sounddesigner detailliert gestaltete akustische Szene wird in dieser Arbeit als *soundscap*e bezeichnet, ebenso die natürliche Klangszene eines Ortes, an dem eine Tonaufnahme erfolgen soll. Nach Schafers (1993) Terminologie stellt die Atmosphäre folglich einen *keynote sound* innerhalb der *soundscape* dar.

2.2 Kategorisierungsmodelle

Um eine geeignete Mikrofonaufnahmetechnik auswählen zu können, ist es zunächst erforderlich, festzulegen, was genau Gegenstand der Tonaufnahme sein soll. Es lassen sich unzählige Varianten natürlicher Soundscapes erdenken. Darüber hinaus wird es am selben Ort zu verschiedenen Zeiten niemals gleich klingen. Natürliche Soundscapes sind dynamisch. Sie sind beispielweise abhängig von Jahres- und Tageszeit, Wetterbedingungen sowie von Einwirkungen durch Menschen, Fauna und Flora. Sowohl die Beschaffenheit und Geometrie eines Raumes oder Platzes als auch Schallreflexionen und -beugungen sind prägend. Das Soundscape ist folglich ein komplexes Zusammenspiel verschiedener Variablen mit unterschiedlichen akustischen Qualitäten. Eine Untergliederung des Soundscapes in verschiedene akustische Wirkungsbereiche soll dabei helfen, die wichtigsten Komponenten systematisch voneinander zu trennen. Dadurch kann die Frage nach einem geeigneten Mikrofonsystem leichter beantwortet werden, denn die jeweiligen Komponenten übernehmen unterschiedliche Funktionen und stellen daher verschiedene Anforderungen an ein Mikrofonsystem. Es existiert bereits eine Vielzahl an Vorschlägen unterschiedlicher Autoren für die Kategorisierung von Soundscapes. Dabei wird deutlich, dass sich diese durch die unterschiedlichen Perspektiven der Autoren voneinander unterscheiden.

Wittek (2012) arbeitet mit einem Ebenen-Modell und teilt das Soundscape dazu in drei Bereiche ein. Hierbei steht die Differenzierung einzelner Bestandteile eines Soundscape im Vordergrund. Jeder Stufe werden charakteristische Signalmerkmale zugeteilt und pro Ebene werden geschlossene und offene Räume betrachtet. Die erste Ebene enthält diffuse Klänge, die nicht lokalisiert werden. Hierzu zählen beispielweise Nachhall und stochastische Signale wie Wind, Blätterrascheln, entfernter Verkehrslärm, etc. Die Diffusität ergibt sich dadurch, dass keines der Signale dieser Ebene aus einer bestimmten Richtung lokalisiert werden kann. Sie kommen gleichzeitig aus vielen, unterschiedlichen Richtungen. Die zweite Ebene enthält Signale, die an sich zwar diskret und damit lokalisierbar sind, aber deren genaue Einfallsrichtungen für die Wahrnehmung des gesamten Soundscapes nicht von Bedeutung sind. Oftmals treten diese Signale konzentriert auf. Dies können beispielsweise frühe Reflexionen, die Laute eines Vogelschwarms oder das Grillenzirpen auf einer Wiese sein. Wiedergabeseitig sollten diese Signale diskret oder über Phantomschallquellen abgebildet werden, um lokalisiert werden zu können. Der dritten Ebene werden schließlich diejenigen Signale des Soundscape zugeordnet, welche zwingend eine korrekte Richtungsabbildung benötigen. Von besonderer Bedeutung ist dies bei der parallelen Aufzeichnung von Bild und Ton, um eine einheitliche Perspektive zu gewährleisten oder bei sich bewegenden Quellen, um eine flüssige Bewegungswiedergabe zu ermöglichen.

Camerer (2000) unterscheidet bei seiner Kategorisierung zwischen vier Bereichen. Diese Einteilung orientiert sich an den unterschiedlichen Vorzugsrichtungen die ein Soundscape aufweisen kann. Bei dieser Perspektive auf ein Soundscape steht der strukturelle Aufbau und die akustische Beschaffenheit des Aufnahmeortes im Vordergrund. Die erste Kategorie bilden die Hintergrund- bzw. Fundament-Atmos. Das Soundscape bildet ein homogenes Schallfeld. Die Signale sind diffus und zeichnen sich durch ihren Hintergrundcharakter aus. Es gibt keine Vorzugsrichtung, die Signale kommen aus vielen verschiedenen Richtungen und lassen sich nicht lokalisieren. Diese Kategorie deckt sich mit der ersten Ebene aus Witteks Modell. Gerichtete Atmos stellen die zweite Kategorie dar. Hierunter fallen Soundscapes, die sich durch einen akustischen Kontrast auszeichnen. Dies kann beispielsweise der Übergangsbereich zwischen zwei sehr unterschiedlichen Soundscapes sein. Verschiedene Richtungen bilden hier einen Akzent aus, weshalb die Perspektive sehr wichtig ist und die Richtungsabbildung an Bedeutung gewinnt. Beispiele für Soundscapes der Kategorie zwei sind Waldränder, Flussufer, Küstenwälder oder Strände. Die dritte Kategorie wird Effekt-Atmos genannt. Hierzu zählen Soundscapes, die einen diskreten akustischen Schwerpunkt aufweisen. Diese sollen mitsamt ihrem Hintergrund perspektivisch korrekt und lokalisierbar erfasst werden. Eine Autohupe im Stadtverkehr, der Start eines Flugzeugs am belebten Flughafen oder ein auffälliger Vogelruf im Wald sind mögliche Beispiele für diese Kategorie. Sie setzt sich somit aus Bestandteilen zusammen, die Wittek (2012) in seinem Modell den Ebenen eins und drei zuordnet. Die vierte Kategorie in Camerers Modell sind Effekte. Dabei geht es um die nahe Perspektive einer weit ausgedehnten Quelle oder die korrekte Abbildung der Bewegung einer Schallquelle. Dies stellt gewissermaßen den Grenzbereich zwischen einem Soundscape und einer einzelnen Schallquelle dar. Man könnte auch sagen, eine Schallquelle wird in Form eines Makro-Soundscapes erfasst. Beispiele sind industrielle Maschinen oder Fahr- und Flugzeuge.

Der Bioakustiker Krause (2016) unterscheidet in einem Soundscape drei Bereiche und bezieht sich dabei ausschließlich auf Outdoor-Soundscapes. Bei seiner Betrachtungsweise steht der Ursprung des Schalls im Vordergrund. Den ersten Bereich nennt er *geophony*. Dieser umfasst diejenigen Signale eines Soundscapes, welche nicht biologischen Ursprungs sind. Sie werden hauptsächlich durch Wetter und Beschaffenheit der Topologie und Geologie bestimmt. Hierunter fallen Klänge wie Wind, Regen, Gewässer. Der zweite Bereich umfasst alle Schallquellen, die biologischen Ursprungs sind. Dies sind also Geräusche, die von der Tier- und Pflanzenwelt ausgehen. Dementsprechend nennt Krause diese Kategorie *biophony*. Alle Klänge, die durch Menschen erzeugt werden, bilden die dritte Kategorie – die *anthrophony*. Diese umfasst daher sämtlichen Schall, der von Geräten, Maschinen, der menschlichen Infrastruktur oder urbanen Räumen ausgeht.

Eine weitere Kategorisierung wurde von Schafer (1993) vorgestellt und bereits in Kapitel 2.1 erwähnt. Diese gliedert ein Soundscape nach der Bedeutung der unterschiedlichen Klänge. So wird zwischen *keynote sound, sound signal, soundmark* und *archytipical sound* unterschieden.

Diese vier Kategorisierungsmodelle zeichnen sich durch eine individuelle Perspektive auf das Soundscape aus und besitzen jeweils eigene Qualitäten. Im Hinblick auf die Zielsetzung dieser Arbeit gilt es zu betrachten, welche Funktion das jeweilige Modell erfüllt und wann es für den Einsatz von Atmosphären in immersiven Medienproduktionen von Nutzen sein kann. Abbildung 1 visualisiert die vier Modelle und stellt ihre jeweiligen Kategorien auf Kreisbahnen dar, um ihre akustischen Eigenschaften herauszustellen. Grundsätzlich gilt es zu beachten, dass die einzelnen Bereiche in der Regel nicht völlig isoliert vorkommen und ihre Übergänge fließend sind. Wittek (2012) nimmt mit seinem Modell eine technische bzw. signalorientierte Perspektive ein. Er fragt: Welche physikalischen Eigenschaften besitzen die Signale im Soundscape und welche Auswirkungen müssen auf der Wiedergabeseite beachtet werden? Diese Herangehensweise hilft dabei, ein grundlegendes Verständnis der benötigten Mikrofonaufnahmetechnik für Atmosphären zu entwickeln. Camerer (2000) nimmt in seinem Modell eine praxisorientiertere Perspektive ein. Er fragt: Wie sehen die Bedingungen am Aufnahmeort aus? Wo wird das Mikrofon-Setup aufgestellt und was gilt es dabei zu beachten? Diese Sichtweise ist somit vor allem bei der praktischen Durchführung einer Atmosphärenaufnahme nützlich und lenkt den Blick und das Ohr auf die konkreten Gegebenheiten vor Ort. Krause (2016) konzentriert sich auf die Schallquellen. Er fragt: Wer ist für den Klang verantwortlich? Wo entsteht er und unter welchen Umweltbedingungen? Dieser Blick auf das Soundscape kann im Vorfeld einer Aufnahme dienlich sein. Der Fokus wird auf das kritische Hören des Soundscapes gelegt und ermöglicht so erst das Erkennen interessanter Aufnahmeorte und akustischer Perspektiven. Schafers (1993) Modell ist auf die Bedeutung der Klänge ausgerichtet. Er fragt: Welche Wirkung hat ein Klang? Was vermittelt oder symbolisiert er? Diese Herangehensweise ist im Sounddesign-Prozess relevant. Um zielgerichtet die Tonspur einer Medienproduktion gestalten zu können, ist es erforderlich, die Funktion und Bedeutung unterschiedlicher Klänge zu kennen und diese bewusst einzusetzen.



 $Abbildung \ 1: \ Sounds cape \ Kategorisierungs modelle.$

3 Bedeutung und Funktion von Atmosphären im Sounddesign-Prozess

Remember my philosophy about backgrounds: Dialog will come and go, music cues will come and go, Hard FX will come and go, but backgrounds are heard all the time, even if it is extremely subtle. The stereophonic envelope of the ambience is crucial, and I give it utmost attention (Yewdall, 2012, S. 353).

Soundeffekte – und damit auch Atmosphären – sind integraler Bestandteil des Storytellings einer Medienproduktion. Im Rahmen eines Sounddesign-Prozesses sind Atmosphären als Bestandteil oder Baustein einer größeren akustischen Szene zu begreifen. Daher stehen sie nicht für sich, sondern sind stets im Kontext einer Medienproduktion und damit anderer Klänge, Bilder und Erlebnisräume zu betrachten. Eine Atmosphärenaufnahme stellt somit kein eigenes, vollständiges Werk dar. Sie steht nicht für sich und muss deshalb spezielle Anforderungen erfüllen, um für die Tongestaltung einer Medienproduktion genutzt werden zu können. Im Rahmen dieser Thesis wird davon ausgegangen, dass Atmosphärenaufnahmen eine gewisse Allgemeingültigkeit besitzen sollen und ihre Nutzung nicht ausschließlich auf eine bestimmte Medienproduktion oder -gattung begrenzt ist. Vielmehr sollen Atmosphären in Form einer Bibliothek für eine große Bandbreite immersiver Medienproduktionen, wie beispielweise Filme, Hörspiele, Games und Virtual Reality Experiences, einsetzbar sein. Die Aufgabe des Sounddesigners ist es, nach funktionalen, ästhetischen und dramaturgischen Gesichtspunkten eine Klangwelt zu gestalten. Denn die akustische Umwelt ermöglicht dem Rezipienten den Zugang und die Verbindung zur akustischen Realität der dargestellten Handlung (Kock, 2019). Stevens und Raybold (2011) vergleichen den Sounddesign-Prozess mit dem Malen eines Gemäldes – sorgfältiges Übereinanderschichten von Vorder-, Mittelund Hintergrund ergeben ein eindrucksvolles Gesamtbild. Das Soundscape wird letztlich im Sounddesign-Prozess aus vielen Einzelklängen spezifisch für die Anforderungen der Produktion komponiert. Um dem Sounddesigner die maximale Kontrolle über die gezielte Gestaltung der akustischen Szene zu ermöglichen, ist es erforderlich, dass die benötigten Klänge oder Klangschichten separat voneinander vorliegen. Diese Einzelklänge nennt Chion (2019, S. 52–53) Elemente der Tonszenerie (elements of auditory setting (EAS)).

Dies unterstreicht, weshalb es sinnvoll ist, Atmosphären als ein Element im Kontext einer größeren akustischen Szene zu begreifen. Diese Definition ist auch bedeutsam für die Aufnahmetechnik natürlicher Atmosphären.

Eine Schlüsselfunktion von Atmosphären ist die akustische Darstellung von Orten. Sie dienen als klanglicher Hintergrund, auf dem mit Hilfe von konkreten Klängen ein detailliertes Soundscape gestaltet werden kann. Diese Darstellung gibt dem Rezipienten Auskunft über Raum und Zeit. Atmosphären leisten in diesem Zusammenhang eine Orientierungsfunktion (Sonnenschein, 2013). Görne (2017, S. 216) spricht hier von "Establishing Sounds". Diese sind in ihrer Funktion vergleichbar mit einem Establishing Shot der Kamera zu Beginn einer Filmszene, die einen neuen Schauplatz vorstellt (Whittington, 2007). Mit Hilfe der akustischen Darstellung des Handlungsortes wird versucht, den Rezipienten in die dramaturgische Welt einzuführen (Wyatt & Amyes, 2005). Die glaubhafte Ausgestaltung akustischer Räume ist hierzu notwendig. Spezifische Umweltgeräusche sowie die Raumgeometrie, Moden und Reflexionsmuster tragen zur unverwechselbaren Akustik eines Ortes bei. Die Umgebungsgeräusche vermitteln Stimmung und "Gefühl" eines Raumes. Die vorhandenen Schallquellen erschaffen im Zusammenspiel mit Schallreflexion und beugung eine konstante Charakteristik (Yewdall, 2012). Insbesondere in non-linearen Medien wie Games oder Virtual Reality kommt der Orientierungsfunktion große Bedeutung zu. Mediennutzer können sich frei bewegen und sind selbst in der akustischen Szene situiert. Räumliche Atmosphären tragen wesentlich dazu bei, dass sich ein Gefühl des "Vor Ort seins" einstellt (Zdanowicz & Bambrick, 2020).

Der Hintergrundcharakter von Atmosphären ist für ihren Einsatz im Sounddesign-Prozess von großer Bedeutung. Der Informationsgehalt einer akustischen Szene muss für den Rezipienten unmittelbar wahrnehmbar sein, ohne besondere Aufmerksamkeit zu erfordern. Idealerweise sind Atmosphären in hohem Maß abstrahiert und auf ein Mindestmaß an Information reduziert (Flückiger, 2007). Sie bilden den Grund einer Figur-Grund-Beziehung. Die Aufnahme eines komplexen Soundscapes mit vielen diskreten Klänge, lässt sich nur bedingt im Sounddesign-Prozess verwenden, da hervortretende Geräusche die Aufmerksamkeit der Rezipienten auf sich ziehen und damit der beabsichtigten Steuerung auf wesentliche Bestandteile der Dramaturgie entgegenwirken (Wyatt & Amyes, 2005). Ferner kann die Aufnahme eines dichten Soundscapes nicht mehr an das dramaturgische Setting angepasst werden, wodurch die Einsatzmöglichkeiten einer solchen Aufnahme stark eingeschränkt ist. Raffaseder (2010) unterscheidet beim Einsatz von Atmosphären zwischen den gegenläufigen Anforderungen der Generalisierung und Differenzierung. Das bedeutet zum einen, dass Atmosphären einen hintergründigen, indirekten Charakter aufweisen müssen, um vielfältig im Sounddesign eingesetzt werden zu können. Zum anderen bedarf es dennoch einer großen Palette atmosphärischer Klangfarben, um unterschiedlichste Orte akustisch interessant und abwechslungsreich darstellen zu können. Die Abstraktion von Atmosphären ist für den Einsatz in Games und Virtual Reality besonders relevant. Da hier der zeitliche Ablauf nicht vorbestimmt ist, müssen Atmosphären in Schleifen abgespielt werden. Auffällige, diskrete Klänge in der Atmosphäre fallen bei Wiederholung deutlich auf und verhindern damit eine glaubwürdige akustische Szene (Stevens & Raybould, 2011). Die Komplexität einer Atmosphäre muss folglich so reduziert sein, dass sie beliebig wiederholt werden kann, ohne dass dies bemerkt wird. Notwendige diskrete Klänge werden in Form von Sound-Objekten in der virtuellen Welt platziert. Diese werden zufällig abgespielt und reagieren auf den Abstand und die virtuelle Umgebung des Mediennutzers. Getrennte Atmosphären und Soundobjekte ermöglichen so die Gestaltung eines dynamischen Soundscapes (Marks, 2017). Flückiger (2007) und Viers (2008) sprechen in diesem Zusammenhang von einer Stereotypisierung der Atmosphäre. Dies stellt ein Qualitätsmerkmal von Atmosphären dar, denn "[...] gerade wegen ihrer unterschwelligen, unbewussten Wahrnehmung prägen sie die von den entsprechenden Orten ausgehende bzw. dort empfundene Stimmung ganz entscheidend mit und sind für die Gesamtwirkung einer Medienproduktion von großer Bedeutung" (Raffaseder, 2010, S. 242).

Atmosphären übernehmen überdies eine wichtige Funktion für die dramaturgische Kontinuität. Der Einsatz einer konsistenten Atmosphäre sorgt für geografische und zeitliche Kontinuität im dramaturgischen Raum über Bildschnitte hinweg. Flückiger (2007) stellt diese verbindende Wirkung als eine der grundlegendsten Funktionen heraus. Eine einheitliche Tongestaltung schafft Kohärenz und verbindet fragmentarische visuelle Ausschnitte zu einem Ganzen. Dies ermöglicht dem Rezipienten einerseits die räumlichzeitliche Orientierung innerhalb der Handlung, andererseits auch die strukturelle Orientierung in Szenen und Sequenzen (Flückiger, 2007). Ändert sich die Atmosphäre über einen Bildschnitt hinweg nicht, so bedeutet dies für den Rezipienten, dass sich zwar die Perspektive auf das Geschehen geändert hat, Ort und Zeit der Handlung jedoch nicht. Dieses Wirkprinzip kann auch dazu genutzt werden einen Orts- oder Zeitsprung zu kommunizieren, eine Szene zu beenden oder eine neue einzuleiten. Der merkliche Wechsel der Atmosphäre wird als Szenenwechsel verstanden, auch wenn dies (noch) nicht über andere Wahrnehmungskanäle kommuniziert wurde. Durch einen Versatz von Bild- und Atmosphärenwechsel kann ein bevorstehender Orts- oder Zeitsprung angekündigt werden. Der Überhang der Atmosphäre in eine neue Szene sorgt für eine zusätzliche dramaturgische Verknüpfung der Szenen (Holman, 2010). Die verbindende Funktion der Atmosphäre ist auch in technischer Hinsicht von Bedeutung. Im Dialogschnitt kann es zu auffälligen Sprüngen des raumstatischen Tons kommen, wenn die Aufnahmen mit unterschiedlichen Mikrofonabständen und -winkeln durchgeführt wurden. Mit Hilfe von Atmosphären können diese Schnitte maskiert und die Kontinuität damit gewahrt werden (Sonnenschein, 2013), sie fungieren in diesem Fall als akustischer Klebstoff oder "auditory perfume" (Holman, 2010, S. 163).

Eine weitere Funktion übernehmen Atmosphären für das Storytelling einer Medienproduktion. Lensing (2017) beschreibt dies als dramaturgisch-kompositorische Funktion. Hierzu kann sich der Sounddesigner der Bedeutungsebenen von Atmosphären bedienen. So ist es beispielsweise auch möglich, über Atmosphären das soziale Milieu, in dem eine Handlung stattfindet, zu präsentieren. Besonders wirkungsvoll ist die Nutzung von Klängen, die Schafer (1993) archetypical sounds nennt. So kann beispielsweise die ruhige Atmosphäre eines Sommerabends mit Grillenzirpen eine friedvolle Stimmung vermitteln, während ein herannahendes Gewitter unheilvoll und bedrohlich wirkt (Holman, 2010). Wetteränderungen können so genutzt werden, um eine zeitliche Entwicklung der Geschichte widerzuspiegeln. DeLouise und Ottenritter (2020) unterstreichen, dass besonders in immersiven Medienproduktionen subtile Atmosphären eine eindrucksvolle Wirkung beim Rezipienten erzielen können. Sie merken weiterhin an, dass der kreative Einsatz von Atmosphären nicht auf fiktionale Produktionen beschränkt ist. Auch für dokumentarische Formate können Atmosphären gewinnbringend eingesetzt werden, sofern die Glaubhaftigkeit Tongestaltung erhalten bleibt. Sonnenschein (2013) spricht sich in diesem der Zusammenhang für die Nutzung von funktionalen Klängen aus, die nicht zwingend einer

authentischen Quelle entstammen müssen. In audiovisuellen Medien sollten Bild und Ton keinen Widerspruch erzeugen (Bild–Ton–Schere) (Kamp, 2007), sofern dies nicht bewusst als Stilmittel eingesetzt wird. Marks (2017) stellt heraus, dass Szenen oftmals erst dann glaubwürdig erscheinen, wenn entsprechende Atmosphären Verwendung finden. So kann beispielsweise eine Szene im Weltall glaubhaft vertont sein, obwohl eine Schallausbreitung gar nicht möglich ist (Zdanowicz & Bambrick, 2020). Darüber hinaus können Atmosphären dazu dienen, die dramaturgische Entwicklung unbewusst zu kommunizieren. Eine passende Atmosphäre, die in unterschiedlichen Intensitäten vorliegt, kann genutzt werden, um den Verlauf der Geschichte durch die Tongestaltung zu unterstützen. Ebenso kann durch das Verschieben der zeitlichen Dimensionen von Bild- und Atmosphärenübergängen Spannung erzeugt werden (Sonnenschein, 2013).

Die bisherigen Betrachtungen bezogen sich auf den Einsatz von Atmosphären innerhalb des diegetischen Raumes einer Erzählung. Atmosphären können jedoch auch durch Verfremdung in den extradiegetischen Raum überführt werden. Dies ermöglicht es, eine neue Bedeutungsebene zu etablieren und mit Hilfe von Atmosphären die subjektive Perspektive von Figuren einzunehmen und Stimmungen und Emotionen auszudrücken (Whittington, 2007). Atmosphären übernehmen hier eine Subjektivierungs-Funktion. Murray (2019) stellt in diesem Zusammenhang heraus, dass Atmosphären als Vehikel zum Einsatz kommen, um Stimmungen oder Gefühle zu vermitteln, die weder im Bild noch im bildsynchronen Ton verankert sind.

Besonders im Rahmen immersiver Audioproduktionen können Atmosphären dazu verwendet werden, den dreidimensionalen akustischen Raum dramaturgisch zu gestalten. In Filmproduktionen kann dadurch der Bildrahmen akustisch erweitert werden. Diese Erweiterung ermöglicht es, den Ton als eigene Erzählebene zu etablieren, der das Bild nicht nur untermalt, sondern insbesondere das kommuniziert, was im Bild nicht zu sehen ist. Dies wird auch durch den Sounddesigner Rydstorm bestätigt: "The good filmmakers [...] are going to use the camera or visual to tell you one thing and the sound to tell you something maybe completely different, like the off-screen world" ((o. J.), zitiert nach Sonnenschein, 2013, S. 170). Chion (2019) prägte in diesem Zusammenhang den Begriff *Extension* und beschreibt damit den Grad, um den der Klangraum den Bildraum erweitert. Mit Hilfe einer variablen Gestaltung der Extension kann die Relation des Rezipienten zum Werk gesteuert werden. Eine geringe Extension evoziert eine distanzierte Haltung, wohingegen große Extension dafür sorgt, dass der Rezipient vom diegetischen Klangraum umgeben wird und sich somit involviert fühlt (Holman, 2010). Die Möglichkeiten der Extension bestanden erstmals mit dem Aufkommen des Surround-Tons und haben durch immersive Audiosysteme weiter zugenommen. Durch geeignete Atmosphären kann die Extension weitreichend gestaltet werden, was den Sounddesign-Prozess um ein wirkungsvolles Ausdrucksmittel bereichert.

Zusammenfassend spielen Atmosphären im Sounddesign-Prozess eine tragende Rolle. Sie übernehmen wichtige Funktionen für die Dramaturgie, die technische Tonqualität und die räumliche Wirkung von Medienproduktionen. Um diese Funktionen erfüllen zu können, müssen Atmosphären spezifische Kriterien erfüllen, denen bei der Planung von Mikrofonaufnahmetechniken Rechnung zu tragen ist.

4 Immersion

Das ozeanische Gefühl, das die klangliche Umhüllung auslöst, befördert die Regression in jene präsymbolischen Schichten, die sich der Kontrolle des Verstandes entziehen (Flückiger, 2007, S. 59).

In diesem Kapitel wird der Begriff *Immersion* näher betrachtet. Im Kontext audiovisueller Medienproduktionen wird dieser vielfältig genutzt und es fällt schwer, eine eindeutige Definition zu geben. Grundsätzlich wird Immersion dabei als anzustrebendes Ziel betrachtet. Ebenso wird sie als qualitative Eigenschaft verstanden und dementsprechend in der Werbung der Audioindustrie genutzt. Nach einer generellen Betrachtung der Bedeutungsebenen von Immersion werden diejenigen Aspekte näher beleuchtet, die für die auditive Ebene einer Medienproduktion relevant sind und im Kontext mit Atmosphären eine Rolle spielen. Hier wird Immersion teilweise synonym mit Attributen wie Präsenz und Umhüllung genutzt (Berg, 2009).

4.1 Immersion in Medienproduktionen

Neben den Bereichen Akustik und Audiotechnologie wurde Immersion im Kontext vieler weiterer Mediengattungen wie Games, VR, Musik, Film und Literatur untersucht. Murray beschreibt Immersion als psychologisches Konzept:

Immersion is a metaphorical term derived from the physical experience of being submerged in water. We seek the same feeling from a psychologically immersive experience that we do from a plunge in the ocean or swimming pool: the sensation of being surrounded by a completely other reality (J. H. Murray, 1998, S. 98).

Dem gegenüber stehen Definitionen, die Immersion als Attribut einer Technologie oder eines Wiedergabesystems definieren, wie man sie beispielsweise bei Slater und Wilbur (1997) findet. Nach einer Studie der verschiedenen Definitionen von Immersion kommen Agrawal et al. (2019) zum dem Schluss, dass sich die verschiedenen Definitionen, die in der Literatur zu finden sind, den beiden Konzepten "psychischer Zustand" oder "technische Systemeigenschaft" zuordnen lassen. Ferner können verschiedene Gründe für Immersion als psychischen Zustand verantwortlich sein. Abbildung 2 zeigt eine Übersicht der beiden Definitions-Konzepte.



Abbildung 2: Definitionskonzepte von Immersion (Agrawal et al., 2019).

Agrawal et al. (2019) postulieren, dass Immersion nicht im Sinne einer technischen Eigenschaft eines Wiedergabesystems verwendet werden sollte. Sie schlagen folgende Definition vor:

Immersion is a phenomenon experienced by an individual when they are in a state of deep mental involvement in which their cognitive processes (with or without sensory stimulation) cause a shift in their attentional state such that one may experience disassociation from the awareness of the physical world (Agrawal et al., 2019, S. 5).

Eine Besonderheit dieser Definition ist, dass eine Sinnesanregung nicht zwingend erforderlich ist, um Immersion zu erleben (Rumsey, 2020). Zur weiteren Differenzierung von Immersion als technische Eigenschaft eines Wiedergabesystems und als psychischer Zustand einer Person, nutzen Agrawal et. al. die Konzepte von *immersiver Tendenz (Immersive tendency)* und *immersivem Potenzial (Immersive potential)* (Agrawal et al., 2019, S. 5). Nach Witmer und Singer (1998) ist immersive Tendenz als Neigung oder Bereitschaft eines Individuums für das Erfahren eines immersiven Erlebnisses zu verstehen. In der Psychologie ist dies auch als Absorptionsfähigkeit bekannt, die von Persönlichkeitseigenschaften und dem Gefühlszustand einer Person abhängt (Tellegen & Atkinson, 1974). Immersives Potenzial definieren Agrawal et. al. (2019) als Eigenschaft eines Wiedergabesystems oder einer Medienproduktion, Immersion hervorrufen zu können. Aus diesem Konzept lässt sich ableiten, dass es problematisch ist, Immersion im Rahmen von Hörversuchen messen zu wollen, ohne die individuellen Eigenschaften und Gefühlszustände der Testpersonen mit auszuwerten. Audioproduktionen können sich daher nur auf das immersive Potenzial auswirken.

Der Begriff vielseitig im Kontext Soundsystemen immersivwird von und Medienproduktionen verwendet. Nach obigen Ausführungen ist jedoch festzuhalten, dass es ein detailreiches Sounddesign und ein technisch hoch entwickeltes Wiedergabesystem allein nicht vermag, Immersion zu erzeugen. Die verschiedenen gestalterischen und technischen Möglichkeiten der Tonwiedergabe können den Grad der Immersion des Rezipienten steigern, also ein hohes immersives Potenzial besitzen, sind allein jedoch kein hinreichendes Mittel um Immersion zu erzeugen. Vielmehr ist Interesse und Faszination des Rezipienten für das Thema oder den Inhalt der Medienproduktion die grundlegende Voraussetzung für die Empfindung von Immersion (Kock, 2019).

4.2 Auditive Dimensionen von Immersion

Soll mit Hilfe des Tons das immersive Potenzial einer Medienproduktion gesteigert werden, so kann dies nur perzeptiv gelingen. Die Gestaltung des dargebotenen Audiomaterials und das eingesetzte Wiedergabesystem sind dafür entscheidenden Faktoren (siehe Kapitel 3 und 6). Es gilt den Rezipienten akustisch an den Ort des Geschehens zu versetzen. Um ein Gefühl des auditiven vor-Ort-seins hervorzurufen, müssen folglich die räumlichen Hörereignismerkmale betrachtet werden. Theile (2014) unterscheidet dabei zwischen Hörereignisrichtung, Entfernung, räumlicher Tiefe, Raumeindruck und Umhüllung. Der Raumeindruck kann weiter differenziert werden in Räumlichkeit und Halligkeit. Unter Räumlichkeit ist eine Quellenverbreiterung durch frühe Reflexionen zu verstehen, was auch

mit apparent source width (asw) bezeichnet wird. Halligkeit beschreibt eine zeitliche Unschärfe, die durch späte Reflexionen und Nachhall erzeugt wird (Theile et al., 2014).

Tabelle 1 zeigt den Zusammenhang zwischen den räumlichen Hörereignismerkmalen und den Schallfeldkomponenten. In Bezug auf das immersive Potenzial sind die Merkmale Umhüllung (Envelopment), Halligkeit, Räumlichkeit und Entfernung von Bedeutung. Es ist anzumerken, dass sich die Merkmale Halligkeit und Räumlichkeit hauptsächlich auf Schall in Innenräumen beziehen, wobei Reflexionen und Nachhall einen wesentlichen Einfluss haben. Im Fall von Atmosphären sind aber auch Außenaufnahmen von besonderem Interesse. Der Klangcharakter im Freien ist ebenfalls räumlich, beruht aber nicht in erster Linie auf Reflexionen und Nachhall. Der Klangeindruck ist grundsätzlich offen und Schallquellen können sehr weit entfernt sein. Der Entfernungseindruck beruht vornehmlich auf dem Höhenabfall durch die natürliche Dissipation und der zunehmenden Diffusität entfernter Schallquellen. Unzählige Schallquellen aus vielen Richtungen verschmelzen zu einer diffusen Gruppe und erzeugen einen räumlichen Klangeindruck (Rumsey, 2001). Reflexionen wie beispielsweise von Böden, Häuserfronten oder Bäumen beeinflussen zwar das Klangbild, sind aber nicht primär für dessen räumlichen Eindruck verantwortlich. Demnach lässt sich festhalten, dass Umhüllung als das wichtigste räumliche Hörereignismerkmal im Zusammenhang mit Atmosphären angesehen werden kann. Atmosphären, die typischerweise in Innenräumen auftreten, wie Applaus in einem Saal oder reges Treiben in einer Bahnhofshalle, stellen hier keinen Widerspruch dar. Letztlich kann jede Raumreflexion als Spiegelschallquelle betrachtet werden (Blauert, 1974) und bei der Abbildung von Atmosphären geht es weniger um eine feine Abstimmung von Direktschall, ersten Reflexionen und Nachhall einzelner Schallquellen im Raum, sondern um die Abbildung sämtlicher Schallquellen in ihrer Gesamtheit.
Höroroignigmorlemol				
norereignismerkinar	Direktschall	Erste Reflexionen	Nachhall	Umhüllender Schall
Richtung	++	+		
Entfernung, Tiefe		++	+	
Räumlichkeit		++		
Halligkeit			++	
Umhüllung			+	++
Klangfarbe	++	+	++	

Komponenten des Schallfelds

Tabelle 1:Zusammenhang von Hörereignismerkmalen und Schallfeldkomponenten
(Theile et al., 2014).

Der Begriff Umhüllung wird in der Literatur teilweise mit Immersion gleichgesetzt (Berg, 2009). Das wesentliche Unterscheidungsmerkmal besteht jedoch darin, dass Immersion einen kognitiven Zustand beschreibt und Umhüllung ein perzeptives Phänomen (Agrawal et al., 2019). Kock (2019) bezeichnet Umhüllung als akustisch-technische Immersion. Wie Berg (2009) in einer Studie dargelegt hat, existiert auch für das Phänomen der Umhüllung keine eindeutige Definition. Eine Gemeinsamkeit der Definitionen ist jedoch die Idee, von Klang umhüllt zu sein. Berg (2009) schlägt daher den Begriff surroundedness vor. Ferner wird definiert, je nachdem Umhüllung unterschiedlich $^{\rm ob}$ Konzertsaalakustik oder Wiedergabesysteme betrachtet werden. Im Zusammenhang mit Konzertsaalakustik wird häufig der Begriff listener envelopment (LEV) genutzt und bezieht sich auf die frühen Reflexionen und den Nachhall, der in einem Saal erzeugt wird. Listener Envelopment wird daher auch für die Beurteilung und Charakterisierung von Konzertsälen herangezogen (Berg & Nyberg, 2008). Im Kontext von Wiedergabesystemen kann akustische Umhüllung auf mehrere Arten dargestellt werden. Neben der Wiedergabe von Direktschall, ersten Reflexionen und Nachhall, können auch mehrere Direktschallquellen, die um den Rezipienten herum platziert werden, eine Art von Umhüllung erzeugen (Leonard, 2018). Eine einzelne Schallquellen mit sehr großer räumlicher Ausdehnung ist ebenfalls denkbar (Berg,

2009). Im Fall von Außenaufnahmen wird sowohl diffuser Schall, als auch konkrete Schallquellen aus vielen Richtungen wiedergegeben und erzeugen ebenfalls *Umhüllung*.

Eine weitere Differenzierung des Merkmals *Umhüllung* in Abhängigkeit vom Wiedergabesystem findet sich bei Sazdov et al. (2007). Sie konnten zeigen, dass sich die räumliche Abbildung in Bezug auf das Merkmal Umhüllung unterscheidet, je nachdem ob das Wiedergabesystem Höhenlautsprecher beinhaltet oder ausschließlich Lautsprecher auf Ohrhöhe. Nur mit einem Wiedergabesystem, das Höhenlautsprecher enthält, gelingt es, den Hörer mit Schall von oben zu "bedecken", während es mit einem Surround-Audiosystem lediglich möglich ist, den Hörer zu "umringen". Sazdov et al. (2007) stellen diese Eigenschaft eines 3D-Audiosystems als besondere Qualität räumlicher Tonwiedergabe heraus und schlagen den Begriff engulfment vor, um die Umhüllung in der vertikalen Ebene zu bezeichnen, während envelopment die Umhüllung auf der horizontalen Ebene beschreibt.

Um das immersive Potenzial einer Medienproduktion mit Hilfe von Atmosphären zu steigern, müssen diese bei der Wiedergabe über ein 3D-Audiosystem einen Grad an *Umhüllung* erzeugen, der die Anforderungen von *envelopment* und *engulfment* erfüllt. Gelingt dies, so können Atmosphären ihre Funktionen im Rahmen des Sounddesigns erfüllen (siehe Kapitel 3) und ein Gefühl des auditiven *vor-Ort-seins* stellt sich beim Rezipienten ein.

5 Psychoakustische Aspekte der Wahrnehmung von Atmosphären

Spatial hearing and hearing the space are important components in orienting in the sonic environment (Pulkki & Karjalainen, 2015, S. 214).

Diejenigen Aspekte der Psychoakustik, die im Zusammenhang mit der Aufnahme, dem Einsatz in der Postproduktion und der Wiedergabe von Atmosphären von besonderer Bedeutung sind, werden in diesem Kapitel näher beleuchtet. Auch wenn die Verwendung von Atmosphären vornehmlich auf die Erzeugung eines umhüllenden Hörereignisses abzielt, so müssen dennoch die psychoakustischen Aspekte der Lokalisation von Schallquellen für ein ganzheitliches Verständnis der Atmosphären-Perzeption betrachtet werden. Wie bereits in Kapitel 2.2 dargelegt wurde, besteht eine natürliche Atmosphäre nicht ausschließlich aus Diffusschall. Sie enthält in der Regel auch konkrete Schallereignisse mit bestimmter Richtung, welche entsprechend lokalisiert werden können. Diese Schallereignisse können für die Atmosphäre prägend sein, für Abwechslung sorgen und Lebendigkeit verleihen. Ebenso können sich Atmosphären dadurch auszeichnen, dass sich eine sehr große Anzahl konkreter Klänge überlagert und so einen Umhüllungseindruck erzeugt. Die genaue Lokalisation jedes einzelnen Klangereignisses ist in diesem Fall nicht entscheidend, dennoch liegt der Wahrnehmung die Funktionsweise der Lokalisation zu Grunde. Letztlich stellen die präzise Lokalisation einer einzelnen Schallquelle und die Wahrnehmung diffuser Umhüllung zwei Pole des räumlichen Hörens dar. Natürliche Atmosphären bewegen sich zwischen ihnen. Die beiden Ohrsignale und die Auswertung dieser Reize im Gehirn entscheiden über die Wahrnehmung einer akustischen Szene.

5.1 Lokalisation einer Schallquelle

Die Richtungswahrnehmung einer Schallquelle beruht grundsätzlich auf der Auswertung von interauralen Laufzeit- und Pegeldifferenzen, sowie spektralen Merkmalen. In Abhängigkeit von Gestalt und Lage des Schallereignisses, werden verschiedene Charakteristika zur Festlegung des Hörereignis-Ortes herangezogen. Dabei kann zwischen horizontalen und vertikalen Merkmalen unterschieden werden. Die Richtungsauflösung variiert in ihrer Güte. Sie ist beispielsweise frontal zum Kopf am höchsten (Görne, 2015). Zur Beschreibung der Orte der tatsächlichen Schallereignisse sowie der wahrgenommenen Hörereignisse bedient man sich eines kopfbezogenen Koordinatensystems. Durch die Winkelangaben Azimut ϕ und Elevation δ wird die Auslenkung vom Ursprung beschrieben, wobei die frontale Blickrichtung als 0°-Referenz definiert ist (Blauert & Braasch, 2008).

5.1.1 Außenohr-Übertragungsfunktion

Auf ihrem Weg zum Ohr erfahren Schallwellen verschiedene lineare Verzerrungen, weshalb sich die Ohrsignale spektral von denen der Schallquelle unterscheiden. Diese Verzerrungen werden durch die akustisch wirksame Anatomie des Menschen hervorgerufen. In Abhängigkeit der Frequenz entstehen durch Torso, Kopf und Form der Ohrmuscheln unterschiedliche Reflexionen, welche zeitverzögert am Ohr eintreffen und Kammfiltereffekte Torso Verdeckungseffekte erzeugen. Ferner rufen Kopf und bei seitlichen Schalleinfallsrichtungen hervor, wodurch es zur Dämpfung am gegenüberliegenden Ohr kommt (Blauert & Braasch, 2008). Je nach Schalleinfallsrichtung entstehen unterschiedliche lineare Verzerrungen, die zur Richtungsbestimmung ausgewertet werden. Über die Außenohr-Übertragungsfunktion (Head-Related Transfer Function, HRTF) lassen sich die spezifischen Filterungen jeder Richtung für jeweils ein Ohr beschreiben (Hellbrück & Ellermeier, 2004). Im Zeitbereich spricht man von der kopfbezogenen Impulsantwort (Head-Related Impulse Response, HRIR), um sämtliche Effekte des Übertragungssystems zu beschreiben. Bedingt durch anatomische Unterschiede besitzt jede Person eine individuelle HRTF. Ist die Außenohr-Übertragungsfunktion bekannt, können Direktsignale mit den richtungsspezifischen Filtern versehen werden, wodurch sich bei Kopfhörerwiedergabe eine Lokalisation in einer bestimmten Richtung erzielen lässt. Dieses Verfahren wird Binaural-Synthese genannt und in Kapitel 6.7 nochmals aufgegriffen.

5.1.2 Lokalisation in der Horizontalebene

In horizontaler Richtung erfolgt die Lokalisation durch interaurale Laufzeit- und frequenzabhängige Pegeldifferenzen. Befindet sich eine Schallquelle außerhalb der 0°-Achse, erreicht der Schall die beiden Ohren zu unterschiedlichen Zeitpunkten, da die Strecken zum

Ohr unterschiedlich lang sind. Hierdurch entstehen zeitliche Differenzen, was wiederum zu unterschiedlichen Signalen am linken und rechten Ohr führt. Die unterschiedlichen Schallaufzeiten werden interaurale Laufzeitdifferenzen bzw. interaural time delay (ITD) genannt. Der Abstand der Ohren zueinander beträgt circa 17 cm, was zu einer maximalen Laufzeitdifferenz von 0,63 ms führt. Die geringste noch wahrnehmbare Laufzeitdifferenz betrögt 0,03 ms und führt zur Lokalisation einer Schallquelle bei 3° bis 5° außerhalb der Mittelachse (Dickreiter & Goeres-Petri, 2013). Die Auswertung der ITDs erfolgt anhand eines Phasenversatzes der Ohrsignale. Aufgrund des Ohrabstandes kann nur ein begrenzter Frequenzbereich sinnvoll ausgewertet werden. Sehr tiefe Frequenzen erzeugen durch ihre großen Wellenlängen nahezu keinen Phasenversatz (Friesecke, 2014). Hohe Frequenzen verursachen Phasenversätze von mehr als 180°, weshalb hier keine eindeutige Richtungszuordnung mehr gelingt. Befindet sich beispielsweise der Quellort einer reinen Sinusschwingung auf horizontaler Ebene bei $\phi = 90^{\circ}$, so erreicht dieses Signal das abgewandte rechte Ohr mit der maximalen Laufzeitdifferenz von 0,63 ms. Eine Schwingung von 1,59 kHz durchläuft während dieser Zeitspanne eine volle Periode, was eine Phasengleichheit beider Ohrsignale bedingt und daher zu einer möglichen Fehllokalisation, nämlich bei $\phi = 0^{\circ}$ führt. Das Gehör kann bis zu dieser Frequenz Phasendifferenzen beider Ohrsignale erkennen, für höhere Frequenzen bedarf es eines Vergleichs der Hüllkurven beider Ohrsignale (Dickreiter & Goeres-Petri, 2013). Dies ermöglicht eine einfachere Lokalisation breitbandiger, transientenhaltiger Signale wie Sprache oder Musik (Görne, 2015).

Neben den interauralen Laufzeitdifferenzen treten zusätzlich interaurale Pegeldifferenzen auf, welche durch Abschattungseffekte von Kopf und Ohrmuschel entstehen (Howard & Angus, 2009). Durch die Position des Kopfes treten Verdeckungseffekte auf, die frequenzabhängige Pegeldifferenzen hervorrufen (Wenzel et al., 2018). Dies spielt jedoch nur für Frequenzen eine Rolle, deren Wellenlängen klein gegenüber den Abmessungen des Kopfes sind. Dieser Effekt stellt ein weiteres Merkmal dar, welches zur Lokalisation ausgewertet wird und wird als *interaurale Pegeldifferenz* bzw. *interaural level difference (ILD)* bezeichnet. Durch größere Pegeldifferenzen können Schallquellen weiter außerhalb der Mittelachse lokalisiert werden. Dieser Zusammenhang von Pegeldifferenz und Schallereignisrichtung wird jedoch stark von der spektralen Zusammensetzung des Signals beeinflusst. Bei breitbandigen Signalen ergeben sich frequenzabhängige Pegelunterschiede, welche – durch unterbewusste Wahrnehmung - die Lokalisation unterstützen. Die Auswertung dieser Klangfarbenunterschiede basiert auf Erfahrungswerten und Lernprozessen, weshalb laut Dickreiter (2013) eine vereinfachte Lokalisation von uns bekannten Stimmen möglich ist.

ITDs sind vornehmlich für die Lokalisation tiefer Frequenzen verantwortlich, ILDs für die Lokalisation hoher Frequenzen (Howard & Angus, 2009). Diese Laufzeit- und Pegeldifferenzen werden beim natürlichen Hören gleichzeitig ausgewertet, der Übergang zwischen beiden Merkmalen gestaltet sich fließend. Die komplexen Zusammenhänge von Pegel Einfallswinkel spektraler Zusammensetzung, und sind stark von der Signalbeschaffenheit abhängig (Dickreiter & Goeres-Petri, 2013). Somit ist festzuhalten, dass Laufzeit- oder Pegeldifferenzen allein nicht zu einer optimalen Lokalisation auf der horizontalen Ebene führen.

Im dreidimensionalen Raum variiert die Genauigkeit der Lokalisation in starkem Maße und ist besonders für den horizontalen Bereich der Blickrichtung optimiert, wobei die zeitliche und örtliche Auflösung begrenzt ist. Laut Görne (2015) beträgt die Lokalisationsschärfe maximal 1° bei frontaler Schalleinfallsrichtung und nimmt bei seitlicher Schalleinfallsrichtung ab. Durch die begrenzte zeitliche Auflösung wird ein Wechsel der Schalleinfallsrichtung von links/rechts nach 150 ms wahrgenommen, ein Wechsel von vorne/hinten erst nach 250 ms. Ferner führt eine Schallquelle direkt von vorn und hinten zu keinerlei ITDs, weshalb es zu Vorne-Hinten-Vertauschungen kommen kann, insbesondere mit geschlossenen Augen. Kopfbewegungen können in diesem Zusammenhang bei der korrekten Lokalisation helfen. Kleinste Drehungen fungieren als "Peilbewegungen", was zu einem zusätzlichen Informationsgewinn führt (Blauert & Braasch, 2008). Die dadurch entstehenden interauralen Laufzeit- und Pegeldifferenzen ermöglichen die Bestimmung einer Schallquellenrichtung mit hoher Sicherheit (Blauert, 1974) und es gelingt, die Unsicherheit bezüglich Vorne-Hinten aufzulösen (Roginska, 2018).

5.1.3 Lokalisation in der Vertikalebene

Im Gegensatz zur horizontalen Lokalisation beruht diese in vertikaler Richtung auf einem anderen Mechanismus und ist weniger genau. Eine Schallquelle auf der vertikalen Ebene führt nicht zu unterschiedlichen Ohrsignalen, da der Schall beide Ohren zeitgleich mit identischer spektraler Zusammensetzung und gleichem Pegel erreicht. Deshalb kann die Lokalisation auf dieser Ebene nicht durch eine Auswertung der Signalunterschiede erreicht werden. Es treten keine interauralen Merkmale auf (Blauert & Braasch, 2008). Folglich muss die Richtungsauswertung auf monoauralen Merkmalen beruhen, die bereits ein Ohr allein aufweist.

Die bereits erwähnten Kammfilterstrukturen treten auch bei Schallquellen in der vertikalen Ebene auf. Es ergeben sich durch die spezifische Feinstruktur der Ohrmuschel richtungsabhängige Änderungen im Spektrum, welche zur Richtungsauswertung herangezogen werden. Dies tritt jedoch erst bei Frequenzen oberhalb von 5 kHz auf, bei denen die Wellenlänge klein gegenüber den Abmessungen des Ohres ist (Howard & Angus, 2009). Laut Dickreiter (1995) ist die Voraussetzung für eine solche Auswertung, dass die spektrale Zusammensetzung aus Blickrichtung bereits bekannt ist. Unterbewusst wird die Abweichung von diesem Spektrum ausgewertet. Hierzu ist eine vorangegangene Lernphase nötig. Auffallend dabei ist, dass die richtungsbedingten Änderungen im Spektrum nicht bewusst wahrgenommen werden und dieselbe Schallquelle aus unterschiedlichen Richtungen als annähernd gleichklingend wahrgenommen wird (Dickreiter & Goeres-Petri, 2013). Die Lokalisationsschärfe einer elevierten Schallquelle ist geringer als die einer Schallquelle in der horizontalen Ebene. Wie in Abbildung 3 verdeutlicht wird, kommt es zu großen Abweichungen zwischen der Lage von Hör- und Schallereignis (Pulkki & Karjalainen, 2015), da weniger Merkmale zur Auswertung herangezogen werden können. Für einen uns unbekannten Sprecher beträgt die Lokalisationsschärfe circa 15° und verschlechtert sich weiter mit zunehmendem Elevationswinkel (Dickreiter & Goeres-Petri, 2013).



Abbildung 3: Lokalisationsschärfe elevierter Sprachsignale (Pulkki & Karjalainen, 2015).

Blauert (1974) konnte anhand seiner Untersuchungen zeigen, dass in Abhängigkeit vom Einfallswinkel auf vertikaler Ebene spezifische Frequenzbänder angehoben werden. Diese Bänder werden richtungsbestimmende oder "Blauert'sche" Bänder genannt. Demnach führt ein hoher Pegel in den Bereichen 300 bis 500 Hz und 3000 bis 5000 Hz zu frontaler Lokalisation. Pegelüberhöhung um 8 kHz bedingen eine Lokalisation nach oben und die Bereiche 800 bis 1600 Hz sowie 10 bis 16 kHz führen zur Lokalisation im hinteren Bereich. Wallis und Lee (2015) konnten die Existenz der Richtungsbestimmenden Bänder in einer weiteren Studie größtenteils bestätigen.

Für die Lokalisation in der vertikalen Ebene gilt festzuhalten, dass sich breitbandige und bekannte Signale einfacher lokalisieren lassen. Dabei ist zu beachten, dass diese genügend Energie im oberen Frequenzbereich aufweisen müssen (Görne, 2015). ILDs und ITDs können herangezogen Richtungsbestimmung nicht werden, weshalb die spektrale zur Zusammensetzung entscheidend ist. Die Lokalisation in der vertikalen Ebenen ist zudem wesentlich unschärfer als auf horizontaler Ebene. Im Zusammenhang mit Atmosphärenaufnahmen ist diese Unschärfe jedoch weniger kritisch, da die exakte Lokalisation einer erhöhten Schallquelle für die angestrebte Wahrnehmung von Umhüllung und Räumlichkeit nicht notwendig ist. Es genügt also beispielsweise, wenn ein Vogel grob oberhalb des Kopfes lokalisiert wird.

5.1.4 Lokalisation mehrerer Schallquellen

Nachdem die grundsätzlichen Mechanismen der Lokalisation einer Schallquelle betrachtet wurden, gilt es nun den komplexeren Fall von mehreren Schallquellen aus unterschiedlichen Richtungen zu betrachten, denn genau dieser Fall ist für eine Atmosphäre prägend. Je nach Komplexität der akustischen Szene können sehr viele Schallquellen an unterschiedlichen Positionen eine dichte akustische Atmosphäre erzeugen.

Die Wahrnehmung mehrerer Schallquellen beruht auf der Auswertung der Ähnlichkeit der akustischen Reize, welche die Ohren erreichen. So fällt die Unterscheidung zweier unterschiedlicher, gleichzeitig dargebotener Stimmen oder Instrumente nicht schwer und jede Schallquelle wird einzeln wahrgenommen und lokalisiert. Denn die abgestrahlten Signale sind sich unähnlich und weisen somit keine Korrelation zueinander auf (Howard & Angus, 2009). Der Wahrnehmung völlig unabhängiger Signale aus unterschiedlichen Richtungen steht die Wahrnehmung identischer Signale aus verschiedenen Richtungen gegenüber. Diese Signale weisen vollständige Korrelation zueinander auf.

Erreichen korrelierte Signale gleichzeitig die Ohren, kommt es zur Summenlokalisation. Das bedeutet, dass die unterschiedlichen Schallquellen als Einheit wahrgenommen werden und ein Hörereignis erzeugen, welches zwischen den beiden Schallereignissen lokalisiert wird. In der Wahrnehmung entsteht eine neue, nicht real existierende Schallquelle, die als virtuelle Schallquelle bzw. als Phantomschallquelle bezeichnet wird (Görne, 2015). Der Präzedenz-Effekt sorgt dafür, dass zwei korrelierte Signale, welche mit kurzem Zeitversatz (0,6 bis 40 ms) beim Hörer ankommen, ebenfalls nicht als getrennte Schallquellen wahrgenommen werden. Der zuerst eintreffende Schall ist entscheidend für die Lokalisationsrichtung, unabhängig davon, aus welcher Richtung der später eintreffende Schall stammt (Blauert & Braasch, 2008). Der Präzedenzeffekt tritt bis zu einer Verzögerung von etwa 40 ms auf. Eine größere zeitliche Differenz führt zur Echobildung, wodurch zwei getrennte Schallereignisse wahrgenommen werden. Durch den "Haas-Effekt" tritt der Präzedenzeffekt selbst dann auf, wenn das zeitlich nachfolgende Signal einen bis zu 10 dB höheren Pegel aufweist (Dickreiter & Goeres-Petri, 2013). Die beiden Phänomene Summenlokalisation und Präzedenzeffekt bilden die Grundpfeiler der Stereofonie, die bei Aufnahme und Wiedergabe zum Einsatz kommen. Durch die Anpassung von Pegel- und Laufzeitunterschieden der korrelierten

Signale kann die Ausbildung von Phantomschallquellen beeinflusst werden. Man spricht daher auch von Laufzeit- und Intensitätsstereofonie. Werden beide Prinzipien genutzt, spricht man von Äquivalenzstereofonie (Görne, 2015).

Von besonderem Interesse für die Aufnahme und Wahrnehmung natürlicher Atmosphären ist der Bereich zwischen vollständiger Korrelation und Dekorrelation. Die Ausgestaltung der Aufnahmetechnik und Wiedergabesysteme hat starken Einfluss auf das Hörerlebnis und wird in den Kapiteln 6 und 8 näher betrachtet. Grundsätzlich führt eine hohe Korrelation zu einem lokalisierbaren Hörereignis, dessen räumliche Ausdehnung durch den Korrelationsgrad beeinflusst wird. Eine geringe Korrelation führt zu einem diffus umhüllenden Hörereignis. Je nach Kategorie einer Atmosphäre (siehe Kapitel 2.2) variiert das Verhältnis von lokalisierbaren Einzelklängen und nicht lokalisierbarem Diffusschall.

5.2 Entfernungswahrnehmung

Für die Entfernungswahrnehmung einer Schallquelle werden vor allem Lautstärke und Hörereignisses Veränderungen im Klangspektrum eines ausgewertet. In diesem Zusammenhang sind abermals Hörerfahrung und Bekanntheit der Schallquelle nötig. Allgemein gilt, dass die wahrgenommene Lautstärke einer Schallquelle mit steigender Entfernung abnimmt. Bei einer kugelförmigen Abstrahlung nimmt der Schalldruckpegel mit jeder Entfernungsverdopplung um 6 dB ab, wohingegen der Verlust bei gerichteter Abstrahlung um näherungsweise 3 dB geringer ist (Dickreiter & Goeres-Petri, 2013). Da natürliche Schallquellen - wie beispielsweise Musikinstrumente - bestimmte Frequenzbänder kugelförmig und andere gerichtet abstrahlen, führt eine wachsende Entfernung zur Schallquelle neben Pegelverminderungen auch $\mathbf{z}\mathbf{u}$ Klangfarbenänderungen. In Untersuchungen von Fletcher und Munson hat sich herausgestellt, dass das Gehör abhängig vom Schalldruckpegel für verschiedene Frequenzbereiche unterschiedliche Empfindlichkeiten aufweist (Fletcher und Munson, zitiert nach Ellermeier & Hellbrück, 2008). Das Gehör reagiert auf tiefe Frequenzen grundsätzlich weniger empfindlich als auf mittlere Frequenzen und die Empfindlichkeit nimmt mit sinkendem Schallpegel noch weiter ab. Somit wird eine Änderung im Spektrum mit sich veränderndem Abstand zur Schallquelle durch diesen Effekt verstärkt.

Durch die Auswertung des Verhältnisses von Direkt- und Diffusschall wird die Entfernungswahrnehmung innerhalb geschlossener Räume unterstützt (Dickreiter, 1995). Der Direktschallpegel nimmt mit zunehmender Entfernung zur Schallquelle ab, wohingegen der Diffusschallpegel gleichbleibt. Raumvolumen und Nachhallzeit haben somit ebenfalls Einfluss auf das Verhältnis von Direkt- und Diffusschall, weshalb zur Auswertung der Informationen auch hier Hörerfahrung notwendig ist. Zusätzlich zur Entfernung der Schallquelle wird auch die Beschaffenheit des Raumes erfasst. Die Richtungswahrnehmung einer Schallquelle ist grundsätzlich präziser als die Wahrnehmung der Entfernung. Da die Einschätzung der Entfernung schneller und präziser über das Auge erfolgen kann, ist das Gehör weniger gefordert, Hörerfahrungen zur präzisen Entfernungswahrnehmung zu sammeln (Dickreiter & Goeres-Petri, 2013).

5.3 Interaurale Korrelation und Umhüllung

Im Rahmen des räumlichen Hörens ist das Verhältnis beider Ohrsignale zueinander von großer Bedeutung. Insbesondere der Grad der Ähnlichkeit der Signale kann genutzt werden, um Rückschlüsse auf ein zu erwartendes Hörereignis tätigen zu können. Eine mathematische Möglichkeit zum Vergleich zweier Signale bietet die Kreuzkorrelationsfunktion (Acoustical Society of America, 2020):

$$R_{x,y}(\tau) = \lim_{T \to \infty} \frac{1}{2T} \int_{T}^{1} x(t) * y(t+\tau) dt$$
 (1)

Als Maß der Ähnlichkeit von Audiosignalen wird häufig die normierte Kreuzkorrelationsfunktion genutzt (Blauert & Braasch, 2008):

$$R_{norm,x,y}(\tau) = \frac{\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{+T} x(t) * y(t+\tau) dt}{x(t)_{eff} * y(t)_{eff}}$$
(2)

Sie stellt den Kreuzkorrelationskoeffizienten dar, der nur Werte zwischen 1 und -1 annehmen kann. Werden für x(t) und y(t) die beiden Ohrsignale genutzt, ergibt sich der interaurale

Korrelationgrad bzw. der interaurale Korrelationskoeffizient, der im englischen Sprachgebrauch mit *interaural cross-correlation (IAC) bzw. interaural cross-correlation coefficient (IACC)* bezeichnet wird. Dementsprechend kann auch die Ähnlichkeit von Mikrofon- bzw. Lautsprechersignalen berechnet werden. In diesem Fall spricht man von *interchannel cross-correlation (ICCC)*. Im Zusammenhang mit der Betrachtung der Ähnlichkeit eines Signalpaares ist auch der Begriff *Kohärenz* geläufig, welcher in Kapitel 8.2 wieder aufgegriffen wird.

Der Korrelationsgrad der Lautsprechersignale, mit denen eine Person beschallt wird, hat Einfluss auf die wahrgenommene Ausdehnung des Schallfelds und die Lage der Hörereignisse. Je größer die Dekorrelation der wiedergegebenen Signale ist, desto größer ist die wahrgenommene Ausdehnung des Schallfeldes und die Umhüllung. Abbildung 4 veranschaulicht einen Versuch von Chernyak und Dubrowsky (1968). Dabei wurden breitbandige Rauschsignale mit unterschiedlichen Korrelationsgraden über Kopfhörer wiedergegeben und die Versuchspersonen zeichneten die wahrgenommene Ausdehnung und die Lage der Hörereignisse auf. Es zeigte sich, dass mit sinkender Korrelation der Rauschsignale die wahrgenommene Ausdehnung des Schallfeldes steigt, bis es schließlich bei einem Korrelationsgrad von 0 zu zwei getrennten Hörereignissen kommt, die jeweils an einem Ohr wahrgenommen werden (Blauert, 1974). Bei der Interpretation der Ergebnisse dieses Versuches gilt es zu beachten, dass die Wiedergabe der Signale über Kopfhörer erfolgte und somit jedes Ohr getrennt beschallt wurde. Im Falle einer Lautsprecherwiedergabe kommt es zu einem Übersprechen, wodurch jedes wiedergegebene Signal immer beide Ohren erreicht. Die Kopfhörerwiedergabe von zwei vollständig dekorrelierten Signalen stellt demnach einen Extremfall dar. Bei Lautsprecherwiedergabe gänzlich dekorrelierter Signale wird es nicht zu komplett dekorrelierten Ohrsignalen kommen. Hirst et al. (2006) gelang es in einer Versuchsreihe mit mehreren Lautsprecherkonfigurationen nicht, einen IACC unter 0,3 zu erzeugen, was diese Überlegungen stützt.



Abbildung 4: Lage der Hörereignisorte bei Kopfhörerwiedergabe von Rauschsignalen mit unterschiedlichen Korrelationsgraden k (Chernyak & Dubrowsky, 1968 zitiert nach Blauert, 1974).

In Medienproduktionen mit hohem immersiven Potenzial werden Atmosphären über mehrere Lautsprecher wiedergeben. Die Auswirkung der Korrelation der Lautsprechersignale kann daher wichtige Erkenntnisse für die Aufnahmetechnik liefern. Frühe Untersuchungen von Damaske (1967) zu quadrophoner Lautsprecherwiedergabe zeigen, dass es bei der Wiedergabe korrelierter Signale über mehr als zwei Lautsprecher ebenfalls zu einer Summenlokalisation kommt und ein einzelnes Hörereignis entsteht. Somit ist die Beschallung mit mehreren Lautsprechern aus unterschiedlichen Richtungen kein hinreichendes Mittel, um Umhüllung zu erzeugen. Von entscheidender Bedeutung ist die Beschaffenheit der wiederzugebenden Signale. Es zeigt sich, dass auch bei Lautsprecherwiedergabe eine Korrelationsgrades $\mathbf{z}\mathbf{u}$ Vergrößerung Verringerung des einer der empfundenen Schalleinfallsgebiete führt und sich damit positiv auf die Diffusität des Schallfelds und die Umhüllung auswirkt (Cremer & Müller, 1976). Abbildung 5 zeigt die von Versuchspersonen wahrgenommenen Schalleinfallsgebiete bei einem Korrelationsgrad von 0,72, 0,35 und 0,15. Dunklere Bereiche weisen auf eine große Häufigkeit hin mit der die entsprechenden Schalleinfallsgebiete wahrgenommen wurden.



Abbildung 5: Auswirkung des Korrelationsgrades bei Lautsprecherwiedergabe (Links: Korrelationsgrad 0,72. Mitte: Korrelationsgrad 0,35. Rechts: Korrelationsgrad 0,15) (Theile, 2001) schematisiert nach (Damaske, 1967).

Damaske (1967) konnte ebenfalls zeigen, dass sich bereits mit vier Lautsprechern ein subjektiv diffuses Schallfeld erzeugen und sich somit eine gute Umhüllung erzielen lässt. Weiterhin empfiehlt er zur Darstellung diffuser Schallfelder nur Einzellautsprecher zu verwenden, da der Einsatz von Phantomschallquellen zu einer instabilen Wahrnehmung führt und empfindlich auf Bewegungen reagiert. Theile (2000) betont ebenfalls die Stabilität des Schallfeldes und Vergrößerung des Sweetspots hin zu einer Sweetarea bei der Wiedergabe dekorrelierter Signale über ein Mehrkanal-Lautsprechersystem. Um einen möglichst umhüllenden Höreindruck zu erzielen, sind - neben der generellen Anforderung möglichst dekorrelierter Signale - unterschiedliche Frequenzbereiche von besonderer Relevanz. Morimoto und Meakawa (1988) konnten in ihren Untersuchungen zeigen, dass grundsätzlich ein Zusammenhang zwischen dem IACC und der wahrgenommenen Umhüllung besteht. Darüber hinaus wurde die Dekorrelation im Frequenzbereich 100 bis 200 Hz als wichtiger spektraler Einflussfaktor identifiziert. So führte bei der Beschallung mit tiefpassgefiltertem dekorreliertem Rauschen die Anpassung des Tiefpassfilters von 500 Hz auf 200 Hz zu einer gesteigerten Umhüllung, die einer Reduktion des IACC um 0,1 entspricht. Dagegen führte die die Anpassung des Tiefpassfilters von 200 Hz auf 100 Hz zu einer wahrgenommenen Umhüllung, die der Reduktion des IACC um 0,3 entspricht. Weiterhin ist anzumerken, dass der IACC nicht imstande ist diese Frequenzgewichtung widerzuspiegeln. Wird der IACC von breitbandigen Signalen bestimmt, so berücksichtigt der Koeffizient nicht, ob verschiedene Frequenzbereiche mehr oder weniger stark korrelieren. Zwei diffuse Schallfelder unterschiedlichen unteren Grenzfrequenzen führten zu verschieden starken mit Umhüllungseindrücken, obschon der gemessene IACC gleich war.

Weiterhin trägt auch die Wiedergabelautstärke zum Umhüllungseindruck bei. Im Rahmen mehrerer Hörversuche fanden Souldre et al. (2002) heraus, dass durch eine höhere Wiedergabelautstärke die wahrgenommene Umhüllung gesteigert werden kann. Der Wiedergabepegel kann natürlich nur innerhalb eines praktisch sinnvollen Rahmens angepasst werden, doch auch innerhalb einer Mischung ist die Berücksichtigung dieser Erkenntnis sinnvoll. Vorausgesetzt in der Mischung liegen diffuse Atmosphärenaufnahmen vor, kann der Grad der Umhüllung durch Pegelanpassungen der Atmosphäre gesteuert werden.

5.4 Hörszenenanalyse

Unter Hörszenenanalyse versteht man ein Modell zur grundlegenden Hörwahrnehmung. Dabei geht es um die Organisation von Klängen in sinnvollen Gruppen. Die Wahrnehmung beruht auf den grundlegenden Gesetzen der Gestaltpsychologie. So können mehrere Anteile einer Hörszene in der Wahrnehmung zu einer Quelle zusammengefasst werden, aber auch konkurrierende Anteile können voneinander separiert werden (Ellermeier & Hellbrück, 2008). Detaillierte Ausführungen dazu finden sich beispielsweise bei Bregmann (1990). Die grundlegenden Prinzipien sind für die Wahrnehmung von Atmosphären in komplexen Soundscapes aufschlussreich und liefern wichtige Erkenntnisse für Sounddesign und Aufnahmetechnik. Grundsätzlich kann zwischen zwei Gruppierungsproblemen unterschieden werden. Die simultane Gruppierung erfolgt bei der Wahrnehmung gleichzeitig bestehender Klänge. Zur Trennung werden dabei folgende Parameter ausgewertet: Ein- und Aussetzen, Harmonizität, Lokalisationsrichtung, gleiche Amplituden- und Frequenzmodulationen und Klangfarbe. Die sequenzielle Gruppierung erfolgt bei der Wahrnehmung von aufeinanderfolgenden Klängen. Dabei können Schallereignisse miteinander verbunden werden (z.B. zu einer Melodie) oder verschiedenen Quellen zugeordnet werden (Ellermeier & Hellbrück, 2008).

Im Zusammenhang mit Atmosphären ist die simultane Gruppierung von besonderem Interesse, da Atmosphären in Medienproduktionen parallel zu anderen Klängen eingesetzt werden. Wie in Abschnitt 5.3 erläutert wurde, kann ein hoher Wiedergabepegel der Atmosphären dienlich sein, um Umhüllung erzeugen. Geht man bei Atmosphären vergleichbar mit Rausch- oder Störsignalen. Es stellt sich die Frage, inwieweit es möglich ist, Atmosphären vergleichsweise laut wiederzugeben zu können ohne andere wichtige Bestandteile, wie beispielsweise Sprache, in der Mischung zu verdecken. Die genauere Betrachtung der Fähigkeit des Hörsinns, zwischen verschiedenen Quellen zu unterscheiden, ist hier dienlich. Ein populäres Bespiel in diesem Zusammenhang ist der Cocktail-Party-Effekt (Cherry, 1953). Hierunter ist zu verstehen, dass man den Schall einer bestimmten Schallquelle aus einem gemischten Schallfeld heraushören kann, obwohl viele ähnliche Störschallquellen vorhanden sind. Diese Fähigkeit beruht auf dem binauralen Hören, weshalb man auch von binauraler Signalerkennung spricht. Blauert und Braasch (2008) merken an, dass Schallereignisse leichter zu unterscheiden sind, wenn diese räumlich getrennt voneinander auftreten, also in unterschiedlichen Richtungen lokalisiert werden. Für den Einsatz von Atmosphären lässt sich daraus folgern, dass sie möglichst aus "keiner Richtung", also diffus wahrgenommen werden sollen, sodass sich präzise lokalisierbare Klänge räumlich deutlich unterscheiden lassen. Hierzu ist es erforderlich, dass die Kanäle der Atmosphäre eine möglichst geringe Korrelation aufweisen. Blauert (1974) beschreibt in diesem Zusammenhang Untersuchungen von Licklider (1948), bei denen Sprach- und Rauschsignale über Kopfhörer wiedergegeben wurden. Die Ergebnisse sind in Abbildung 6 zusammengefasst. Korrelieren die Rauschsignale beider Kanäle, fallen diese in der räumlichen Wahrnehmung mit dem Sprachsignal zusammen und die Trennung fällt schwer. Sind die Rauschsignale dagegen dekorreliert, kommt es zu einer räumlichen Trennung und die binaurale Signalerkennung gelingt leichter.



Abbildung 6: Hörereignislage von Sprach- und Rauschsignalen bei unterschiedlichen Korrelationsgraden k (Licklider, 1948 zitiert nach Blauert, 1974).

Ein ähnliches Phänomen ist in geschlossenen Räumen zu beobachten: Die binaurale Nachallunterdrückung. Dabei reduziert sich die wahrgenommene Halligkeit beim binauralen Hören im Vergleich zum monoauralen Hören (Blauert & Braasch, 2008). Daher ist es möglich, mit einem hohen Anteil an Reflexionen und Nachhall den Eindruck von Umhüllung zu erzeugen ohne Verdeckungseffekte befürchten zu müssen. Dies gilt besonders beim Einsatz von mehrdimensionalen Wiedergabesystemen, mit denen es möglich ist, verschiedene Reflexionsmuster aus unterschiedlichen Richtungen wiederzugeben.

6 Wiedergabesysteme

But a truly immersive reproduction needs to do more than reproduce localization. It needs to reproduce the spatial quality and the envelopment that we hear in nearly any space (Griesinger, 2018, S. 1).

Die Wiedergabe von Atmosphärenaufnahmen ist über eine Vielzahl von Lautsprechersystemen möglich. Dabei unterscheiden sich die verfügbaren Systeme nicht bloß durch die Anzahl und Position der verwendeten Lautsprecher, sondern ebenso durch die genutzten Formate, die als Zwischenstufe von Aufnahme und Wiedergabe aufgefasst werden können. Die vorliegende Arbeit beschränkt sich bei der Wiedergabe auf zwei definierte Lautsprechersysteme, die sich aufgrund ihres Aufbaus und ihrer Verbreitung besonders für die Wiedergabe und Untersuchung von Atmosphären in immersiven Medienproduktionen eigenen. Diese werden in Abschnitt 6.2 vorgestellt. Für die weiteren Betrachtungen im Rahmen dieser Arbeit, wird dabei vom einfachsten Fall einer direkten Zuordnung von Aufnahme- und Wiedergabekanal ausgegangen. Auf die Untersuchung verschiedener Panning-Verfahren, sowie aktueller objekt- oder szenenbasierter Formate wird dabei bewusst verzichtet. Ihre vielfältigen Signalverarbeitungsstufen bringen zusätzliche Variablen mit sich, die sich ihrerseits auf die akustische Darstellung von Atmosphären auswirken können und deren Analyse den Rahmen dieser Arbeit übersteigen würden.

Im Zusammenhang mit immersiven Medienproduktionen sind die Bezeichnungen "3D" und "3D-Audiosystem" verbreitet. Bei kritischer Betrachtung fällt auf, dass eine korrekte dreidimensionale Darstellung die vollständige Reproduktion von Breite, Höhe und Tiefe erfordert. Der Eindruck von Tiefe wird in Audioproduktionen durch Reflexionen, Hallfahnen, Lautstärke und Klangfarbe erzielt und kann auch bei monophoner Beschallung vermittelt werden. Die Darstellung von Breite und Höhe ist in der Regel auf eine obere Halbkugel beschränkt, wobei die Lautsprechersysteme eine gebogene Fläche mit zwei Dimensionen abdecken. Dennoch wird im Rahmen dieser Arbeit am Begriff "3D-Audio" festgehalten, da dieser zum einen weit verbreitet und prägnant ist und zum anderen eine einfache Unterscheidung zu Surround-Wiedergabesystemen ermöglicht. "3D" steht hier folglich für das Vorhandensein elevierter Lautsprecher und nicht für eine physikalisch korrekte dreidimensionale Darstellung.

6.1 Grundsätzliche Anforderungen

Systeme, die zur Wiedergabe immersiver Medienproduktionen zum Einsatz kommen, erweitern die bekannten Mehrkanalsysteme für Surround um eine weitere Dimension. Durch zusätzliche Lautsprecher, die sich ober- und teilweise auch unterhalb der Ohrhöhe befinden, werden Schalleinfallsrichtungen in der Medianebene ermöglicht. Aufgrund dieser Erweiterung hat sich der Begriff *3D-Audiosystem* etabliert. Die Intention eines solchen Systems lässt sich folgendermaßen zusammenfassen:

"Ein 3D-Audiosystem hat das Ziel Hörereignisse zu erzeugen, die in Richtung, Entfernung, Ausdehnung und ihrer akustischen Umgebung frei manipulierbar sind und für möglichst viele Hörer mit möglichst gleichen Eigenschaften wahrnehmbar sind" (Melchior, 2020).

Im Zusammenhang mit der Wiedergabe von Atmosphären stellt sich die Frage, inwieweit kanalbasierte Wiedergabesysteme in der Lage sind, Umhüllung zu erzeugen und welche Faktoren des Wiedergabesystems hierbei eine besondere Rolle spielen. Umhüllende Atmosphärenaufnahmen zeichnen sich durch einen hohen Anteil diffusen Schalls aus. Ein Wiedergabesystem für immersive Medienproduktionen muss folglich in der Lage sein, ein subjektiv diffuses Schallfeld zu erzeugen. Subjektiv diffus bedeutet in diesem Zusammenhang, dass Rezipienten den Eindruck haben, der Schall komme aus allen Richtungen, auch wenn objektiv nur diskrete Schalleinfallsrichtungen in Form von Lautsprechern an festen Positionen vorhanden sind (Damaske, 1967). Theile und Wittek (2011b) merken dabei an, dass diffuser Schall auch diffus wiedergegeben werden muss. Aktuelle Wiedergabesysteme sind dazu in der Lage, sofern die wiederzugebenden Lautsprechersignale ausreichend unterschiedlich sind, was bedeutet, dass sie über einen weiten Frequenzbereich dekorrelieren. Dies bestätigen auch Lauterslager und Nuyten:

"In a multi-channel loudspeaker system, a diffuse soundfield can be achieved by feeding a number of loudspeakers with incoherent signals. Both the degree of coherence between the signals feeding the loudspeakers and the loudspeaker layout, influence the diffuseness of the soundfield. [...] If a high correlation degree is measured, it can be expected that the signal is heard from a distinct direction. At a low correlation degree, the direction of the signal will be difficult to establish" (Lauterslager & Nuyten, 1976, S. 3).

6.2 Etablierte Lautsprechersetups

Audioproduktionen Im Rahmen immersiver sind derzeit verschiedenste Lautsprecherkonfigurationen verbreitet. Viele Hersteller entwickeln ihre eigenen Setups, außerdem vergrößert sich die Bandbreite möglicher Wiedergabesysteme laufend. Ein einheitlicher Standard konnte sich bisher nicht etablieren, wodurch die Auswahl und Planung geeigneter Mikrofonaufnahmeverfahren erschwert wird. Idealerweise können Atmosphärenaufnahmen auf vielen Lautsprecherkonfigurationen adäquat wiedergegeben werden und bieten die Sicherheit, auch auf zukünftigen Systemen eingesetzt werden zu können. Einen ersten Schritt in Richtung Standardisierung geht die ITU mit ihrer Empfehlung ITU-R BS.2051-2 (ITU, 2018). Doch auch darin wird eine Vielzahl möglicher Konfigurationen vorgeschlagen, weshalb in dieser Arbeit nur die Lautsprecherkonfigurationen der Systeme D und J näher betrachtet werden. Dabei bietet sich die von der ITU vorgeschlagene Schreibeweise in der Form (U+M+B) an, um die verschiedenen Systeme zu bezeichnen. M steht dabei für die Anzahl der Lautsprecher im Middle-Layer, also auf Ohrhöhe, U für die Lautsprecher im Upper-Layer, also mit positivem Elevationswinkel und B für Lautsprecher im Bottom-Layer, also mit negativem Elevationswinkel. Abbildung 7 zeigt das Wiedergabesystem D (4+5+0), die genauen Lautsprecherpositionen sind in Tabelle 2 aufgeführt. Sein Aufbau ist vollständig kompatibel zum bekannten Surroundsystem 5.1, wie es in der ITU-R BS.775-3 definiert ist (ITU, 2012). Dieser Umstand ermöglicht es, bestehende Surroundsysteme um vier Höhenkanäle zu erweitern, um ein Wiedergabesystem für immersive Medienproduktionen aufzubauen. Hierdurch kann der zeitliche und finanzielle Aufwand einer Umrüstung geringgehalten werden und es besteht vollständige Abwärtskompatibilität. Darüber hinaus ist diese Lautsprecherkonfiguration kompatibel zu den kommerziellen Formaten Auro 3D und Dolby Atmos (Auro Technologies, 2015; Dolby Laboratories, 2018). Aus diesen Gründen ist das

Wiedergabesystem D weit verbreitet und stellt somit ein geeignetes System für die Untersuchung von Atmosphären in immersiven Audioproduktionen dar.



Abbildung 7: Wiedergabesystem D (4+5+0) nach ITU-R BS.2051-2. Links: Upper-Layer. Rechts: Middle-Layer (ITU, 2018).

Kürzel	Kanal	Lautsprecherposition		
	Kallai	Azimut-Winkel [°]	Elevations-Winkel [°]	
L	Left	+30	0	
\mathbf{R}	Right	-30	0	
С	Center	0	0	
\mathbf{LFE}	LFE	_	_	
Ls	Left Surround	+100 +120	0	
Rs	Right Surround	-100120	0	
Ltf	Left top front	+30 $+45$	+30 +55	
Rtf	Right top front	-3045	+30 +55	
Ltr	Left top rear	+110 +135	+30 $+55$	
Rtr	Right top rear	-110135	+30 +55	

 $Tabelle \ 2: \qquad Lauts precherkon figuration \ D \ (4+5+0) \ nach \ ITU-R \ BS. 2051-2 \ (ITU, \ 2018).$

Ebenfalls soll das Wiedergabesystem J (4+7+0) näher betrachtet werden. Es erweitert das Wiedergabesystem D (4+5+0) um zwei seitliche Surroundlautsprecher im Middle-Layer und die hinteren Surroundlautsprecher rücken im Vergleich zu Wiedergabesystem D (4+5+0) weiter nach hinten. Den prinzipiellen Aufbau zeigt Abbildung 8, die Lautsprecherpositionen sind in Tabelle 3 aufgelistet. Besonders die seitlichen Surroundlautsprecher im Middle-Layer machen dieses Wiedergabesystem interessant für die Wiedergabe von Atmosphären. Denn die Lautsprecherpositionen von +/- 90° haben einen hohen Anteil an der umhüllenden Wirkung, wenn diffuser Schall wiedergeben wird (Berg & Nyberg, 2008; Furuya et al., 2005; Soulodre et al., 2003) und wie in Kapitel 5.1.2 vorgestellt wurde, erzeugen Schalleinfallsrichtungen von Azimut 90° die größten ILDs.



Abbildung 8: Wiedergabesystem J (4+7+0) nach ITU-R BS.2051-2. Links: Upper-Layer. Rechts: Middle-Layer (ITU, 2018).

Kürzel	Kanal	Lautsprecherposition		
		Azimut-Winkel [°]	Elevations-Winkel [°]	
L	Left	+30	0	
R	Right	-30	0	
\mathbf{C}	Center	0	0	
LFE	LFE	_	_	
Lss	Left side surround	+85 +110	0	
Rss	Right side surround	-85110	0	
Ls	Left Surround	+100 +120	0	
Rs	Right Surround	-100120	0	
Ltf	Left top front	+30 $+45$	$+30 \ldots +55$	
Rtf	Right top front	-3045	$+30 \ldots +55$	
Ltb	Left top back	+110 +135	$+30 \ldots +55$	
Rtb	Right top back	-110135	+30 +55	

Tabelle 3: Lautsprecherkonfiguration J (4+7+0) nach ITU-R BS.2051-2 (ITU, 2018).

6.3 Einsatz der Surroundkanäle

Um genauere Erkenntnisse über die Wirkungsweise eines mehrkanaligen Wiedergabesystems zu gewinnen, ist es sinnvoll, zunächst Wiedergabesysteme ohne Höhenlautsprecher zu betrachten. Bereits Damaske (1967) konnte zeigen, dass sich mit vier gleichmäßig verteilten Lautsprechern auf Ohrhöhe ein relativ diffuses Schallfeld erzeugen lässt und ein Umhüllungseindruck entsteht. Lauterslager und Nuyten (1976) nutzten dagegen eine fünfkanalige Lautsprecherkonfiguration mit den Positionen +/- 30°, +/- 90° und 180° und konnten so ebenfalls Umhüllung erzeugen. Morimoto (1997) fand zudem heraus, dass insbesondere die hinteren Lautsprecher von Bedeutung sind, wenn der Eindruck akustischer Umhüllung erzeugt werden soll. Diese und weitere Forschungsarbeiten führten letztlich zur bekannten 5.1 Lautsprecherkonfiguration, wie sie in der *ITU-R BS.775-3* (ITU, 2012) standardisiert wurde. Im Rahmen einer systematischen Untersuchung nutzten Hiyama et al. (2002) dekorrelierte Rauschsignale, um die Mindestanzahl und Positionierung von Lautsprechern auf Ohrhöhe zur Erzeugung eines diffusen Schallfeldes zu ermitteln. Dabei kamen sie zu dem Ergebnis, dass bei konzentrischer Verteilung der Lautsprecher in gleichen Abständen mindestens sechs Lautsprecher erforderlich sind. Werden dagegen symmetrische Lautsprecherpaare genutzt, welche unterschiedliche Abstände zueinander haben, sind bereits Lautsprecher ausreichend. Überdies vier konnten sie zeigen, dass die Lautsprecherkonfiguration nach ITU-R BS.775-3, mit Ausnahme des Center-Kanals, eine optimale Positionierung der Lautsprecher zu Erzeugung diffuser Schallfelder aufweist. Weiterhin wurde festgestellt, dass auch mit Mikrofonsystemen Signale erzeugt werden können, mit denen die positiven Eigenschaften dieser Lautsprecherkonfiguration zur Diffusschallwiedergabe, ausgenutzt werden (Hamasaki & Hiyama, 2003).

Im Fall von Heim-Wiedergabesystemen kann angemerkt werden, dass diese in der Regel vorteilhaftere Voraussetzungen für die Wiedergabe diffuser Atmosphären aufweisen, als für die Wiedergabe präzise lokalisierbarer Schallquellen zwischen den Surroundkanälen. Dies liegt zum einen in einer suboptimalen Lautsprecheraufstellung – die nicht den Empfehlungen der ITU folgt – begründet, zum anderen in den akustischen Bedingungen des Wiedergaberaumes. Im Gegensatz zu Produktionsstudios und Kinosälen handelt es sich um kleinere Räume mit parallelen, schallharten Wänden, wodurch es zu zahlreichen Reflexionen kommt, die für den Eindruck eines subjektiv diffusen Schallfeldes nützlich sind. Um die Diffusschallwiedergabe weiter zu unterstützen, können für die Surroundkanäle Dipol-Lautsprecher eingesetzt werden. Dabei wird der Schall in zwei Richtungen abgestrahlt und erreicht den Hörer nur über Reflexionen. Das Prinzip ist in Abbildung 9 dargestellt. Diese Art der Wiedergabe fördert den Umhüllungseindruck, beeinflusst jedoch auch die Lokalisation negativ, da keine direkt Schallabstrahlung zum Hörer erfolgt (Holman, 2008).



Abbildung 9: Dipol-Lautsprecher im 5.1 Wiedergabesystem (Holman, 2008).

Einen weiteren Spezialfall stellt der Einsatz der Surroundlautsprecher in Kinowiedergabesystemen dar. Die Herausforderung dieser Anwendung liegt darin, dass eine große Publikumsfläche möglichst gleichmäßig zu beschallen ist. Um dies zu erreichen, kommen für die Surround-Kanäle Lautsprecher-Gruppen zum Einsatz, die parallelgeschaltet und über eine große Fläche verteilt werden (Peter, 2020). Eine solche Anordnung zeigt Abbildung 10. Für die Erzeugung eines umhüllenden Klangeindrucks ist die Verwendung von Lautsprechergruppen nachteilig, da so vollständig korrelierte Signale über viele Lautsprecher wiedergegeben werden. Überdies sind die Schalllaufzeiten der einzelnen Lautsprecher zu jedem Platz unterschiedlich, sodass sich Kammfilter und damit Klangverfärbungen ergeben (Holman, 2008). Lautsprechergruppen für Surroundkanäle kommen im Kino nicht nur beim 5.1-Format zum Einsatz, sondern werden auch für das Surround-Format 7.1 und das immersive Format Dolby Atmos genutzt. Im Fall von Dolby Atmos werden für Atmosphären in der Regel kanalbasierte "Beds" eingesetzt, die über Lautsprechergruppen wiedergegeben werden. Diese "Beds" entsprechen denen der 5.1 und 7.1 Systeme. Werden Signale in die Beds gemischt, verhalten sich diese wie bei einer kanalbasierten Mischung. Für die Höhenkanäle werden zwei Lautsprechergruppen genutzt, die jeweils in einer Linie unter Decke der angebracht sind (Watson, 2013). Die Problematik Mehrfachwiedergabe vollständig korrelierter Signale der und Erzeugung von Kammfilterstrukturen ist somit auch hier gegeben. Dolby Atmos bietet durch seinen objektorientierten Ansatz jedoch die Möglichkeit, jeden Lautsprecher einzeln anzusteuern.

Die Signale einer möglichst einhüllenden Atmosphärenaufnahme könnten folglich als "Objekte" definiert und somit über einzelne Lautsprecher anstatt über Lautsprecher-Gruppen wiedergeben werden. Dadurch wird die Mehrfachwiedergabe gänzlich korrelierter Signale vermieden, was sich positiv auf die Umhüllung und Klangfarbe auswirkt. Erste Versuche von Andriessen und Hoffmeister (2017) unterstützen diesen Ansatz, es sind jedoch weitere systematische Untersuchungen nötig, bevor diese Theorie als allgemeingültige Praxisempfehlung aufgefasst werden kann.



Abbildung 10: Beispielhafte Anordnung eines 5.1 Wiedergabesystems im Kino (Peter, 2020).

6.4 Einsatz der Höhenkanäle

Nachdem gezeigt werden konnte, dass sich bereits mit Surroundsystemen diffuse Schallfelder und Umhüllung erzeugen lassen, stellt sich die Frage, inwieweit immersive Wiedergabesysteme mit Höhenkanälen das Hörerlebnis weiter steigern und zu einer höherwertigen Wiedergabe von Räumlichkeit führen können. Erste Untersuchungen zum Einfluss der Lautsprecherpositionen auf die perzeptive Umhüllung, die auch Lautsprecher in der Höhen-Ebene beinhalteten, konnten einen signifikanten Zuwachs der Umhüllung durch elevierte Lautsprecherpositionen nachweisen. Hierbei kamen Lautsprecher an den Positionen 0° azimut, 90° elevation oder +/- 90° azimut, 80° elevation zum Einsatz. Zudem konnte gezeigt werden, dass Rezipienten deutlich zwischen diffusen Schallfeldern, die mit und ohne Höhenlautsprecher erzeugt wurden, unterscheiden können (Furuya et al., 2001; Wakuda et al., 2003). Weitere Erkenntnisse zum Einfluss der Höhenkanäle auf die Umhüllung lieferten Untersuchungen von Power et al. (2014). Sie verglichen verschiedene, mehrkanalige Wiedergabesysteme mit und ohne Höhenkanälen in Hinblick auf das Attribut Umhüllung miteinander, wobei vier Höhenlautsprecher mit Elevationswinkel von 35° eingesetzt wurden. Es zeigte sich, dass der Einsatz von Höhenlautsprechern die Umhüllungswahrnehmung deutlich steigern kann. Dieses Ergebnis deckt sich ebenfalls mit Aussagen von Tsingos et al. (2010). In einer größeren Vergleichsstudie der Universität Southhampton wurden verschiedene 3D-Wiedergabesysteme mit vielfältigen Lautsprecherkonfigurationen untersucht. Zum einen bestätigt diese Studie die Ergebnisse von Power et al., dass sich Höhenlautsprecher positiv auf die Umhüllung auswirken, zum anderen zeigt sie, dass auch die Anzahl der Lautsprecher ein signifikanter Faktor ist und mehr Lautsprecher zu einer gesteigerten Wahrnehmung von Umhüllung führen. Dies gilt sowohl für Lautsprecher auf Ohrhöhe als für solche in der Höhenebene. Es gilt jedoch zu beachten, dass neben der bloßen Anzahl eine möglichst gleichmäßige Verteilung der Lautsprecher vorteilhaft ist (Cousins et al., 2015).

In *ITU-R BS.2051-2* wird für die Lautsprecher der Höhenebene ein großer Bereich von 30 bis 50° Elevation angegeben und es stellt sich die Frage, welche Elevationswinkel zu bevorzugen sind. Theile und Wittek (2011b) merken an, dass eine Höhenlokalisation nur am oberen Rand des Upper-Layer möglich ist. Dies stützen auch Untersuchungsergebnissen von Lee (2011), nach welchen vertikale Phantomschallquellen zwischen den Lautsprechern der Mittel- und Höhenebene weniger präzise gebildet werden können als solche auf horizontaler Ebene. Diese Erkenntnisse sprechen dafür, für die Lautsprecher des Upper-Layer einem merklich höheren Winkel als 30° zu nutzen. In Kapitel 0 wurde aufgezeigt, dass die Wahrnehmung elevierter Schallquellen von einer hohen Lokalisationsunschärfe geprägt ist. Höhere Elevationswinkel der Lautsprecher können dafür sorgen, dass trotz einer hohen Abweichung, die Wahrnehmung einer elevierten Schallquelle erhalten bleibt. Auch wenn eine stabile Lokalisation in der Höhe für die Wiedergabe von diffusen Atmosphären weniger relevant ist, kann der höhere Elevationswinkel dennoch dienlich sein. Der obere Halbraum innerhalb des Wiedergabesystems wird durch einen Elevationswinkel von 45° der Höhenlautsprecher gleichmäßiger abgedeckt als durch einen Elevationswinkel von 30°. Cousins et al. (2015) konnten zeigen, dass eine möglichst gleichmäßige Verteilung der Lautsprecher entscheidend ist, zur Erzeugung eines diffusen, umhüllenden Schallfelds. Es muss jedoch bedacht werden, dass dies nur innerhalb des Sweetspots zutrifft – außerhalb des Sweetspots wurden mit niedrigen Elevationswinkeln bessere Ergebnisse erzielt.

Ferner gibt \mathbf{es} Anwendungsfälle von Höhenlautsprechern, bei denen die Abstrahlcharakteristika der verwendeten Lautsprecher eine besondere Rolle spielen. Sowohl bei Heim- als auch in Kinowiedergabesystemen werden die Höhenlautsprecher teilweise direkt unter der Decke montiert. Diese Montageart wird beispielsweise für Dolby Atmos Heim-Systeme vorgeschlagen (Dolby Laboratories, 2018) und findet sich auch an einigen Hörpositionen von Dolby Atmos Kino-Systemen wieder (Peter, 2020). Es gilt zu berücksichtigen, dass die akustische Achse des Lautsprechers dabei senkrecht nach unten zeigt. Typischerweise besitzen Mehrwegelautsprecher außerhalb dieser Achse keinen ausgewogenen Frequenzgang und es kommt hier insbesondere in den Übergangsbereichen der Treiber zu Sprüngen im Abstrahlverhalten. Abbildung 11 zeigt das Abstrahlverhalten eines 3-Wege-Lautsprechers und verdeutlicht die Unterschiede zwischen horizontaler- und vertikaler Abstrahlung. Lund (2020) gibt in diesem Zusammenhang zu bedenken, dass sich im Falle eines Wiedergabesystems mit vier elevierten Kanälen und direkter Deckenmontage, die akustische Achse nicht auf den Sweetspot ausgerichtet ist. Ein Rezipient im Sweetspot weicht sowohl horizontal als auch vertikal von der akustischen Hauptachse ab, weshalb mit deutlichen Klangverfärbungen zu rechnen ist. Im Sinne einer möglichst homogenen Schallabstrahlung sollte folglich die akustische Achse der Lautsprecher stets auf den Sweetspot ausgerichtet werden. Ist dies nicht möglich und eine senkrechte Deckenmontage erforderlich, sollten Lautsprecher mit einer möglichst breiten, frequenzunabhängigen Abstrahlcharakteristik zum Einsatz kommen.



Abbildung 11: Schalldruck-Isobaren der frequenzabhängigen Abstrahlwinkel eines 3-Wege Lautsprechers. Oben: Abstrahlcharakteristik horizontal. Unten: Abstrahlcharakteristik vertikal. Modell: Neumann KH 310 (NEUMANN, o. J.).

6.5 Einsatz des Centerkanals

Es stellt sich die Frage, inwieweit es sinnvoll ist, den Centerkanal bei der Aufnahme und Wiedergabe von Atmosphären zu nutzen. Das Ziel der Atmosphären-Wiedergabe besteht allgemein darin, ein subjektiv möglichst diffuses Schallfeld zu erzeugen. Ein zusätzlicher Lautsprecher aus einer anderen Richtung ist diesem Ziel grundsätzlich dienlich. Jedoch muss berücksichtigt werden, dass hierzu ein weiteres, dekorreliertes Lautsprechersignal erforderlich ist. Dieser Umstand lässt folglich die Mikrofonaufnahmetechnik komplizierter werden, da jeder zusätzlich zu erfassende, dekorrelierte Kanal das Mikrofonsetup größer und impraktikabler werden lässt (Theile & Wittek, 2011a), was in Kapital 7 näher betrachtet wird.

Darüber hinaus wurde in Kapitel 6.3 aufgerührt, dass die Position der Lautsprecher von entscheidender Bedeutung ist. Es ist eine möglichst homogene Beschallung anzustreben, um so ein isotropes Schallfeld im Sweetspot anzunähern. Durch die Lautsprecherpositionen von +/- 30° der Kanäle L und R der Wiedergabesysteme, welche im Rahmen dieser Arbeit betrachtet werden, ist der frontale Bereich ohnehin schon mehr als ausreichend abgedeckt und ein zusätzlicher Centerlautsprecher brächte kaum einen Vorteil. Vielmehr käme es durch das Ungleichgewicht von drei frontalen und zwei rückwärtigen Lautsprechern zur Ausbildung einer frontalen Vorzugsrichtung, die der Umhüllung nicht dienlich ist (Yewdall, 2012). Denn es sind besonders die seitlichen und rückwärtigen Schalleinfallsrichtungen, die sich positiv auf die Umhüllung auswirken. Außerdem sollte die räumliche Verteilung der Lautsprecher möglichst gleichmäßig sein (Cousins et al., 2015; Morimoto, 1997; Soulodre et al., 2003).

Ferner erfüllt der Centerkanal in mehrkanaligen Wiedergabesystemen – besonders wenn diese bildbezogen sind – eine wichtige Aufgabe: In der Regel beinhaltet er relevante Informationen einer Medienproduktion, wie beispielsweise Kommentar, Gesang oder O-Ton und garantiert eine klare Lokalisation, auch für Rezipienten außerhalb des Sweetspot (Peter, 2020). Die Verständlichkeit der Informationen ist entscheidend, weshalb der Centerkanal meist exklusiv für diese genutzt wird. Die Wiedergabe von Atmosphären über den Centerkanal würde die Verständlichkeit dagegen herabsetzen. Etablierte Mikrofonsetups zur Aufnahme von Surround-Atmosphären, wie beispielweise das IRT-Kreuz oder ORTF-Surround, verzichten daher auf ein Mikrofon für den Centerkanal (Wittek, 2012).

Aus den aufgeführten Gründen kann geschlussfolgert werden, dass der Einsatz des Centerkanals für Atmosphären als Sounddesignkomponente in immersiven Medienproduktionen nicht empfehlenswert ist. Daher verzichten die in Kapitel 8 praktisch untersuchten Mikrofonsetups auf den Einsatz eines Mikrofons für den Centerkanal.

6.6 Tieftonwiedergabe

Die in Kapitel 6.2 vorgestellten Wiedergabesysteme beinhalten einen optionalen LFE (Low Frequency Enhancement) Kanal. Dieser Kanal beinhaltet tieffrequente Soundeffekte und wird separat über einen Subwoofer wiedergegeben. Er ist ausdrücklich nicht als tieffrequente Erweiterung der anderen Kanäle zu verstehen und deren Tieftonanteil sollte nicht in den LFE-Kanal gemischt werden (Fisher, 2005). Morimoto und Maekawa (1988) konnten zeigen, dass tieffrequente, dekorrelierte Signale den Umhüllungseindruck besonders fördern. Würde man die tieffrequenten Anteile einer dekorrelierten Atmosphärenaufnahme in den LFE-Kanal mischen, käme es zu einer Monobildung, mit der die tieffrequente Dekorrelation verloren ginge. Darüber hinaus kann nicht ausgeschlossen werden, dass es durch die Summation von gegenphasigen Signalanteilen zu Interferenzeffekten kommt, welche zu Verzerrungen des Spektrums führen würden. In der Postproduktion ist also stets von Vollbereichslautsprechern auszugehen, durch die sich eine Summation der Tieftonanteile erübrigt. Das Problem der monophonen Summation ergibt sich jedoch zwangsweise beim Einsatz eines Subwoofers in Kombination mit Bass-Management. Dabei dient der Subwoofer nicht nur der Wiedergabe der Signale des LFE-Kanals, sondern auch der tieffrequenten Erweiterung der übrigen Kanäle. Besonders bei kleineren Heimwiedergabesystemen kommen oftmals Lautsprecher mit begrenztem Frequenzbereich zum Einsatz, welche die Verwendung von Bass-Management notwendig machen (Holman, 2008). Die Verwendung eines einzelnen Subwoofers für die Anwendung von Bass-Management beruht auf den Überlegungen, dass tieffrequente Anteile nicht lokalisiert werden können und somit keinen Beitrag zum Räumlichkeitseindruck und zur Umhüllung leisten. Dies wurde beispielsweise von Borenius (1985) und Kügler und Theile (1992) untersucht, wobei eine Übergangfrequenz zum Subwoofer von 100 bis 140 Hz empfohlen wurde. Dies steht jedoch im Widerspruch zu den Ergebnissen jüngerer Arbeiten, die verdeutlichen, dass tieffrequenter Schall, sofern er dekorreliert ist, aus unterschiedlichen Richtungen deutlich wahrgenommen werden kann und einen relevanten Beitrag zur Wahrnehmung von Räumlichkeit und Umhüllung leistet. Es wird empfohlen mindestens zwei Subwoofer aus unterschiedlichen Richtungen einzusetzen oder bei Verwendung eines einzelnen Subwoofers, die obere Grenzfrequenz auf maximal 63 Hz festzulegen (Griesinger, 2018; Martens, 1999; Martens et al., 2004). Somit lässt sich festhalten, dass für einen maximalen Umhüllungseindruck idealerweise Vollbereichslautsprecher eingesetzt werden und auf Bass-Management verzichtet wird. Ist

dies in kleineren Heimwiedergabesystemen nicht zu realisieren, kann durch den Einsatz mehrerer Subwoofer dennoch ein Umhüllungseindruck mit tiefen Frequenzanteilen realisiert werden.

6.7 Entwicklung einer virtuellen Abhörumgebung in MAX

Mit Hilfe einer virtuellen Abhörumgebung werden im Rahmen dieser Arbeit die Eigenschaften verschiedener Mikrofonaufnahmesetups miteinander verglichen. Eine Dokumentation der aufgenommenen Atmosphären und eingesetzten Mikrofonarrays findet sich in Kapitel 8.1. Die Entwicklung der virtuellen Abhörumgebung erfolgte in MAX, einer grafischen Programmierumgebung der Firma Cycling '74, die hauptsächlich in Musik- und Multimediaprojekten Anwendung findet. Ziel ist es, die Wiedergabe von verschiedenen Atmosphärenaufnahmen über die Wiedergabesysteme D (4+5+0) und J (4+7+0) zu simulieren und dabei die Vorgaben der ITU-R BS.2051-2 zu erfüllen. Jeder Lautsprecherkanal wird durch eine virtuelle Punktschallquelle repräsentiert. Auf die Simulation der Raumakustik eines realen Wiedergaberaumes wird dabei bewusst verzichtet. Hierdurch kann das Wiedergabeszenario, wie es in einem reflexionsarmen Raum stattfände, simuliert werden. Der Einfluss der Lautsprecherpositionen kann somit unabhängig von akustischen Eigenschaften eines Wiedergaberaumes untersucht werden. Grundsätzlich ist davon auszugehen, dass sich frühe Reflexionen und Nachhall eines realen Wiedergaberaumes positiv auf die subjektive Diffusität auswirken. Besonders durch die frühen Reflexionen werden weitere dekorrelierte Signale aus unterschiedlichen Richtungen erzeugt. Die Simulation eines reflexionsarmen Raumes stellt somit einen Extremfall dar, der sich jedoch für eine Untersuchung besser eignet und den Faktor Raumakustik gänzlich ausschließt.

Wie zu Beginn des Kapitels erläutert, erfolgt die Wiedergabe der Mikrofonsignale dabei ausschließlich über eine direkte Kanalzuordnung. Im Sweetspot des simulierten Wiedergabesystems befindet sich ein virtueller, fixierter Kunstkopf. Dieser generiert Ohrsignale, die stellvertretend für diejenigen einer Person im Sweetspot des Wiedergabesystems stehen. Im Zuge einer Korrelationsanalyse können diese Ohrsignale ausgewertet werden und liefern damit Erkenntnisse über den Einfluss verschiedener Mikrofonarrays (siehe Kapitel 8.3). Zur Umsetzung dieser Anwendung werden die Wiedergabesignale mit den HRIRs gefaltet, die der Position eines Lautsprechers des Wiedergabesystems entsprechen. Die Binauralsynthese erfolgt dabei mit Programmstrukturen aus der Softwarebibliothek SPAT, einer Entwicklung des französischen Forschungsinstitutes für Akustik und Musik IRCAM. SPAT erweitert MAX um Funktionen und Signalprocessing-Algorithmen zur räumlichen Tonbearbeitung (Carpentier, 2018; Ircam, o. J.). Weitere Informationen zum Funktionsumfang von SPAT finden sich beispielsweise bei Carpentier et al. und in den umfangreichen Handbüchern (2015; Ircam, 2018, 2020). Abbildung 12 zeigt die Oberfläche des MAX-Patches in der Präsentationsansicht. Die Audiodateien können wahlweise in mono einzelnen Wiedergabekanälen zugeordnet oder als Mehrkanaldatei eingeladen werden. Dabei gilt die Kanalreihenfolge nach ITU-R BS.2051-2. Loop-Wiedergabe, Gain-Anpassung, sowie die Aufzeichnung des Ausgangssignal als kopfbezogene Stereodatei sind ebenfalls möglich. Zur Überprüfung der eingeladenen Dateien, können die wichtigsten Informationen wie Dateiname, Länge, Kanalanzahl, Samplerate und Bit-Tiefe angezeigt werden.



Abbildung 12: Virtuelle Abhörumgebung in MAX (Präsentationsansicht).

Im Bereich HRTF-Setup können die Datensätze ausgetauscht werden, die bei der Binauralsynthese zum Einsatz kommen. Der Max-Patch kann dazu HRTF-Datensätze im SOFA-Format verarbeiten. Beim SOFA-Format handelt es sich um einen offen Standard, welcher in AES69-2015 definiert ist (Audio Engineering Society, 2015). Vielfältige HRTF-Datensätze stehen hierdurch für Forschung und Entwicklung frei zur Verfügung. Für die weiteren Untersuchungen im Rahmen dieser Arbeit wird ein Datensatz genutzt, welcher mit dem Kunstkopfmodell KU-100 der Firma Neumann im reflexionsarmen Raum des Instituts für Nachrichtentechnik an der TH Köln erstellt wurde. Als Lautsprecher kam ein Genelec 8260A im Abstand von 3,2 Metern zum Einsatz. Dabei wurde der gesamte sphärische Bereich mit einer Auflösung von 2° erfasst. Es handelt sich folglich um einen Fernfeld HRTF-Datensatz mit einer hohen räumlichen Auflösung (Bernschütz, 2013). Die so erfassten Positionen können über die jeweiligen HRTFs exakt beschrieben werden, Zwischenpositionen müssen jedoch durch Interpolation errechnet werden. Daher ist es hilfreich, die genauen Positionen des HRTF-Datensatzes zu kennen, um so mögliche Interpolationsfehler vermeiden zu können. Um diese Informationen aus dem HRFT-Datensatz auslesen zu können, wurde eine Anwendung³ in MATLAB geschrieben. Durch Einbindung der SOFA-API (Sofaconventions, 2020) und der ITA-Toolbox (Berzborn et al., 2017), kann das SOFA-Dateiformat in MATLAB verarbeitet werden.

Abbildung 13 stellt zunächst alle erfassten Positionen des HRTF-Datensatzes dar. Die räumliche Auflösung ist ausreichend hoch, um die nötigen Lautsprecherpositionen abzubilden. Um jedoch sicherzustellen, dass keine Ungenauigkeiten durch Interpolation auftreten, wurden die relevanten Messpunkte für diejenigen Lautsprecherpositionen ausgelesen, die in den zulässigen Bereichen der *ITU-R BS.2251-2* liegen und in den MAX-Patch übertragen. Die genutzten Messpunkte beider Lautsprecherkonfigurationen sind in Abbildung 15 dargestellt.

Im MAX-Patch kann über die Schaltfläche *Spat5-Viewer* eine Positionierungsübersicht angezeigt werden, die der Überprüfung der virtuellen Lautsprecherpositionen dient. Sie ist in Abbildung 16 und Abbildung 17 dargestellt. Da eine Anpassung der Positionen mit der Maus möglich ist, können sie im MAX-Patch über eine entsprechende Schaltfläche zurückgesetzt werden.

³ Der gesamte Quellcode der Anwendung findet sich im Anhang dieser Arbeit (A1.1 SOFA-Analyse).



Abbildung 13: Messpunkte des HRTF-Datensatzes der virtuellen Abhörumgebung.



Abbildung 14: HRTF-Messpunkte der Lautsprecherkonfiguration (4+5+0).


Abbildung 15: HRTF-Messpunkte der Lautsprecherkonfiguration (4+7+0).



Abbildung 16: Virtuelle Lautsprecherpositionen (4+5+0) in MAX.



Abbildung 17: Virtuelle Lautsprecherpositionen (4+7+0) in MAX.

7 Mikrofon- und Aufnahmetechnik

Für die Umwandlung von Schall in elektrische Schwingungen gibt es eine ganze
Reihe von Verfahren mit jeweils verschiedenen technischen Realisierungen.
Das führt zu einer relativ großen Vielfalt an Mikrofontypen, von denen jedes
bestimmte Vor- und Nachteile hat, sich also für konkrete Situationen mehr oder
weniger eignet (Dickreiter, 1995, S. 81).

Die primäre Funktion von Mikrofonen ist die Umwandlung von Schall in elektrische Schwingungen. Neben physikalisch-technischen Parametern muss beim Einsatz von Mikrofontechnik auch stets die Aufnahmesituation, die Praktikabilität und die künstlerische Gestaltung berücksichtigt werden. Durch die Wahl und den Aufbau der Mikrofontechnik wird das spätere Klangbild durch den Tonschaffenden bestimmt. Dabei kann es nicht eine "richtige" Lösung geben und jede Aufnahmesituation sollte individuell betrachtet werden. Besonders im Rahmen der mobilen Tonaufzeichnung müssen stets Aspekte der Praktikabilität berücksichtigt werden. Dieses Kapitel stellt die technischen Daten von Mikrofonen heraus, die speziell für Atmosphärenaufnahmen bedeutsam sind. Darüber hinaus werden die Konzepte einiger Mikrofonsysteme betrachtet, welche gezielt für den Einsatzweck der Atmosphärenaufnahme entwickelt wurden.

7.1 Anforderungen an die Mikrofontechnik

7.1.1 Technische Anforderungen

Bei der Tonaufzeichnung im Freien sind die Mikrofone den Umweltbedingungen ausgesetzt. Insbesondere durch Wind kann es leicht zu tieffrequenten Störungen und Übersteuerungen kommen. Daher ist ein hochwirksamer Windschutz für die erfolgreiche Atmosphären-Aufzeichnung von großer Bedeutung. Grundsätzlich können bei der Unterdrückung von Windstörungen zwei übliche Vorgehensweisen unterschieden werden. Dies sind die Umhüllung des Wandlers mit einem offenporigen Schaumstoff oder die Unterbringung des Mikrofons in einem geschlossenen Volumen mit Hilfe eines Windkorbs (Schneider, 2008). Wuttke (2000) merkt diesbezüglich an, dass für Druckempfänger ein Schaumstoffwindschutz wirksamer ist und für Druckgradientenempfänger ein Windschutzkorb. Das Aufbringen eines zusätzlichen langhaarigen Fells kann in beiden Fällen Windturbulenzen weiter reduzieren. Es gilt bei der Anwendung zu berücksichtigen, dass Windschütze sich stets auf den Frequenzgang der Mikrofone auswirken. Schaumstoffe führen zu einer Absenkung der Höhen. Durch Windschutzkörbe kommt es dagegen zu einer Welligkeit im Frequenzgang bei hohen Frequenzen und bei Druckgradientenempfängern darüber hinaus zu einer Tiefenabsenkung, dabei wirken sich große Windkörbe weniger stark aus als kleine. Die Notwendigkeit eines Windschutzes sollte daher stets vor Ort geprüft werden, um negative Einflüsse auf den Frequenzgang durch einen unnötigen Windschutz zu vermeiden. Grundsätzlich sind Druckempfänger weniger anfällig für Windstörungen, was bei

Des Weiteren sind elastische Mikrofonaufhängungen empfehlenswert, um Störungen durch Trittschall zu vermeiden. Erschütterungen und Windböen können insbesondere bei großen Stativen oder Auslegern für tieffrequente Störungen sorgen. Auch in diesem Zusammenhang zeigen sich Druckempfänger als weniger sensibel, was auf die im Vergleich zu Druckgradientenempfängern festere Membranaufhängung zurückzuführen ist (Ballou & Safari, 2012).

Ein für die Mikrofonen weiteres wichtiges Kriterium Auswahl von für Atmosphärenaufnahmen ist die Empfindlichkeit. Sie wird auch als Feld-Übertragungsfaktor bezeichnet und gibt die effektive Wechselspannung an, die ein Mikrofon abgibt, wenn es einem Schalldruck von 1 Pa (94 dB SPL) ausgesetzt wird. Er wird in der Regel für 1 kHz angegeben (Boré & Peus, 1999). Für Atmosphärenaufnahmen ist eine hohe Empfindlichkeit nützlich, da die auftretenden Schallpegel meist gering sind, wie beispielsweise bei der Aufnahme einer ruhigen Natur-Atmosphäre. Ferner sind die vom Mikrofon selbst abgegebenen Störgeräusche zu berücksichtigen. Diese sollte zur Aufnahme leiser Atmosphären möglichst ausfallen. In Datenblättern hierzu der gering ist Ersatzgeräuschpegel und/oder der Geräuschpegelabstand finden. Der $\mathbf{z}\mathbf{u}$ Ersatzgeräuschpegel beziffert dabei den Pegel der Geräuschspannung ohne Einwirkung von Schall (Dickreiter, 1995) und der Geräuschpegelabstand das in dB ausgedrückte Verhältnis von Ersatzgeräuschpegel und Feld-Übertragungsfaktor.

Ferner wird für die Auswahl von Mikrofonen oftmals der Frequenzgang herangezogen. Diesbezüglich gilt es zu beachten, dass der Frequenzgang eines Mikrofons im Reflexionsarmen Raum in Richtung der Hauptachse bei kurzen Abständen zur Schallquelle gemessen wird. Daher besitzt er nur für den Direktschall einer Quelle Gültigkeit. Die Frequenzgänge für seitlichen Schalleinfall können variieren. Der Frequenzgang für diffusen Schall kann völlig anders verlaufen. In der Regel wird der Diffusfeld-Frequenzgang jedoch nicht im Datenblatt angegeben. Für Atmosphärenaufnahmen befinden sich die Mikrophone meist nicht im Nahfeld der Schallquellen, weshalb der Frequenzgang für Direktschall wenig Aussagekraft besitzt. Abweichende Diffusfeld-Frequenzgänge sind daher für die unterschiedliche Klanglichkeit zweier Mikrofontypen verantwortlich, die ansonsten über identische Frequenzgänge laut Datenblatt verfügen (Boré & Peus, 1999; Wuttke, 2000).

7.1.2 Anforderungen für Atmosphären als Sounddesignkomponente

Ausgehend von den Ausführungen der vorangegangenen Kapitel werden an dieser Stelle die wichtigsten Anforderungen zusammengefasst, die sich durch die Funktion der Atmosphären im Sounddesign-Prozess für die Aufnahmetechnik ergeben.

- Gleichmäßige, homogene Abbildung ohne Vorzugsrichtung → Gleichmäßiger Aufbau des Mikrofonsystems, Verzicht auf den Center-Kanal.
- Möglichst diffuse Abbildung \rightarrow Großer Abstand zu konkreten Schallquellen, möglichst wenig Direktschall aufnehmen.
- Die Erzeugung umhüllender Schallfelder erfordert wiedergabeseitig möglichst dekorrelierte Signale → Hohe Kanaltrennung der Mikrofonsignale, große Pegel- und Laufzeitunterschiede zwischen den Kanälen des Mikrofonsystems.
- Praktikabilität bei Aufnahmen im Freien → Aufnahmeequipment muss mobil einsetzbar sein inklusive mobiler Stromversorgung, Betriebssicherheit und Schutz vor Wind- und Körperschall.

7.2 Spezielle Mikrofonsysteme zur Atmosphärenaufnahme

In diesem Abschnitt werden exemplarische Mikrofonsysteme vorgestellt, die speziell für die Anwendung der Atmosphärenaufnahme entwickelt wurden.

7.2.1 ESMA-3D

Die Grundidee des "Equal Segment Microphone Array" (ESMA) ist, dass alle nebeneinander liegenden Mikrofone einen gleich großen Öffnungswinkel besitzen. Hierdurch entstehen Stereo-Segmente, deren Aufnahmewinkel alle gleich groß sind (Lee, 2018). Dabei kann die Anzahl verwendeter Mikrofone variieren. Ziel ist eine möglichst gleichmäßige Abbildung und Lokalisation über den 360°-Raum. Die Anzahl der Stereo-Segmente, welche im Mikrofonarray enthalten sind, bestimmt dabei die Auflösung. Idealerweise werden die Aufnahmen des "ESMA-Array" über ein Lautsprechersystem wiedergegeben, das eine zum Mikrofonsystem identische Winkelanordnung besitzt (Lee, 2016). Beim "ESMA-3D" steht somit eine hohe Lokalisationsgenauigkeit im Vordergrund. Für den Aufbau kommen Richtmikrofone mit Nierencharakteristik zum Einsatz und die Kapselabstände können variieren. In einem Vergleich mehrerer "Esma-3D"-Mikrofonarrays mit 4 Nierenmikrofonen und Kapselabständen von 24, 30 und 50 cm wurde festgestellt, dass ein Aufbau mit 50 cm Kapselabstand die höchste Lokalisationsschärfe und gleichmäßigste Abbildung liefert (Lee, 2016). Zur Erfassung der elevierten Kanäle wird ein koinzidentes Verfahren eingesetzt. Die Nierenmikrofone des Bottom-Layers können dazu um eine weitere Niere ergänzt werden, wodurch sich in der Vertikalebene ein X/Y-Aufbau ergibt. Diesen Aufbau veranschaulicht Abbildung 18. Alternativ kann ein Mikrofon mit Achtercharakteristik eingesetzt werden, wodurch in der Vertikalebene ein M/S-System mit hoher Kanaltrennung entsteht. Dies ermöglicht die nachträgliche Anpassung der Abbildungseigenschaften in der Postproduktion (Lee, 2018). Das Mikrofonsystem "ESMA-3D" zeichnet sich durch einen kompakten Aufbau aus, der in der Praxis hilfreich sein kann. Im Fokus steht eine hohe Lokalisationsschärfe, weshalb es sich primär für Atmosphärenaufnahmen eignet, die von einer hohen Anzahl konkreter, sich bewegender Schallquellen geprägt sind. Es gilt zu berücksichtigen, dass es bei der Wiedergabe zu Verzerrungen der Lokalisation kommt, wenn die Wiedergabewinkel nicht denen der Aufnahmewinkel entsprechen, was beim (4+5+0)-Wiedergabesystem der Fall ist.



Abbildung 18: Koinzidente Mikrofonanordnung in der Vertikalebene beim Mikrofonsystem "ESMA 3D" (Lee, 2018).

7.2.2 Ortf-3D

Das Mikrofonsystem "Ortf-3D" von Schoeps zeichnet sich durch einen sehr kompakten Aufbau aus, bei dem alle Mikrofone in einem gemeinsamen Windkorb mit elastischer Aufhängung untergebracht werden. Im Rahmen der mobilen Atmosphärenaufnahme kann es so mit nur einem Stativ eingesetzt werden und überzeugt mit hoher Praktikabilität. Es wird als "8-Kanal-Array mit hoher Signaltrennung, optimaler Diffusfeld-Dekorrelation und guter 3D-Abbildung bei hoher Stabilität im Wiedergaberaum" (Wittek & Theile, 2016, S. 3) bezeichnet. Das "ORTF-3D" besteht aus zwei Layern, wobei jeweils vier Mikrofone mit Supernierencharakteristik zum Einsatz kommen. Auf ein Mikrofon für den Center-Kanal wird verzichtet. Pro Layer werden die Mikrofone in einem Rechteck mit Kantenlängen von zehn mal zwanzig Zentimetern angeordnet (Wittek & Theile, 2016). Dabei besteht kein Abstand zwischen den Layern, wodurch sich in der Vertikalebene ein koinzidentes und horizontal ein gemischtes Stereoverfahren bildet. Den Aufbau zeigt Abbildung 19. Die Mikrofone werden um 45° nach oben bzw. unten geneigt, wodurch sich in der Vertikalebene ein XY-Stereopaar mit einem Öffnungswinkel von 90° ergibt. Die hohe Richtwirkung der Supernierenmikrofone sorgt laut Wittek (2015) trotz der koinzidenten Anordnung für eine gute Abbildung und Diffusfeld-Dekorrelation. Im Bottom-Layer des "ORTF-3D" kommen kompakte Kleinmembran-Kondensatormikrofone mit frontaler Besprechungsrichtung zum Einsatz (Schoeps CCM 41). Im Height-Layer werden dagegen Mikrofone mit einer seitlichen Besprechungsrichtung eingesetzt (Schoeps CCM 41V). Dieser Umstand ermöglicht es, die Mikrofone kompakt übereinander zu platzieren und die Bauform des Mikrofonsystems klein zu halten.



Abbildung 19: Mikrofonsystem "ORTF 3D" im Windkorb (Wittek, o. J.).

7.2.3 Doppel U-Fix

Das Mikrofonsystem "Doppel U-Fix" ist eine Entwicklung von Camerer (2020) und wurde für die Wiedergabe über ein (4+5+0)-Wiedergabesystem konzipiert. Auffällig ist, dass ein zusätzliches Mikrofon für den Center-Kanal verwendet wird. Insgesamt kommen neun Mikrofone mit der Richtcharakteristik breite Niere zum Einsatz. Die Überlegung dabei ist, möglichst nahe andie hochwertige Tieftonübertragung von Kugelmikrofonen heranzukommen und dabei dennoch eine Richtwirkung aufrechtzuerhalten. Diese Richtwirkung ist dafür verantwortlich, dass trotz hoher Kanaltrennung die Abmessungen des Gesamtsystems überschaubar bleiben. Die genauen Kapselabstände und Versatzwinkel sind in Abbildung 20 aufgeführt. Der Abstand zwischen den Layern beträgt 25 bis 30 cm. Hauptaugenmerk bei der Konzeption des Mikrofonsystems ist die bestmögliche Anpassung von Aufnahme und Wiedergabewinkeln. Diese Abstimmung ermöglicht eine sehr gute Reproduktion bewegter Schallquellen bei geringer Winkelverzerrung. Dadurch ist dieses Mikrofonsystem besonders prädestiniert für die Aufzeichnung abwechslungsreicher und lebendiger Atmosphären mit vielen, sich bewegenden, konkreten Schallquellen. Weiterhin zeichnet sich das "Doppel U-Fix" durch eine hohe Praktikabilität aus. Es ist möglich, das gesamte Mikrofonarray auf einem Stativ zu platzieren. Die Mikrofone des Bottom- und Height-Layers sind jeweils in einem eigenen Windkorb mit elastischer Aufhängung untergebracht. Durch die Vorverkabelung mittels Multicore-Kabel ist das "Doppel U-Fix" besonders schnell einsatzbereit. Abbildung 21 zeigt das Mikrofonsystems während einer Atmosphärenaufnahme im Wald.



Abbildung 20: Kapselabstände und Versatzwinkel des Mikrofonsystems "Doppel U-Fix", links: Bottom-Layer, rechts: Height-Layer (Camerer, 2020).



Abbildung 21: Das Mikrofonsystem "Doppel U-Fix" im Einsatz (Camerer, 2020).

8 Testaufnahmen und Analyse

So, in the end, the purpose of field recording is to apply thought to the action of recording, to reflect upon the sound and its expression, and to consider how all of this affects those who hear it (Virostek, 2013, S. 67).

Zum Zwecke der praktischen Erprobung verschiedener Mikrofonsysteme zur Atmosphärenaufnahme im Freien wurden im Rahmen dieser Arbeit verschiedene Testaufnahmen durchgeführt. Auszüge dieser Aufnahmen dienen als Begleitmaterial zur Arbeit und ermöglichen es dem Leser, einen Eindruck von den klanglichen Abbildungseigenschaften verschiedener Mikrofonsysteme zu bekommen. Darüber hinaus werden die Aufnahmen einer technischen Signalanalyse unterzogen, um zu untersuchen, inwieweit sich theoretische Überlegungen zu angestrebten Abbildungseigenschaften in der Praxis widerspiegeln und ob diese bei der Planung von Mikrofonsystemen für Atmosphärenaufnahmen helfen können.

8.1 Dokumentation der erstellten Testaufnahmen

Dieses Kapitel gibt einen Überblick der Rahmenbedingungen, unter denen die Testaufnahmen entstanden sind. Aufnahmeorte und -zeiten sowie die eingesetzte Aufnahmetechnik werden vorgestellt.



Abbildung 22: Atmosphärenaufnahme im Reinhardswald.

8.1.1 Konzept

Im Rahmen der Testaufnahmen wurden sechs Mikrofonsysteme praktisch erprobt. Dabei kamen Mikrofonarrays zum Einsatz, die für die Wiedergabe über ein (4+5+0)- oder (4+7+0)-Wiedergabesystem ausgelegt sind. Aus Gründen der Materialverfügbarkeit und Praktikabilität war es nicht möglich, alle Mikrofonsysteme parallel aufzuzeichnen. Daher wurden die verschiedenen Mikrofonarrays unmittelbar nacheinander am selben Aufnahmeort aufgezeichnet. Die Testaufnahmen entstanden unter realen Bedingungen im Freien, weshalb ausschließlich auf mobiles Equipment zurückgegriffen wurde. Um dennoch einen direkten Vergleich zu ermöglichen, wurde parallel zu jedem Mikrofonsystem sowohl ein Kunstkopf (Neumann KU81i) als auch ein A-Format Ambisonics-Mikrofon (Sennheiser AMBEO VR-Mikrofon) aufgezeichnet. Die Kunstkopfaufnahmen werden im weiteren Verlauf dieser Arbeit für signaltechnische Untersuchungen genutzt. Die Aufnahmen des Ambisonics-Mikrofons werden für zukünftige Vergleiche zur Verfügung gestellt. Die Decodierung der aufgezeichneten A-Format-Signale stellt ein umfangreiches Themengebiet dar und kann im Rahmen dieser Arbeit nicht in der nötigen Tiefe behandelt werden. Insgesamt wurden Aufnahmen an drei unterschiedlichen Orten durchgeführt. Es wurden Takes von jeweils mindestens 15 Minuten Dauer erstellt, die nachträglich auf circa zwei Minuten gekürzt wurden. Dies diente dazu, unerwünschte Störgeräusche, wie Flugzeuge, vorbeifahrende Autos und starke Windstörungen zu entfernen. Es wurde mit einer Bittiefe von 24 Bit und einer Abtastrate von 48 kHz aufgezeichnet.

8.1.2 Aufnahmeorte

Die ersten Testaufnahmen erfolgten im Habichtswald bei Dörnberg in der Nähe von Kassel (Ort 1). Die genauen Koordinaten sind Abbildung 23 zu entnehmen. Die Aufnahmen an diesem Ort entstanden in der Nacht vom 03.11.2020 auf den 04.11.2020, zwischen 23:00 und 2:30 Uhr. Dieser Zeitraum wurde gewählt, um Störeinflüsse durch Passanten und Verkehr zu minimieren. Die Wetterlage war winterlich frisch, mit vereinzelten Regenschauern und frischem, böigem Wind. Die hier entstandenen Aufnahmen sind vornehmlich geprägt vom Rascheln der Blätter und Zweige im Wind. In Windpausen ist entfernter Autobahnverkehr wahrzunehmen.



Bilder © 2020 GeoBasis-DE/BKG,GeoContent,Maxar Technologies,Kartendaten © 2020 GeoBasis-DE/BKG (©2009) 200 m Abbildung 23: Aufnahmeort 1 – Habichtswald (Quelle: Google Maps).

Weitere Aufnahmen wurden im Reinhardswald erstellt (Ort 2). Mit einer Fläche von über 200 km² stellt er eines der größten zusammenhängenden Waldgebiete Deutschlands dar und die Umgebung ist nur dünn besiedelt. Damit bietet dieses Gebiet grundsätzlich gute Voraussetzungen für die Aufnahme natürlicher Waldatmosphären in Deutschland. Der genaue Standort der Aufnahme ist in Abbildung 24 dargestellt. Die Aufzeichnungen erfolgten am 04.11.2020 zwischen 14:00 und 17:00 Uhr. Die Wetterlage war insgesamt ruhig, mit nur leichter Bewölkung und einer schwachen, böigen Brise. Die hier entstandenen Aufnahmen werden durch leichten Wind und Blätterrascheln sowie verschiedene Vogelrufe geprägt. Jedoch sind in der Ferne vereinzelt auch ungewünschte Geräusche von Forstarbeiten zu hören.



Neben naturnahen Aufnahmen wurden zusätzlich Atmosphären in einem urbanen Gebiet erstellt. Um eine Mindestdistanz zu konkreten Klängen einhalten zu können und um Störungen und Unterbrechungen durch Passanten zu vermeiden, wurde ein erhöhter Standort auf einem Parkhausdeck in der Kasseler Innenstadt gewählt (Ort 3). Der Standpunkt der Mikrofone ist in Abbildung 25 eingezeichnet. Die Aufnahmen an diesem Standort erfolgten am 05.11.2020 zwischen 14:30 und 18:00 Uhr. Die Wetterlage war ruhig und es kam lediglich leiser Zugwind auf. Darüber hinaus war es sehr trüb und es herrschte dichter Nebel. Die Fußgängerzone der Innerstadt war nur mäßig besucht, weshalb die Atmosphären vor allem durch die Geräuschkulisse der Straßenbahn sowie entfernter Hauptverkehrsstraßen geprägt werden. Vereinzelt sind Passanten, Baulärm und Rabenvögel zu hören.



Bilder © 2020 Google,Bilder © 2020 AeroWest,GeoBasis-DE/BKG,GeoContent,Maxar Technologies,Kartendaten © 2020 20 m GeoBasis-DE/BKG (© 2009)

Abbildung 25: Aufnahmeort 3 - Kasseler Innenstadt, Parkhausdeck (Quelle: Google Maps).

8.1.3 Aufnahmetechnik

Für die mobile Tonaufzeichnung kamen drei Recorder der Firma Zoom zum Einsatz. Dabei handelte es sich um zwei Achtkanal-Recorder vom Typ F8n sowie einen Vierkanal-Recorder vom Typ F4. Die Recorder wurden per Timecode gekoppelt, um einen synchronen Aufzeichnungsbeginn aller Geräte zu gewährleisten. Die Stromversorgung erfolgte über einen externen Akkumulator. Innerhalb eines Recorders konnten die Gain-Regler gekoppelt werden, und ermöglichten so eine gleichmäßige Vorverstärkung der angeschlossenen Mikrofone. Die benötigte 48V-Phantomspeisung wurde von den Recordern bereitgestellt. Die eingesetzten Mikrofontypen sind in Tabelle 4 aufgelistet. Der Aufbau der Mikrofonsysteme wird in Abschnitt 0 erläutert. Dabei ist zu berücksichtigen, dass sicherlich viele weitere Mikrofonarrays für eine praktische Erprobung in Frage kommen würden, im Rahmen dieser Arbeit jedoch nur eine Auswahl getestet werden konnte. Es wurde versucht, mit Hilfe des zur Verfügung stehenden Materials eine möglichst große Bandbreite an Mikrofonsystemen abzudecken.

Nr.	Mikrofonsysteme	Mikrofontypen	Details
01	Hamasaki + Height	4x Neumann U89i, 4x Schoeps CMC5/MK4	 4x Großmembran-Kondensator mit Wechselcharakteristik (hier: Acht), kombinierter Fell- Schaum Windschutz 4x Kleinmembrankondensator, Niere, kombinierter Fell-Schaum Windschutz
	Kunstkopf	Neumann KU81i	
	Ambisonics A-Format	Sennheiser Ambeo VR	kombinierter Fell-Schaum Windschutz
02	Kugel-Array	8x Schoeps CMC5/MK2	8x Kleinmembrankondensator, Kugel, kombinierter Fell-Schaum Windschutz
	Kunstkopf	Neumann KU81i	
	Ambisonics A-Format	Sennheiser Ambeo VR	kombinierter Fell-Schaum Windschutz
03	Grenzflächen-Array	6x Schoeps BLM 03C 4x Schoeps CMC5/MK2	 6x Kleinmembrankondensator, kleine Grenzfläche, Fellwindschutz 4x Kleinmembrankondensator, Kugel, kombinierter Fell-Schaum Windschutz
	Kunstkopf	Neumann KU81i	
	Ambisonics A-Format	Sennheiser Ambeo VR	kombinierter Fell-Schaum Windschutz
04	Nieren-Array	8x Schoeps CMC5/MK4	8x Kleinmembrankondensator, Niere, kombinierter Fell-Schaum Windschutz
	Kunstkopf	Neumann KU81i	
	Ambisonics A-Format	Sennheiser Ambeo VR	kombinierter Fell-Schaum Windschutz
05	Nieren-Array II	10x Schoeps CMC5/MK4	10x Kleinmembrankondensator, Niere, kombinierter Fell-Schaum Windschutz
	Kunstkopf	Neumann KU81i	
	Ambisonics A-Format	Sennheiser Ambeo VR	kombinierter Fell-Schaum Windschutz
06	Doppel IRT-Kreuz	8x Schoeps CMC5/MK4	8x Kleinmembrankondensator, Niere, kombinierter Fell-Schaum Windschutz
	Kunstkopf	Neumann KU81i	
	Ambisonics A-Format	Sennheiser Ambeo VR	kombinierter Fell-Schaum Windschutz

Tabelle 4:Mikrofonsysteme der Testaufnahmen.

8.1.4 Aufbau der verwendeten Mikrofonsysteme

In diesem Kapitel werden der systematische Aufbau der verwendeten Mikrofonsysteme sowie praktische Hintergründe zur Umsetzung vorgestellt. Anhand von grafischen Darstellungen können Kapselabstände, Versatzwinkel und Richtcharakteristiken abgelesen werden. Dabei werden die in Tabelle 4 vorgestellten Bezeichnungen der Mikrofonsysteme übernommen.

8.1.4.1 Hamasaki + Height

Das Mikrofonsystem "Hamasaki + Height" besteht aus einem Hamasaki-Square, welches um vier Höhenkanäle erweitert wird. Das Hamasaki-Square wird aus vier Mikrofonen mit aufgebaut, für die Achter-Charakteristik Höhenkanäle können verschiedene Druckgradientenempfänger zum Einsatz kommen. Für die Aufnahmen dieser Arbeit wurden Mikrofone mit Nierencharakteristik gewählt. Das Hamasaki-Square wird häufig für die Aufnahme von Diffusschall in Konzertsälen eingesetzt und bietet durch die Achtercharakteristik die Möglichkeit, frontalen und rückseitigen Schall auszublenden. Den schematischen Aufbau des Mikrofonsystems zeigt Abbildung 26. Es wurde ein quadratischer Grundaufbau gewählt. Der Kapselabstand betrug 2,5 Meter innerhalb des Height- und Bottom-Layers. Der Abstand zwischen den Layern betrug einen Meter. Die Nierenmikrofone des Height-Layers wurden um 45° abgewinkelt.



Abbildung 26: Mikrofonsystem Hamasaki + Height.

8.1.4.2 Kugel-Array

Dieses System wurde aus acht Druckempfängern mit Kugelcharakteristik aufgebaut. Es ist auch unter dem Namen 2L-Cube (benannt nach Morten Lindberg) bekannt (Kim, 2018). Durch die Kugelcharakteristik sind einzig Laufzeitunterschiede für die Dekorrelation der Mikrofonkanäle verantwortlich. Um auch für tiefe, langwellige Frequenzen dekorrelierte Signale aufzeichnen zu können, müssen entsprechend große Mikrofonabstände gewählt werden. Dies führt zu einem großen Aufbau des Mikrofonsystems, wodurch die Praktikabilität eingeschränkt wird. Reine Druckempfänger zeichnen sich jedoch durch ihre gute Tieftonwiedergabe aus und werden aufgrund ihrer räumlichen Abbildungsqualitäten auch häufig für Musikaufnahmen eingesetzt (Dickreiter, 1995). Abbildung 27 schematisiert den Aufbau des Mikrofonsystems. Innerhalb der beiden Layer wurden Kapselabstände von vier Metern gewählt. Der Abstand zwischen Bottom- und Height-Layer wurde durch den maximalen Auszug der Stative auf 3 m begrenzt. Als Mindestabstand der Mikrofone des Bottom-Layers vom Boden wurden hierbei 0,5 Meter festgelegt. Um eine die Homogenität innerhalb des Systems zu wahren, wurden keine größeren Abstände zwischen den Stativen genutzt. Zudem haben Untersuchungen von Rumsey und Lewis (2002) ergeben, dass Aufnahmen mit Systemen, welche noch größere Kapselabstände verwenden, subjektiv schlechter bewertet werden.



Abbildung 27: Mikrofonsystem Kugel-Array.

8.1.4.3 Grenzflächen-Array

Beim "Grenzflächen-Array" handelt es sich um eine Abwandlung des "Kugel-Arrays". Die vier Mikrofone des Bottom-Layers werden durch sechs Grenzflächen-Mikrofone, die hexagonal angeordnet sind, ersetzt. Aufnahmen dieses Mikrofonsystems sind für die Wiedergabe über ein (4+7+0)-Lautsprechersystem bestimmt. Die Grenzflächen werden direkt auf dem Boden platziert, wodurch der Mehraufwand durch zwei zusätzliche Mikrofone überschaubar bleibt, denn es sind keine zusätzlichen Stative nötig. Darüber hinaus erhöht sich hierdurch der vertikale Kapselabstand um 0,5 Meter. Die Grenzflächen-Mikrofone sollten auf einer möglichst ebenen, schallharten Fläche angebracht werden (Dickreiter, 1995). Dies ist nicht in jeder Aufnahmesituation im Freien zu realisieren. Darüber hinaus sind beispielsweise feuchte, bewachsene oder mit Laub bedeckte Flächen nur bedingt schallhart, wodurch die Einsatzmöglichkeiten von Grenzflächen-Mikrofonen für Atmosphärenaufnahmen limitiert sind. Der Aufbau des "Grenzflächen-Arrays" ist Abbildung 28 zu entnehmen. Der Aufbau des Height-Layers entspricht dem des "Kugel-Arrays". Die Grenzflächen des Bottom-Layers sind in einem gleichmäßigen Sechseck angeordnet, die Kapselabstände betragen auch hier vier Meter.



Abbildung 28: Mikrofonsystem Grenzflächen-Array.

8.1.4.4 Nieren-Array I

Das "Nieren-Array I" wird aus acht Mikrofonen mit Nieren-Charakteristik aufgebaut. Wie in Abbildung 29 dargestellt wird, ist das Array vollständig symmetrisch in Form eines Würfels aufgebaut. Die Kapselabstände betragen jeweils einen Meter und die Öffnungswinkel jedes Mikrofonpaares 90°.



Abbildung 29: Mikrofonsystem Nieren-Array I.

8.1.4.5 Nieren-Array II

Für die Wiedergabe über ein (4+7+0)-Lautsprechersystem wird das "Nieren-Array I" um zwei weitere Mikrofone mit Nierencharakteristik im Bottom-Laver erweitert. Durch die zusätzlichen Mikrofone ändern sich die Öffnungswinkel in diesem Layer von 90° auf 60°. Somit stehen zwei weitere Signale zur Verfügung. Aufgrund der geringeren Öffnungswinkel ist jedoch eine höhere Korrelation der Mikrofonpaare des Bottom-Layers zu erwarten. Es stellt sich somit die Frage, ob die zusätzlichen Mikrofone des "Nieren-Array II" vorteilhaft sind. Wie Abbildung 30 zu entnehmen ist, erfolgt der Aufbau nicht in einem gleichmäßigen Sechseck, sondern in asymmetrischer Ausführung. Zwei Mikrofone erweitern den Bottom-Layer des "Nieren-Array I" zu den Seiten hin, wobei Kapselabstände von einem Meter eingehalten werden. Diese Form des Aufbaus ergibt sich aus praktischen Überlegungen: Der Grundaufbau des "Nieren-Array I" mit vier Stativen kann unter Anpassung der Öffnungswinkel übernommen werden. Es werden lediglich zwei weitere Stative für die zusätzlichen Mikrofone benötigt. Gerade im Außeneinsatz mit mobilem Equipment ist geringer Material- und Transportaufwand ein wichtiges Kriterium. Sollte dagegen ein gleichmäßiger, hexagonaler Aufbau realisiert werden, wären sechs Stative für den Bottom-Layer und vier Stative für den Height-Layer nötig.



Abbildung 30: Mikrofonsystem Nieren-Array II.

8.1.4.6 Doppel-IRT

Das Mikrofonsystem "Doppel-IRT" ist als 3D-Erweiterung des bekannten IRT-Kreuz zur Atmosphärenaufnahme zu verstehen (Weinzierl, 2008). Dazu wird jeweils ein IRT-Kreuz für den Bottom- und Height Layer verwendet. Dieses System ist aus insgesamt acht Mikrofonen mit Nierencharakteristik aufgebaut. Die Mikrofone sind um 45° nach oben bzw. unten abgewinkelt, sodass sich auch für die Mikrofonpaare aus Bottom- und Height-Layer ein Öffnungswinkel von 90° ergibt. Die Kapselabstände betragen überall 25 cm. Der Aufbau ist in Abbildung 31 skizziert. Es handelt sich um ein sehr kompaktes Mikrofonsystem, das besonders wegen seiner hohen Praktikabilität für Atmosphärenaufnahmen interessant ist. Mit Hilfe von mehreren Stereo-Schienen kann der gesamte Aufbau an nur einem Stativ befestigt werden, wodurch es sehr transportabel wird und sich auch in schwierigen Aufnahmesituationen mit beengten Platzverhältnissen einsetzen lässt.



Abbildung 31: Mikrofonsystem Doppel IRT-Kreuz.

8.2 Signaltechnische Analyse der Abbildungseigenschaften

Im Zusammenhang mit der Planung und Durchführung stereofoner Mikrofonaufnahmen ist die Abschätzung der zu erwartenden Abbildungseigenschaften von großer Relevanz für den Tonschaffenden. Bei der Betrachtung von Atmosphärenaufnahmen sind vornehmlich die räumlichen Abbildungseigenschaften von Interesse. Grundsätzlich wird versucht, über die Untersuchung der Ähnlichkeit zweier Mikrofonsignale eine Aussage über die zu erwartende räumliche Breite oder Umhüllung einer Mikrofonanordnung treffen zu können. Die räumliche Abbildung hängt von der Korrelation der Mikrofonsignale ab (Riekehof-Böhmer et al., 2010) und eine breite, umhüllende Abbildung erfordert eine möglichst geringe Korrelation (Griesinger, 2001). Wie bereits im Zusammenhang mit der interauralen Korrelation angesprochen wurde (siehe Kapitel 5.3), werden in der Literatur und Praxis verschiedene Methoden zur Untersuchung der Verwandtschaftsbeziehung von Signalen herangezogen. kurzen Vorstellung der mathematischen Grundlagen Nach einer verschiedener Vergleichsmethoden werden die Abbildungseigenschaften der praktisch erprobten

Mikrofonsysteme mit Hilfe der Kohärenzfunktion des *Schoeps Image Assist* abgeschätzt. In einem weiteren Schritt werden die Aufnahmen der eingesetzten Mikrofonsysteme anhand einer Kohärenzanalyse signaltechnisch ausgewertet. Dazu wurden verschiedene MATLAB-Skripte erstellt, mit denen die Ähnlichkeit zweier Mikrofonsignale untersucht werden kann. Durch die grafische Darstellung der Kohärenzfunktion wird ein Vergleich zwischen Vorhersage und realer Aufnahme ermöglicht. Dabei steht die Frage im Vordergrund, inwieweit sich die vorhergesagten Abbildungseigenschaften eines Mikrofonsystems anhand natürlich erstellter Atmosphärenaufnahmen im Freien nachvollziehen lassen.

8.2.1 Theoretische Grundlagen

8.2.1.1 Korrelationsgrad

In Kapitel 5.3 wurde die normierte Kreuzkorrelationsfunktion für kontinuierliche Signale angegeben (2). Erfolgt die Berechnung mit Hilfe von Digitaltechnik, so handelt es sich um zeitdiskrete Signale. Daher kann zusätzlich die normierte Kreuzkorrelationsfunktion für zeitdiskrete Signale angeben werden, welche auch als $p_{sg}^{E}(m)$ bezeichnet wird, mit den diskreten Signalen *s* und *g* und der Energie *E* (Riekehof-Böhmer et al., 2010). Sie ist definiert als:

$$p_{sg}^{E}(m) = \frac{\sum_{n=-\infty}^{+\infty} s^{*}(n) * g(n+m)}{\sqrt{E_{s} * E_{g}}}$$
(3)

Die Signaldauer τ wird also durch den Laufindex m ersetzt. Als Korrelationsgrad bezeichnet man den normierten Kreuzkorrelationskoeffizienten. Das ist der Wert, den die normierte Kreuzkorrelationsfunktion an der Stelle $\tau = 0$ bzw. m = 0 annimmt. Mit Hilfe eines Korrelationsgradmessers wird in der Studiotechnik häufig die Beziehung zweier Signale untersucht (Gernemann-Paulsen, o. J.). Ein Sinus- und ein Kosinussignal gleicher Frequenz ergeben einen Korrelationsgrad von 0, da diese Signale um 90° zueinander phasenverschoben sind. Identische Signale ergeben einen Korrelationsgrad von 1 und identische um 180° zueinander verschobene Signale ergeben einen Korrelationsgrad von -1. Der Korrelationsgrad ist in erster Linie dazu geeignet, die Phasenlage von zwei Signalen zueinander zu beschreiben und wird daher auch als Phasenmessgerät bezeichnet (Friesecke, 2007). Diese Geräte arbeiten in der Regel mit einer Integrationszeit von circa 0,5 Sekunden, weshalb für impulsartige Signale keine Aussage getroffen werden kann. Mikrofone liefern breitbandige Signale mit statistischem Charakter, weshalb eine momentane Phasendifferenz nur wenig aussagekräftig ist. Für koinzidente Mikrofonanordnungen kann die Betrachtung der Phasenlage sinnvoll sein, für räumlich getrennte Mikrofonanordnungen kann der Korrelationsgradmesser jedoch keine eindeutigen Aussagen treffen (Sengpiel, 2002). Darüber hinaus ist die Korrelation räumlich getrennter Mikrofonanordnungen generell frequenzabhängig (Elko, 2001). Durch die Normierung auf die Gesamtenergie bei der Berechnung des Kreuzkorrelationskoeffizienten findet die spektrale Leistungsverteilung der Signale keine Berücksichtigung (Riekehof-Böhmer et al., 2010).

8.2.1.2 Kohärenzgrad

Aus der Kreuzkorrelationsfunktion (3) kann der Kohärenzgrad k abgeleitet werden. Blauert (1974) definiert diesen als:

$$k = \max |p_{sg}^{E}(m)| \tag{4}$$

Die Maximalwertbildung führt dazu, dass keine Aussage mehr über die Phasenbeziehung der Signale getroffen werden kann. Der Kohärenzgrad ist ebenfalls abhängig von der Gesamtenergie innerhalb der Signale und die Frequenzabhängigkeit wird nicht berücksichtigt (Riekehof-Böhmer et al., 2010).

8.2.1.3 Kohärenzfunktion

Zur Berücksichtigung der Frequenzabhängigkeit kann die Kohärenzfunktion $C_{xy}(f)$ betrachtet werden, da sie eine Funktion der Frequenz ist. Nach Bendat und Piersol (1980) ist diese definiert als:

$$C_{xy}(f) = \frac{|P_{xy}(f)|^2}{P_{xx}(f) * P_{yy}(f)} = \frac{|P_{xy}(f)|}{\sqrt{P_{xx}(f) * P_{yy}(f)}}$$
(5)

Dabei sind $P_{xx}(f)$ und $P_{yy}(f)$ die Autoleistungsdichtespektren der Signale x bzw. y. Bei $P_{xy}(f)$ handelt es sich um das Kreuzleistungsdichtespektrum der Signale x und y. In der digitalen Signalverarbeitung existieren verschiedene Methoden zur Berechnung des Leistungsdichtespektrums. Eine häufig verwendete Berechnungsart ist die Methode nach Welch (1967), die beispielsweise in (Djuric & Kay, 2017) näher erläutert wird. Bei der hier vorgestellten Form der Kohärenzfunktion wird der des Betrag Kreuzleistungsdichtespektrums quadriert und normiert, weshalb auch die Bezeichnung magnitude-squared-coherence (msc) geläufig ist. Die Betragsbildung verhindert auch hier, dass eine Aussage über die Phasenlage der zwei Signale getroffen werden kann. Im Gegensatz zum Kohärenzgrad kann nun jedoch eine Betrachtung bei jeder Frequenz erfolgen.

8.2.1.4 Nicht-quadrierte Kohärenzfunktion

Elko (2001) führt zusätzlich die nicht quadrierte Variante $\gamma_{xy}(f)$ der Kohärenzfunktion ein. Sie ist definiert als:

$$\gamma_{xy}(f) = \frac{P_{xy}(f)}{\sqrt{P_{xx}(f) * P_{yy}(f)}} = \frac{P_{xy}(f)}{P_{xx}^{\frac{1}{2}}(f) * P_{yy}^{\frac{1}{2}}(f)}$$
(6)

Diese Variante der Kohärenzfunktion ermöglicht zusätzlich die Auswertung der Phasenlage von zwei Signalen. Simplifiziert ausgedrückt liefert die nicht quadrierte Kohärenzfunktion den Korrelationsgrad pro Frequenz (Riekehof-Böhmer et al., 2010).

8.2.1.5 Diffuse Field Image Prädiktor (DFI-P)

2010 stellten Riekehof-Böhmer et al. mit dem Diffuse Field Image Predictor (DFI-P) eine neue Option zur Voraussage der wahrgenommenen räumlichen Breite einer stereofonen Mikrofonanordnung Dabei erfolgt die Angabe der erwartenden vor. zu Abbildungseigenschaften anhand eines einzelnen Zahlenwerts, wodurch eine einfache und schnelle Vergleichbarkeit ermöglicht werden soll. Die Berechnung des DFI-P beruht auf der nicht quadrierten Kohärenzfunktion. Im nächsten Schritt wird die nicht quadrierte Kohärenzfunktion mit einer Wichtungsfunktion $\chi(f)$ gewichtet. Es werden nur Frequenzen zwischen 40 und 1500 Hz zur Berechnung des DFI-P herangezogen, da tieffrequente

Dekorrelation für die Wahrnehmung räumlicher Breite und Umhüllung von großer Relevanz ist (siehe Kapitel 5.3). Dabei werden 40 Hz mit +9 dB überbewertet und es folgt die Gewichtung mit -3 dB pro Oktave. Der DFI-P wird definiert als:

$$DFI-P = \frac{1}{n} * \sum_{f=40}^{1500 \, Hz} [\gamma_{xy}(f) * \chi(f)]^2 \tag{7}$$

Mit *n* als Anzahl der FFT-Stützstellen, $\gamma_{xy}(f)$ als nicht quadrierte Kohärenzfunktion und $\chi(f)$ als Wichtungsfunktion. Der DFI-P kann Werte zwischen 0 und 1 annehmen, wobei ein kleiner Wert auf eine geringe Korrelation im Frequenzbereich 40 bis 1500 Hz hinweist. Ein hoher Wert weist entsprechend auf eine hohe Korrelation in diesem Frequenzbereich hin (Riekehof-Böhmer et al., 2010).

8.2.2 Vorhersage von Abbildungseigenschaften mit Hilfe des Schoeps Image Assistant

Für die Planung und den Aufbau von Mikrofonsystemen ist es nützlich, schon im Vorfeld der Aufnahme die zu erwartenden Abbildungseigenschaften abschätzen zu können. Besonders verbreitet ist die Vorhersage des stereofonen Aufnahmewinkels und die Darstellung von Lokalisationskurven. Detaillierte Ausführungen dazu sind beispielsweise bei Williams (2004, 2013) zu finden. Heute existieren verschiedene Softwaretools und Apps⁴, welche, nach Angabe von Richtcharakteristik, Öffnungswinkel und Kapselabstand eines Mikrofonpaares, die entsprechenden Daten und Graphen generieren. Im Zusammenhang mit Atmosphärenaufnahmen sind vornehmlich die räumlichen Abbildungseigenschaften von Interesse, wenn ein möglichst hoher Grad an Umhüllung angestrebt wird. Das Tool Schoeps Image Assistant (SCHOEPS Mikrofone, o. J.-b) bietet unter anderem die Möglichkeit, die Kohärenzfunktion eines Mikrofonpaares im Diffusfeld darzustellen. Anhand dieser Funktionalität sollte es möglich sein, die Ähnlichkeit eines Mikrofonpaares innerhalb eines Mikrofonarrays abzuschätzen. Für die in Abschnitt 8.1.4 vorgestellten Mikrofonsysteme

⁴ Beispielsweise *Neumann Recording Tools* (Georg Neumann GmbH Berlin, 2020), *MARRS* (The University of Huddersfield Enterprises Ltd., 2020), *Schoeps Image Assistant* (SCHOEPS Mikrofone, 2020).

müssen demnach alle benachbarten Mikrofonpaare eines Arrays betrachtet werden. Dies schließt sowohl die Mikrofonpaare innerhalb des Height- bzw. Bottom-Layers als auch diejenigen Paare mit ein, die sich zwischen beiden Layern bilden lassen. Zur Benennung der Mikrofonsignale werden dabei die Bezeichnungen der Lautsprecherkanäle, denen die Signale bei der Wiedergabe zugeordnet werden, übernommen.



Abbildung 32: Kohärenzfunktionen des Schoeps Image Assistant für das Mikrofonsystem "Hamasaki + Height".

Für das Mikrofonsystem "Hamasaki + Height" ergeben sich vier verschiedene Mikrofonpaarungen, die betrachtet werden können. Die Kanäle L–Ls und R–Rs bilden ein Mikrofonpaar aus Achten mit einem Kapselabstand von 2,5 Metern und einem Öffnungswinkel von 0°. Die Kanäle L–R und Ls–Rs unterscheiden sich dahingehend, dass der Öffnungswinkel hier 180° beträgt. Innerhalb des Height-Layers ergeben sich durch den symmetrischen Aufbau vier Mikrofonpaare aus jeweils zwei Nieren mit einem Kapselabstand von einem Meter und einem Öffnungswinkel von 90°. Zwischen Bottom- und Height-Layer bilden die Signale L–Ltf, R–Rtf, Ls–Ltr und Rs–Rtr je ein Mikrofonpaar bestehend aus Niere und Acht mit einem Kapselabstand von 2,5 Metern und einem Öffnungswinkel von 45°. Die Verläufe der Kohärenzfunktionen, welche mit dem *Schoeps Image Assistant* generiert wurden, sind in Abbildung 32 übereinander aufgetragen. Es wird deutlich, dass alle Mikrofonpaare bis etwa 150 Hz eine geringe Kohärenz von unter 0,2 aufweisen. Zu tieferen Frequenzen hin steigen die Verläufe unterschiedlich stark an. Die Verschiedenheit der Mikrofonpaare spiegelt sich in der großen Varianz der Kurvenverläufe wider. Besonders die Mikrofonpaare L–R und Ls–Rs, mit ihren um 180° verdrehten Achten, fallen durch einen welligen Kurvenverlauf auf. Bei etwa 130 Hz tritt ein lokales Maximum mit einer Kohärenz von etwa 0,55 auf; hier liegt die Wellenlänge im Bereich des Kapselabstands von 2,5 Metern. Mit Ausnahme des Mikrofonpaares aus Niere und Acht steigen alle Kurven zu den tiefsten Frequenzen hin deutlich an. Insgesamt zeigt sich eine niedrige Kohärenz bis 150 Hz für alle Mikrofonpaare. Die Funktionsverläufe lassen darauf schließen, dass sich mit dem Mikrofonsystem "Hamasaki + Height" ausreichend dekorrelierte Signale aufzeichnen lassen, um ein umhüllendes Klangbild zu erzeugen. Dabei ist jedoch die Ausbildung eines lokalen Kohärenzmaximums bei 130 Hz für die Mikrofonpaare L-R und Ls-Rs als suboptimal einzustufen, da dieses in einem spektralen Bereich liegt, der für einen umhüllenden Klangeindruck besonders wichtig ist (siehe Kapitel 5.3). Die hohen Kohärenzwerte für Frequenzen unterhalb von 50 Hz liegen in einem spektralen Bereich, in dem für den Fall natürlicher Atmosphärenaufnahmen mit geringem Nutzpegel zu rechnen ist. Zudem ist zu beachten, dass Druckgradientenempfänger und Mikrofonverstärker in der Regel kein lineares Übertragungsverhalten bis 0 Hz aufweisen und sich stattdessen ein Hochpassverhalten zeigt (Wuttke, 2000).



Abbildung 33: Kohärenzfunktionen des Schoeps Image Assistant für die Mikrofonsysteme "Kugel-Array" und "Grenzflächen-Array".

Abbildung 33 zeigt die Verläufe der Kohärenzfunktion von Mikrofonpaaren mit Kugelcharakteristik bei verschiedenen Kapselabständen. Anhand dieser Kurven lassen sich Mikrofonsysteme "Kugel-Array" und "Grenzflächen-Array" näher betrachten. die Grenzflächen-Mikrofone weisen eine Halbkugelcharakteristik auf, werden hier jedoch anhand einer Kugelcharakteristik betrachtet, da der Schoeps Image Assistant keine spezielle Halbkugelcharakteristik anbietet. Sowohl beim "Kugel-Array" als auch beim "Grenzflächen-Array" betragen die Kapselabstände der Mikrofonpaare innerhalb des Bottom- bzw. Height-Layers vier Meter. Die Abstände in der Vertikalebene betragen drei Meter beim "Kugel-Array" und dreieinhalb Meter beim "Grenzflächen-Array". Es zeigt sich, dass alle Mikrofonpaare bis mindestens 50 Hz stark inkohärent sind. Größere Kapselabstände verschieben diese Grenze weiter zu tieferen Frequenzen. Die Unterschiede können insgesamt als gering beschrieben werden. Kleine Verbesserungen in Richtung geringerer Kohärenz bei tieferen Frequenzen gehen folglich mit sehr großen Mikrofonabständen einher, weshalb je Aufnahmesituation ein praktikabler Kompromiss gefunden werden muss. nach Zusammenfassend lassen die Funktionsverläufe darauf schließen, dass bereits mit Kapselabständen von drei Metern eine gute Dekorrelation bis 50 Hz erreicht wird. Daraus kann gefolgert werden, dass sich sowohl das "Kugel-Array" als auch "Grenzflächen-Array" gut eignen, um umhüllende Atmosphären aufzuzeichnen.



Abbildung 34: Kohärenzfunktionen des Schoeps Image Assistant für die Mikrofonsysteme "Nieren-Array I", "Nieren-Array II" und "Doppel-IRT".

Die Verläufe der Kohärenzfunktion verschiedener Mikrofonpaare mit Nierencharakteristik sind in Abbildung 34 dargestellt. Sie eigenen sich um Abbildungseigenschaften der Mikrofonsysteme "Nieren-Array I", "Nieren-Array II" und "Doppel-IRT" abzuschätzen. Da das "Nieren-Array I" vollständig symmetrisch aufgebaut ist, ergeben sich für jedes Mikrofonpaar dieselben Verhältnisse. Der Öffnungswinkel beträgt 90°, bei einem Kapselabstand von einem Meter. Die zugehörige Kohärenzfunktion zeigt im Vergleich die niedrigsten Kohärenzwerte. Bis circa 150 Hz sind die Signale vollständig inkohärent, danach steigt die Kohärenz steil an und flacht bei mittleren Kohärenzwerten ab. Durch zwei zusätzliche Kanäle im Bottom-Layer betragen die Öffnungswinkel im "Nieren-Array II" nur 60°. Dies äußert sich im Verlauf der Kohärenzfunktion in der Form, dass im Vergleich zum Öffnungswinkel von 90°, der Anstieg der Kohärenz ab circa 150 Hz steiler verläuft und mit 0,75 insgesamt höhere Werte in den unteren Frequenzen erreicht. Die Mikrofonpaare des "Doppel–IRT" weisen durch die vergleichsweise kurzen Kapselabstände von 25 cm einen deutlich anderen Kurvenverlauf auf. Bis circa 600 Hz sind die Signale inkohärent und steigen danach steil an bis zu mittleren Kohärenzwerten. Im besonders relevanten Spektralbereich von 100 bis 200 Hz treten im Vergleich zum "Nieren–Array I" und "Nieren–Array II" hohe Kohärenzwerte auf. Dies weist auf eine signifikant schlechtere Fähigkeit bei der Darstellung von Umhüllung hin. Hier zeigt sich die Kompromisshaftigkeit, welche der kompakte Aufbau des "Doppel-IRT"-Mikrofonsystems mit sich bringt. Es lässt sich festhalten, dass im Vergleich der erprobten Mikrofonsysteme mit Nierencharakteristik vom "Nieren-Array I" der stärkste Umhüllungseindruck zu erwarten ist. Der aus praktischen Gründen erfolgte asymmetrische Aufbau des "Nieren-Array II" bringt einen Anstieg der Kohärenz bei tiefen Frequenzen mit sich. An dieser Stelle kann nicht geklärt werden, ob die zusätzlichen Kanäle des Bottom-Layers vorteilhaft für den Umhüllungseindruck sind, auch wenn die Kohärenzfunktion höhere Werte aufweist. Durch größere Kapselabstände im Bottom-Layer könnte dies kompensiert werden, was jedoch die Praktikabilität stark einschränkt. Das Mikrofonsystem "Doppel-IRT" scheint weniger geeignet, um möglichst umhüllende Atmosphären aufzuzeichnen.

Tabelle 5 listet die Werte des DFI-P der Mikrofonpaare auf. Diese Werte werden vom *Schoeps Image Assistant* zusätzlich zur Kohärenzfunktion angegeben. Der DFI-P zeigt hier eine gute Übereinstimmung mit den Verläufen der Kohärenzkurven. Die Mikrofonpaarungen des "Kugel–Array" und des "Grenzflächen–Array" erzielen die niedrigsten Werte, ebenso wie einzelne Paare des "Hamasaki+Height". Die Mikrofonpaare des "Nieren-Array I & II" sowie weitere Paare des "Hamasaki+Height" weisen leicht erhöhte DFI-P-Werte auf. Das kompakte "Doppel-IRT" erzielt in diesem Vergleich den höchsten Wert. Es ist anzumerken, dass besonders die Werte im unteren Bereich dicht beieinander liegen, was eine Differenzierung anhand des DFI-P erschwert. Die Verläufe der Kohärenzfunktion ermöglichen in diesem Zusammenhang eine detailliertere Betrachtung.

Abschließend lässt sich festhalten, dass die Betrachtung der Diffusfeldkohärenz und des DFI-P mit Hilfe des *Schoeps Image Assistant* darauf schließen lässt, dass die Mikrofonsysteme "Kugel–Array" und "Grenzflächen–Array" den größten Umhüllungseindruck erzeugen können. Das "Nieren–Array I", "Nieren–Array II" und das "Hamasaki + Height" bewegen sich bei diesem Vergleich im Mittelfeld. Das Mikrofonsystem "Doppel–IRT" lässt den geringsten Umhüllungseindruck erwarten.

Mikrofonarray	Mikrofonpaare		
	$Acht/Acht - 2,5 \text{ m Abstand} - 0^{\circ} Öffnungswinkel$		
Hamaali - Haisht	$Acht/Acht - 2,5 \text{ m Abstand} - 180^{\circ} Öffnungswinkel$		
Hamasaki + Height	Niere – $2,5~{\rm m}$ Abstand – 90° Öffnungswinkel		
	Acht/Niere – 1 m Abstand – 45° Öffnungswinkel	0,009	
	Kugel/Kugel - 4 m Abstand		
Kugel-Array Grenzflächen-Array	Kugel/Kugel – 3,5 m Abstand		
· ·	Kugel/Kugel – 3 m Abstand	0,002	
Nieren–Array I	Niere/Niere – 1 m Abstand – 60° Öffnungswinkel		
Nieren–Array II	Niere/Niere – 1 m Abstand – 90° Öffnungswinkel	0,014	
Doppel-IRT	Niere/Niere – 0,25 m Abstand – 90° Öffnungswinkel		

 Tabelle 5:
 Vorhersage des DFI-P mit Hilfe des Schoeps Image Assistant.

8.2.3 Signalanalyse zur Betrachtung der Abbildungseigenschaften realer Atmosphären-Aufnahmen

Nachdem im vorangegangenen Kapitel die räumlichen Abbildungseigenschaften der erprobten Mikrofonsysteme mit Hilfe des Schopes Image Assistant theoretisch betrachtet werden in diesem Kapitel Bedingungen erstellten wurden. die unter realen signaltechnisch Atmosphärenaufnahmen untersucht. Nach der Vorstellung der Analysemethodik, die in MATLAB umgesetzt wurde, werden die Auswertungsergebnisse mit denen der Vorhersage verglichen. Ergänzend werden die Werte des DFI-P gegenübergestellt. Dabei steht die Frage im Fokus, inwieweit sich die Ergebnisse der Vorhersage anhand realer Atmosphärenaufnahmen nachvollziehen lassen, denn letztendlich ist die Dekorrelation der Wiedergabesignale entscheidend für den Umhüllungseindruck. Für die Aufnahmepraxis von Atmosphären ist es somit von großer Bedeutung, die tatsächliche Korrelation der Aufnahmesignale im Vorfeld abschätzen zu können.

8.2.3.1 Methodik zur Kohärenzauswertung in MATLAB

Mit Hilfe eines MATLAB-Skripts werden die Kohärenzfunktionen mehrerer Mikrofonpaare berechnet und grafisch dargestellt. Der komplette Quellcode ist im Anhang einzusehen (siehe A1.2 Kohärenz–Auswertung). Zunächst werden die Audiodaten der Aufnahme in MATLAB geladen. Es besteht die Möglichkeit, einen beliebigen Zeitausschnitt festzulegen. Die Berechnung der Kohärenzfunktion erfolgt über die Methode *mscohere(...)*, die bereits in MATLAB implementiert ist. Die Berechnung der benötigten Leistungsdichtespektren beruht auf dem Verfahren nach Welch (MathWorks Deutschland, 2020b) (siehe Abschnitt 8.2.1.3). Das folgende Listing zeigt den Code zur Berechnung der Kohärenzfunktion eines Mikrofonpaares L-R:

[Cohere_L_R,f]=	<pre>mscohere(au_</pre>	L,au_R,hann	(16384),[]	,16384,fs);
-----------------	-------------------------	-------------	------------	-------------

Listing 1: Funktionsaufruf der magnitude-squared-coherence in MATLAB.

Da die Kohärenzfunktion eine Funktion der Frequenz ist, erfolgt die Umsetzung in den Frequenzbereich über eine Fast-Fourier-Transformation (FFT). Dabei entscheiden die FFT-Parameter, welche beim Funktionsaufruf von *mscohere()* übergeben werden, über die Frequenzauflösung und untere Grenzfrequenz. Die FFT-Size beträgt hier 16384 Samples, was bei einer Audio-Samplerate von 48 kHz zu einer Auflösung von circa 2,93 Hz führt. Durch diese hohe Auflösung kann eine Betrachtung mit hoher Genauigkeit auch bei tiefen Frequenzen erfolgen (Kuttner, 2015). Es kommt eine Hanning-Fensterfunktion zum Einsatz, die in MATLAB standardmäßig mit 50-prozentiger Überlappung arbeitet. Durch die hohe Frequenzauflösung kommt es im Bereich mittlerer und hoher Frequenzen zu einer Anhäufung von Funktionswerten, wodurch der Kurvenverlauf bei einer Darstellung mit logarithmischer Abszisse nur schwer abzulesen ist. Aus diesem Grund werden die Funktionswerte in Frequenzblöcke mit 1/3-Oktav-Bandbreite zusammengefasst und es findet eine arithmetische Mittelwertbildung pro Block statt. Abschließend erfolgt die grafische Ausgabe der Kurvenverläufe der Kohärenzfunktionen für alle Mikrofonpaare. Aus Gründen der Übersichtlichkeit werden die Funktionskurven in Abhängigkeit des Layers des Mikrofonsystems auf verschiedene Graphen aufgeteilt.

8.2.3.2 Methodik zur DFI-P-Berechnung in MATLAB

Die Berechnung des DFI-P eines Mikrofonpaares erfolgt ebenfalls mit Hilfe eines MATLAB-Skripts, wobei Formel (7) genutzt wird. Der komplette Quellcode ist im Anhang einzusehen (siehe A1.3 DFI-P Berechnung). Nach dem Laden der Audiodaten, ist es zunächst erforderlich, die nicht quadrierte Kohärenzfunktion nach Formel (6) zu berechnen. Dazu werden die Autoleistungsspektren $P_{xx}(f)$ und $P_{\nu\nu}(f)$ sowie das Für Kreuzleistungsdichtespektrum $P_{xv}(f)$ berechnet. die Berechnung der Autoleistungsspektren wird die MATLAB- Funktion pwelch() genutzt und für die Berechnung des Kreuzleistungsdichtespektrums die Funktion cpsd(). Beide Funktionen beruhen auf der Berechnungsmethode nach Welch (MathWorks Deutschland, 2020a, 2020c). Die Berechnungen finden im Frequenzbereich statt, weshalb beim Funktionsaufruf FFT-Parameter übergeben werden. Dies zeigt Listing 2:

```
pxx = pwelch(audio_DFI_1_neu,hann(16384),[],16384,fs);
pyy = pwelch(audio_DFI_2_neu,hann(16384),[],16384,fs);
```

```
[pxy,f_welch] = cpsd(audio_DFI_1_neu,audio_DFI_2_neu,hann(16384),[],16384,fs);
```

Listing 2: Funktionsaufrufe zur Berechnung von Leistungsdichtespektren in MATLAB.

Anschließend werden die berechneten Leistungsdichtespektren in Formel (6) eingesetzt, was in MATLAB durch folgenden Code realisiert wird:

my_mnsc = pxy./((pxx.^0.5).*(pyy.^0.5));

Listing 3: Berechnung der nicht-quadrierten Kohärenz in MATLAB.

In einem weiteren Schritt wird nun die Wichtungsfunktion berechnet. Dabei werden nur Frequenzen zwischen 40 Hz und 1,5 kHz betrachtet und es erfolgt eine Gewichtung mit -3 dB pro Oktave. Riekehof-Böhmer et al. (2010) beschreiben nicht den exakten Verlauf der Wichtungsfunktion, weshalb der Verlauf anhand des Funktionsgraphen abgelesen wurde. Daraufhin wurde der Nulldurchgang der Wichtungsfunktion auf 400 Hz festgelegt. In Darstellung lässt die Wichtungsfunktion logarithmischer sich anschließend alsGeradengleichung darstellen. Abbildung 35 zeigt den zugehörigen Funktionsverlauf der in MATLAB erzeugten Wichtungsfunktion. Die Pegel der Gewichtungen werden zur DFI-P in Linearfaktoren Berechnung des umgewandelt. Die Begrenzung des Frequenzbereichs wird realisiert, indem die Wichtungsfaktoren außerhalb des Bereichs von 40 Hz bis 1,5 kHz auf null gesetzt werden.



Abbildung 35: Wichtungsfunktion zur Berechnung des DFI-P in MATLAB.

Außerdem wird zur Berechnung des DFI-P noch die Anzahl der Stützstellen benötigt, an denen die Wichtungsfunktion definiert wird. Dazu werden alle Wichtungsfaktoren ungleich null gezählt, was durch eine Zählschleife mit bedingter Anweisung programmatisch umgesetzt wird. Abschließend können alle Zwischenwerte in Formel (7) eingesetzt werden, um den DFI-P zu berechnen. Die Begrenzung des Frequenzbereichs gelingt über die Wichtungsfaktoren, da ungewünschte Bereiche durch die Multiplikation mit null nicht in die Summation eingehen. Die Programmierung der Zählschleife und der abschließenden Berechnung des DFI-P zeigt das folgende Listing:

```
% Anzahl der FFT-Stützstellen an denen die Wichtungsfunktion definiert ist
n_DFI = 0;
for i = 1:length(f_gains)
    if f_gains(i,3) ~= 0 % ungleich 0
        n_DFI = n_DFI+1;
    end
end
%print DFI value
my_dfi = 1/n_DFI * (sum((real(my_mnsc).*f_gains(:,3)).^2));
fprintf('DFI-Prädiktor: %.3f\n',my_dfi)
```

Listing 4: Zählschleife und abschließende Berechnung des DFI-P in MATLAB.

8.2.4 Vergleich von Vorhersage und Praxis

8.2.4.1 Verlauf der Kohärenzfunktion

Zunächst werden die durchschnittlichen Verläufe der Kohärenzfunktionen verschiedener Mikrofonsysteme betrachtet. Dabei werden die jeweiligen Kohärenzwerte über die Dauer eines Takes von zwei Minuten gemittelt. Anhand dieser Darstellung lässt sich überprüfen, ob sich die Kurvenverläufe der Kohärenzfunktionen, welche mit dem Schoeps Image Assistant vorhergesagt wurden, grundsätzlich in den Kohärenzfunktionen der realen Aufnahmen widerspiegeln. Abbildung 36zeigt die Kurvenverläufe der Atmosphärenaufnahme, welche mit dem "Kugelarray" im Habichtswald entstanden ist. Es fällt auf, dass innerhalb eines Layers die Kohärenzkurven der einzelnen Mikrofonpaare unterschiedlich verlaufen, obwohl die Mikrofonpaare des "Kugel-Array" identisch aufgebaut sind. Die Kohärenzkurven der Mikrofonpaare zwischen den Layern erzeugen insgesamt weniger starke Abweichungen. Alle Kurven steigen unterhalb von 200 Hz mehr oder weniger stark an. In Abbildung 37 sind die gemittelten Kohärenzkurven der Aufnahme dargestellt, die ebenfalls mit dem "Kugel-Array" im Reinhardswald erstellt wurde. Bis circa 200 Hz zeigen alle Mikrofonpaare niedrige Kohärenzwerte, darunter unterscheiden sich die

Kurvenverläufe im Vergleich zur Aufnahme aus dem Habichtswald deutlich. Nach einem Anstieg der Kohärenzwerte ab 200 Hz, fallen die Werte ab etwa 100 Hz wieder ab, um dann ab 40 bis 50 Hz wieder anzusteigen. Dieser Verlauf ist bei den vertikalen Mikrofonpaaren besonders deutlich ausgeprägt. Abbildung 38 veranschaulicht die Kohärenzverläufe der "Kugel-Array"-Aufnahme aus der Kasseler Innenstadt. Es zeigen sich wiederum deutlich abweichende Kurvenverläufe. Die Mikrofonpaare innerhalb des Bottom- und Height-Layers besitzen bis circa 30 Hz durchgehend niedrige Kohärenzwerte, wohingegen die vertikalen Mikrofonpaare ab etwa 150 Hz stark ansteigende Kurvenverläufe aufweisen. Im Vergleich der drei Aufnahmeorte zeigt sich beim "Kugel-Array", dass bis circa 200 Hz durchgehend niedrige Kohärenzwerte für alle Mikrofonpaare erreicht werden. Unterhalb von 200 Hz unterscheiden sich die Kohärenzkurven der unterschiedlichen Aufnahmen deutlich voneinander. In allen drei Aufnahmesituation kam exakt derselbe Mikrofonaufbau zum weshalb die verschiedenen Kohärenzverläufe auf Unterschiede Einsatz, in der Schallfeldzusammensetzung bei der Aufzeichnung zurückzuführen sind. Keine der Darstellungen weist eine hohe Übereinstimmung mit den in Abbildung 33 aufgezeigten Kurvenverläufen des Schoeps Image Assistant auf.



Abbildung 36: Über zwei Minuten gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Habichtswald mit dem "Kugel-Array".



Abbildung 37: Über zwei Minuten gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Reinhardswald mit dem "Kugel-Array".



Frequenz (Hz) Abbildung 38: Über zwei Minuten gemittelte Kohärenzverläufe der Atmosphärenaufnahme in der Kasseler Innenstadt mit dem "Kugel-Array".

Ort 3, Aufnahme 2, "Kugel-Array" – Ausschnitt: Sekunde 0 bis 120

Aufnahmen des "Kugel-Arrays" miteinander verglichen wurden, kann die Betrachtung kürzerer Zeitabschnitte eines Aufnahme-Takes weitere Erkenntnisse über die Auswirkungen von örtlichen Schallfeldänderungen auf den Verlauf der Kohärenzkurven liefern. Abbildung 39 und Abbildung 40 zeigen jeweils die Kohärenzkurven eines fünfsekündigen Ausschnitts der "Kugel-Array"-Aufnahme aus dem Reinhardswald. Es fällt auf, dass die Kurvenverläufe stark variieren. Hervorzuheben sind die erhöhten Kohärenzwerte zwischen 2 und 4 kHz im ersten Ausschnitt. Diese könnten auf Vogelrufe zurückzuführen sein, welche im Zeitabschnitt 00:15 bis 00:20 deutlich wahrnehmbar sind (Hörbeispiel 1). Davon abgeschen lassen sich durch einen subjektiven Hörvergleich mit Ausschnitt zwei (Hörbeispiel 2) keine weiteren konkreten Schallereignisse erkennen, mit denen die Unterschiede der Kohärenzverläufe unmittelbar erklärt werden können. Dieses Beispiel verdeutlicht, dass nicht nur die Schallfelder verschiedener Atmosphärenaufnahmen, welche mit demselben Mikrofonsystem aufgezeichnet werden, abweichende Kohärenzkurven hervorrufen. Es sind ebenso Makroänderungen innerhalb des Schallfeldes einer Atmosphäre für die Varianz der Kohärenzfunktionen verantwortlich.


Abbildung 39: Über fünf Sekunden gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Reinhardswald mit dem "Kugel-Array", Ausschnitt 00:15 – 00:20.



Abbildung 40: Über fünf Sekunden gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Reinhardswald mit dem "Kugel-Array", Ausschnitt 00:30 – 00:35.

Um die mit Hilfe des "Kugel-Array" getätigten Feststellungen zu verifizieren, wird anhand des "Doppel-IRT" ein weiteres Mikrofonsystem exemplarisch betrachtet. Abbildung 41 zeigt die Verläufe der Kohärenzfunktionen der Aufnahme vom Habichtwald, Abbildung 42 die der Aufnahme vom Reinhardswald und Abbildung 43 diejenigen der Atmosphärenaufnahme der Kasseler Innenstadt; jeweils gemittelt über die gesamte Take-Länge von zwei Minuten. Die Kohärenzkurven der verschiedenen Aufnahmen zeigen völlig verschiedene Verläufe. Anhand der Vorhersagen aus Kapitel 8.2.2 wären mehrheitlich höhere Kohärenzwerte beim "Doppel-IRT" zu erwarten gewesen, was sich jedoch nicht bestätigt. Besonders die Kohärenzkurven der Atmosphärenaufnahme des Reinhardswalds (Abbildung 42) weist innerhalb des Bottom- und Height-Layers bis 100 Hz sehr geringe Kohärenzwerte auf. Somit bestätigt sich auch hier die Annahme, dass verschiedene Atmosphären zu deutlich unterschiedlichen Kohärenzkurven führen können, auch einidentisches wenn Mikrofonsystem zum Einsatz kommt. Insgesamt weisen alle Kohärenzkurven deutliche Unterschiede in Abhängigkeit des Aufnahmeortes zueinander auf. Die über die Take-Länge gemittelten Kohärenzverläufe aller erprobten Mikrofonsysteme sind im Anhang zu finden (siehe A2 Kohärenzfunktionen der Mikrofonsysteme).



Ort 1, Aufnahme 6, "Doppel-IRT Kreuz" – Ausschnitt: Sekunde 0 bis 120

Abbildung 41: Über zwei Minuten gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Habichtswald mit dem "Doppel-IRT".



Ort 2, Aufnahme 6, "Doppel-IRT Kreuz" – Ausschnitt: Sekunde 0 bis 120

Abbildung 42: Über zwei Minuten gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Reinhardswald mit dem "Doppel-IRT".



Abbildung 43: Über zwei Minuten gemittelte Kohärenzverläufe der Atmosphärenaufnahme in der Kasseler Innenstadt mit dem "Doppel-IRT".

Die Betrachtung zweier fünfsekündiger Ausschnitte der Atmosphärenaufnahme vom Reinhardswald, welche ebenfalls mit dem "Doppel-IRT" erstellt wurden, zeigt ein ähnliches Bild. In Abbildung 44 sind die Kohärenzverläufe des Ausschnitts 00:45 bis 00:50 dargestellt. Die höchsten Kohärenzwerte aller Mikrofonpaare von maximal 0,6, werden im Frequenzbereich zwischen 100 und 400 Hz erreicht. Bei allen anderen Frequenzen treten deutlich niedrigere Kohärenzwerte auf. Einen Höreindruck dieses Ausschnitts liefert Hörbeispiel 3. 30 Sekunden später, im Ausschnitt 01:20 bis 1:25, weisen die Kohärenzkurven einen signifikant abweichenden Verlauf auf. Dies kann anhand von Abbildung 45 nachvollzogen werden. Hier steigen die Kohärenzkurven der Mikrofonpaare innerhalb des Bottom- und Height-Layers ab circa 1 kHz kontinuierlich an und sind ab 90 Hz nahezu vollständig kohärent. Die Kohärenzfunktionen der Mikrofonpaare, welche sich aus der Kombination von Bottom- und Height-Layer ergeben, steigen erst ab 200 Hz an und weisen einen welligeren Verlauf auf. Das zu diesem Ausschnitt zugehörige Hörbeispiel 4 lässt subjektiv keine konkreten Ursachen dafür erkennen.

Zusammenfassend kann gesagt werden, dass die Kohärenzkurven aller erprobten Mikrofonsysteme in hohem Maße von den akustischen Gegebenheiten am Aufnahmeort abhängen. Dabei führen nicht nur die Atmosphären verschiedener Orte, welche mit einem identischen Mikrofonsystem aufgenommen werden, zu abweichenden Kohärenzverläufen, sondern es ruft auch der dynamische Verlauf innerhalb einer Atmosphäre beträchtliche Änderungen im Kurvenverlauf hervor. Aus diesem Grund kann ein Vergleich der erprobten Mikrofonsysteme untereinander nicht durchgeführt werden, da diese nicht zeitgleich aufgezeichnet werden konnten. Durch die Dynamik der Atmosphären würde ein Vergleich zu falschen Ergebnissen führen. Ferner kann resümiert werden, dass anhand der Diffusfeldkohärenzkurven des *Schoeps Image Assistant* nicht auf die tatsächlichen Kohärenzverläufe einer Atmosphärenaufnahme im Freien geschlossen werden kann.



Abbildung 44: Über fünf Sekunden gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Habichtswald mit dem "Doppel-IRT", Ausschnitt 00:45 – 00:50.



Abbildung 45: Über fünf Sekunden gemittelte Kohärenzverläufe der Atmosphärenaufnahme im Habichtswald mit dem "Doppel-IRT", Ausschnitt 01:20 – 01:25.

Ort 1, Aufnahme 6, "Doppel-IRT" – Ausschnitt: Sekunde 45 bis 50

8.2.4.2 DFI-P-Werte im Vergleich

In einem weiteren Schritt werden nun die zugehörigen DFI-P-Werte der oben näher betrachteten Mikrofonsysteme "Kugel-Array" und "Doppel-IRT" untersucht. Es soll überprüft werden, ob sich die große Varianz der Kohärenzkurven ebenfalls in den DFI-P-Werten niederschlägt. Dazu werden die DFI-P-Werte tabellarisch erfasst. Es werden die mit dem *Schoeps Image Assistant* ermittelten Werte, denjenigen gegenübergestellt, die anhand der Atmosphärenaufnahmen in MATLAB berechnet wurden. Dabei werden sowohl die über die gesamte Take-Laufzeit gemittelten DFI-P-Werte aufgeführt, als auch die der in Abschnitt 8.2.4.1 vorgestellten fünfsekündigen Ausschnitte.

Tabelle 6 enthält die zum "Kugel-Array" gehörigen DFI-P-Werte. Der vom *Schoeps Image Assistant* bestimmte Wert beträgt 0,001. Die in MATLAB berechneten Werte schwanken zwischen 0,013 und 0,296. Es fällt auf, dass innerhalb einer Atmosphärenaufnahme große Schwankungen in Abhängigkeit des Mikrofonpaars auftreten, obwohl diese innerhalb der Layer identisch aufgebaut waren. Weiterhin zeigt sich, dass auch ein Wechsel des Aufnahmeortes deutlich andere DFI-P-Werte mit sich bringt. Die Betrachtung der fünfsekündigen Abschnitte zeigt, dass es auch hier zu Schwankungen kommt. Dabei fallen die Abweichungen von Mikrofonpaar zu Mikrofonpaar innerhalb eines Ausschnitts größer aus als die Abweichungen von Ausschnitt zu Ausschnitt.

Tabelle 7 enthält die entsprechenden DFI-P-Werte für das Mikrofonsystem "Doppel-IRT". Hier zeigt sich insgesamt ein ähnliches Bild zu Tabelle 6. Über alle Mikrofonpaare und Aufnahmeorte treten signifikante Schwankungen auf. Der vom *Schoeps Image Assistant* bestimmte Wert beträgt für das Mikrofonsystem "Doppel-IRT" 0,158. Die in MATLAB berechneten Werte schwanken zwischen 0,074 und 0,573.

Insgesamt kann resümiert werden, dass die DFI-P-Werte die Schwankungen der Kohärenzfunktionen ebenfalls widerspiegeln. Die DFI-P-Werte decken einen großen Zahlenbereich ab und unterscheiden sich deutlich von den Werten, die vom *Schoeps Image Assistant* ermittelt werden. Im Vergleich beider Wertetabellen fällt auf, dass die DFI-P-Werte des "Doppel-IRT" insgesamt höher ausfallen als diejenigen des "Kugel-Arrays".

		DFI-P							
System	Mikrofon- paar	Image Assistant	ø Ort 1	ø Ort 2	ø Ort 3	Ort 2 00:15 - 00:20	Ort 2 00:30 – 00:35		
"Kugel- Array"	L-R	0,001	0,060	0,021	0,050	0,045	0,052		
	L–Ls	0,001	0,064	0,083	0,064	0,102	0,115		
	R–Rs	0,001	0,074	0,086	$0,\!059$	0,111	0,118		
	Ls-Rs	0,001	0,101	0,021	$0,\!056$	0,047	0,064		
	Ltf-Rtf	0,001	0,092	0,013	0,044	0,053	0,062		
	Ltf–Ltr	0,001	0,043	0,043	0,060	0,074	0,064		
	Rtf–Rtr	0,001	0,033	0,041	0,060	0,078	0,067		
	Ltr-Rtr	0,001	0,073	0,014	0,056	0,071	0,073		
	L–Ltf	0,001	0,095	$0,\!054$	0,211	0,096	0,074		
	R–Rtf	0,001	0,097	0,046	$0,\!198$	0,114	0,075		
	Ls–Ltr	0,001	0,121	0,088	0,282	0,115	0,112		
	Rs–Rtr	0,001	0,110	0,073	0,296	0,119	0,101		

Tabelle 6:	DFI-P-Werte	des	Schoeps	Image	Assistant	und	der	Berechnung	für	das
	Mikrofonsystem "Kugel-Array".									

		DFI-P						
System	Mikrofon- paar	Image Assistant	ø Ort 1	ø Ort 2	ø Ort 3	Ort 1 00:45 – 00:50	Ort 1 01:20 – 01:25	
	L-R	$0,\!158$	0,253	0,138	0,381	0,166	0,400	
	L–Ls	$0,\!158$	0,232	$0,\!143$	$0,\!357$	$0,\!130$	0,344	
	R–Rs	$0,\!158$	$0,\!187$	$0,\!140$	$0,\!407$	$0,\!137$	0,314	
	Ls-Rs	$0,\!158$	$0,\!304$	0,074	$0,\!280$	0,134	$0,\!456$	
	Ltf-Rtf	$0,\!158$	0,270	$0,\!151$	$0,\!348$	$0,\!154$	0,378	
Doppel-	Ltf–Ltr	$0,\!158$	$0,\!278$	$0,\!171$	0,300	$0,\!145$	$0,\!455$	
IRT"	Rtf–Rtr	$0,\!158$	$0,\!248$	0,215	0,367	0,169	0,343	
	Ltr-Rtr	$0,\!158$	0,400	0,090	0,227	0,202	0,555	
	L–Ltf	$0,\!158$	$0,\!126$	0,263	$0,\!520$	$0,\!107$	$0,\!158$	
	R–Rtf	$0,\!158$	0,170	$0,\!281$	$0,\!573$	0,186	0,208	
	Ls–Ltr	$0,\!158$	0,228	0,241	0,560	0,192	0,310	
	Rs–Rtr	0,158	0,301	0,213	$0,\!540$	0,330	0,339	

Tabelle 7:DFI-P-Werte des Schoeps Image Assistant und der Berechnung für das
Mikrofonsystem "Doppel-IRT".

8.3 Analyse der interauralen Kohärenz

In diesem Kapitel wird die Kohärenz von Ohrsignalen betrachtet. Wie in Kapitel 5.3 bereits angesprochen wurde, kann der Grad der Ähnlichkeit der Ohrsignale genutzt werden, um Rückschlüsse auf ein zu erwartendes Hörereignis zu ziehen. In der Literatur wird dazu oftmals die IACC zur Auswertung herangezogen. Dabei erfolgt die Berechnung der IACC in der Regel anhand von binauralen Impulsantworten, welche am Aufnahmeort erfasst werden. Da im Rahmen dieser Arbeit Atmosphärenaufnahmen untersucht werden sollen, die synthetisch binauralisiert werden, wird zur Betrachtung der Ohrsignale die Kohärenzfunktion genutzt (siehe 8.2.1.3). Im Gegensatz zur IACC erlaubt die Kohärenzfunktion zudem eine detailliertere Betrachtung, da jede Frequenz separat berechnet wird. Die praktische Umsetzung erfolgt dabei mit der in Abschnitt 8.2.3.1 vorgestellten Methodik. Der Quellcode der genutzten MATLAB-Skripte findet sich im

Anhang (A1.4 Binaural-Analyse Wiedergabesystem 4+5+0; A1.5 Binaural-Analyse Wiedergabesystem 4+7+0; A1.6 Binaural-Analyse zum Vergleich der Wiedergabesysteme).

8.3.1 Funktionstest mit Rauschsignalen und verschiedenen Lautsprecherkonfigurationen

In einem ersten Schritt wird die Funktionalität der Kohärenzanalyse von Ohrsignalen mit Hilfe von dekorrelierten Rauschsignalen überprüft. Zudem wird so untersucht, ob sich die theoretischen Ausführungen zu mehrkanaligen Wiedergabesystemen aus Kapitel 6 anhand einer Kohärenzanalyse nachvollziehen lassen. Mittels eines Signalgenerators aus einer Digital Audio Workstation (DAW) wurden zunächst einminütige, dekorrelierte Signale aus rosa Rauschen erzeugt. Die Rauschsignale wurden im Anschluss mit Hilfe der virtuellen Abhörumgebung in MAX (siehe Kapitel 6.7) über verschiedene Lautsprecherkonfigurationen wiedergegeben und binauralisiert. Die so generierten Ohrsignale wurden daraufhin einer Kohärenzanalyse Abbildung die Kurvenverläufe der unterzogen. 46zeigt Kohärenzfunktionen synthetischer Binauralsignale, bei denen unterschiedlich viele Lautsprecher des (4+5+0)-Wiedergabesystems zum Abspielen der dekorrelierten Rauschsignale genutzt wurden. Es ist zu erkennen, dass die alleinige Wiedergabe über den Center-Kanal zu vollständig kohärenten Ohrsignalen führt. Dies stützt die Überlegungen, den Center-Kanal bei der Atmosphärenaufnahme nicht zu berücksichtigen, da er keinen relevanten Beitrag zur Wahrnehmung von Umhüllung leistet (siehe Kapitel 6.5). Weiterhin fällt auf, dass alle Kohärenzkurven zu tiefen Frequenzen hin stark ansteigen. Dies lässt sich darauf zurückführen, dass die Abmessungen des Kopfes bei tiefer werdenden Frequenzen zunehmend klein gegenüber der Wellenlänge sind und somit immer geringere Unterschiede zwischen den Ohren erzeugt werden. Ferner zeigt sich, dass bei ausschließlicher Nutzung der Lautsprecher L und R höhere Kohärenzwerte auftreten, als bei zusätzlicher Nutzung der Surround- und Height-Lautsprecher. Hier bestätigt sich, dass zusätzliche Lautsprecher – bei der Wiedergabe von dekorrelierten Signalen – zu größeren Unterschieden der Ohrsignale führen und damit für die Darstellung eines umhüllenden Schallfelds dienlich sind. Bezüglich des Mehrwerts der Höhenlautsprecher liefert diese Kohärenzanalyse weniger eindeutige Ergebnisse. Vergleicht man die Kohärenzfunktion, die sich durch die Nutzung der Lautsprecher L, R, Ls und Rs ergibt, mit derjenigen, die zusätzlich vier Höhenlautsprecher nutzt, fällt auf, dass bis circa 300 Hz das Wiedergabesystem mit Höhenlautsprechern zu gleichen oder niedrigeren Kohärenzwerten der Ohrsignale führt. Unterhalb von 300 Hz werden jedoch durch das Wiedergabesystem ohne Höhenlautsprecher niedrigere Kohärenzwerte erzeugt.



Abbildung 46: Kohärenzvergleich synthetischer Binauralsignale, erstellt durch Wiedergabe von dekorreliertem rosa Rauschen über verschiedene Lautsprecher des (4+5+0)-Wiedergabesystems; gemittelt über 60 Sekunden.

Abbildung 47 stellt die Ergebnisse der Kohärenzanalyse bei Nutzung verschiedener Lautsprecherkombinationen des (4+7+0)-Wiedergabesystems zur Binauralsynthese dar. Hier zeigt sich insgesamt ein ähnliches Bild wie beim (4+5+0)-Wiedergabesystem. Innerhalb eines Layers führen zusätzliche Lautsprecher in weiten Frequenzbereichen meist zu niedrigeren Kohärenzwerten. So liegen die Werte der Kurve L–R über denen der Kurve L– R–Ls–Rs. Kommen zusätzlich die Lautsprecher Lss und Rss zum Einsatz (violette Kurve), werden in vielen Frequenzbereichen noch niedrigere Kohärenzwerte erreicht, in einigen Bereichen treten nun jedoch auch höhere Kohärenzwerte auf. In Kombination mit den Lautsprechern des Height-Layers ergeben sich leicht differierende Kurvenverläufe, ein deutlicher Mehrwert durch zusätzliche Höhen-Lautsprecher lässt sich jedoch nicht erkennen.



Abbildung 47: Kohärenzvergleich synthetischer Binauralsignale, erstellt durch Wiedergabe von dekorreliertem rosa Rauschen über verschiedene Lautsprecher des (4+7+0)-Wiedergabesystems; gemittelt über 60 Sekunden.

Abschließend werden die Verläufe der Kohärenzfunktionen beider Wiedergabesysteme in einer Grafik dargestellt. Mit Ausnahme des Center- und LFE-Kanals werden für diesen Vergleich alle virtuellen Lautsprecher zur Wiedergabe dekorrelierter Rauschsignale genutzt. Zusätzlich zu rosa Rauschen wurde auch weißes Rauschen bei der Binauralsynthese genutzt. Die Kurvenverläufe werden in Abbildung 48 wiedergegeben. Es zeigt sich, dass alle Kurven dicht beieinander liegen. Die Unterschiede, die zwischen rosa und weißem Rauschen auftreten, sind minimal. Mit der hier erfolgten Kohärenzanalyse ist die vermutete Überlegenheit des Wiedergabesystems (4+7+0) in Bezug auf die Erzeugung möglichst inkohärenter Ohrsignale somit nicht nachweisbar. Alle Kohärenzanalysen wurden zudem mit unterschiedlich langen Testsignalen und zufälliger Verteilung der Rauschsignale auf die virtuellen Lautsprecher durchgeführt, um systematische Fehler durch die erzeugten Testsignale auszuschließen. Abschließend kann festgehalten werden, dass die Kohärenzanalyse von Ohrsignalen grundsätzlich geeignet ist, um die Unterschiede der Quellsignale, welche bei der Binauralsynthese zum Einsatz kommen, darzustellen. Allerdings fällt es schwer, die Ergebnisse zu interpretieren, da die Kurvenverläufe dicht beieinander verlaufen und stark schwanken. Die Eignung des Verfahrens für den Vergleich verschiedener Mikrofonsysteme für Atmosphärenaufnahmen ist daher nur eingeschränkt gegeben.



Abbildung 48: Kohärenzvergleich synthetischer Binauralsignale, erstellt durch Wiedergabe von dekorreliertem Rauschen über die Wiedergabesysteme (4+5+0) und (4+7+0); gemittelt über 60 Sekunden.

8.3.2 Vergleich zwischen Kunstkopfaufnahmen und binauralisierten Mikrofonarrays

Exemplarisch herfür sollen die beiden Mikrofonsysteme "Kugel-Array" und "Doppel-IRT" untersucht werden. Dazu wurden die Aufnahmen der Mikrofonarrays mit Hilfe der virtuellen Abhörumgebung in MAX (siehe Kapitel 6.7) binauralisiert. Dies ermöglicht einen direkten Vergleich von Mikrofonarray und synchron aufgezeichnetem Kunstkopf. Idealerweise ist es möglich, durch eine Kohärenzanalyse der Ohrsignale festzustellen, welches Mikrofonarray bei der Wiedergabe möglichst unterschiedliche Ohrsignale erzeugt und so den größten Umhüllungseindruck hervorrufen kann. Dabei kann die synchrone Kunstkopfaufnahme als Referenz betrachtet werden, da sie den Höreindruck der Atmosphäre vor Ort unmittelbar überträgt. Beim vorangegangenen Test wurde bereits deutlich, dass selbst bei idealisierten Wiedergabesignalen die Ohrsignale bei tiefen Frequenzen zunehmend kohärent sind. Daher ist auch bei realen Aufnahmen ein ähnlicher Kohärenzverlauf zu erwarten. Dies bestätigen Abbildung 49 und Abbildung 50. Dargestellt werden die Kohärenzfunktionen der Kunstkopfaufnahme (durchgezogene Linie) und die des binauralisierten Mikrofonsystems (gestrichelte Linie), welches synchron dazu aufgezeichnet wurde. In Abbildung 49 sind die Kohärenzverläufe des "Kugel-Arrays" für alle drei Aufnahmeorte dargestellt, jeweils gemittelt über die Take-Länge von zwei Minuten. In Abbildung 50 sind die entsprechenden Kurven für das Mikrofonsystem "Doppel-IRT" illustriert. Insgesamt fällt auf, dass alle Kurven relativ dicht beieinander liegen. Die Auswirkungen der verschiedenen Aufnahmeorte scheinen sich in den Kohärenzverläufen der Ohrsignale wesentlich schwächer niederzuschlagen, als in den Kohärenzverläufen der Mikrofonpaare eines Arrays (siehe Kapitel 8.2.4.1). Auch die Kohärenzverläufe der binauralisierten Mikrofonsysteme verlaufen, mit Ausnahme einiger Spitzen, recht dicht zu denen der Kunstkopfaufnahmen. Im Bereich zwischen 50 und 400 Hz weichen die Kohärenzkurven des binauralisierten "Doppel-IRT" stärker von den Kohärenzkurven der synchronen Kunstkopfaufnahmen ab, als diejenigen des binauralisierten "Kugel-Arrays". Dies kann als Indiz für die Überlegenheit des "Kugel-Arrays" in Bezug auf räumliche Abbildungseigenschaften interpretiert werden. Allerdings sind die Abweichungen allgemein als gering einzustufen, weshalb das Präferieren eines Mikrofonsystems anhand dieses Vergleichs der Kohärenzverläufe von Ohrsignalen wenig sinnvoll erscheint. Die Vergleiche der Kohärenzfunktionen aller erprobten Mikrofonsysteme und der zugehörigen Kunstkopfaufnahmen sind im Anhang $\mathbf{z}\mathbf{u}$ finden (A3)Kohärenzfunktionen der Kunstkopfaufnahmen und der binauralisierten Mikrofonsysteme im Vergleich).

Beim Betrachten kürzerer Zeitabschnitte fällt auf, dass auch hier die Schwankungen der Kohärenzkurven geringer ausfallen. Allerdings kommt es gelegentlich zu starken Einbrüchen im Kohärenzverlauf der Kunstkopfaufnahmen. Ein Beispiel hierzu ist in Abbildung 51 dargestellt. In dem kurzen dargestellten Zeitfenster kommt es aufgrund von Windstörungen am linken Kanal des Kunstkopfes zu geringen Kohärenzwerten im unteren Frequenzbereich. Dies ist anhand von Hörbeispiel 5 nachzuvollziehen. Zehn Sekunden später ist die Störung vorüber und die Kohärenzkurve erreicht wieder hohe Werte im tieffrequenten Bereich, was in Abbildung 52 dargestellt ist (Hörbeispiel 6). Da es am Aufnahmeort 1 (Habichtswald) generell sehr windig war, sind die geringeren Kohärenzwerte der Kunstkopfaufnahme hier vermutlich auf häufigere Windstörungen zurückzuführen. Zudem gestaltete sich das Anbringen eines wirksamen Windschutzes am Kunstkopf als schwierig.



Abbildung 49: Kohärenzverläufe von Kunstkopf und binauralisiertem "Kugel-Array", gemittelt über zwei Minuten.



Abbildung 50: Kohärenzverläufe von Kunstkopf und binauralisiertem "Doppel-IRT", gemittelt über zwei Minuten.



Abbildung 51: Kohärenzverläufe von Kunstkopf und binauralisiertem "Kugel-Array", gemittelt über fünf Sekunden, Ausschnitt 00:30 – 00:35.



Abbildung 52: Kohärenzverläufe von Kunstkopf und binauralisiertem "Kugel-Array", gemittelt über fünf Sekunden, Ausschnitt 00:45 – 00:50.

8.4 Diskussion der Analyseergebnisse

Durch die Untersuchung der Kohärenzfunktionen verschiedener Atmosphärenaufnahmen und Mikrofonsysteme konnte gezeigt werden, dass der Verlauf der theoretischen Diffusfeldkohärenz nicht anhand realer Atmosphärenaufnahmen nachempfunden werden kann. Ferner wurde festgestellt, dass die Kohärenzverläufe der aufgezeichneten Signale stark variieren und dynamisch sind. Dabei führen sowohl unterschiedliche Aufnahmeorte als auch natürliche Schwankungen innerhalb einer Atmosphäre zu deutlichen Änderungen im Kurvenverlauf der Kohärenzfunktionen. Hieraus kann geschlussfolgert werden, dass verschiedene Mikrofonsysteme synchron aufgezeichnet werden müssen, wenn ein Vergleich anhand der Kohärenzfunktion erfolgen soll. Die Vorhersage der Diffusfeldkohärenz, wie sie durch den Schoeps Image Assistant erfolgt, beruht auf idealisierten, theoretischen Annahmen. Bei der Berechnung wird von einem idealen diffusen Schallfeld ausgegangen (Elko, 2001). Ein ideales Diffusfeld ist homogen und isotrop, was bedeutet, dass die Schallenergie aus allen Richtungen im Mittel gleich ist (SCHOEPS Mikrofone, o. J.-a). Diese Betrachtung mag für den Nachhall in einem Konzertsaal eine plausible Annäherung darstellen, für die Praxis der Atmosphärenaufnahme scheint sie jedoch weniger hilfreich. Die Beschaffenheit des Schallfeldes einer Atmosphäre im Freien ist sehr komplex und vielseitig. Zwar kommt es auch hier durch Reflexionen an Böden, Hauswänden, Bergen, Bäumen etc. zu diffusen Schallfeldanteilen, diese erfüllen jedoch nicht die Anforderungen eines idealen diffusen Schallfelds. In der Praxis entsteht ein gemischtes Schallfeld mit direkten und diffusen Anteilen. Die Diffusität einer Atmosphäre ergibt sich vor allem daraus, dass Schallquellen nicht lokalisiert werden und der Klang von "überall" herzukommen scheint. Die genaue Unterscheidung zwischen Direktschall und diffusem Nachhall, wie sie in geschlossenen Räumen möglich ist, lässt sich bei Atmosphären im Freien nur schwerlich umsetzen.

Darüber hinaus gilt es zu berücksichtigen, dass bei der Vorhersage der Diffusfeldkohärenz von idealen Richtcharakteristiken der Mikrofone ausgegangen wird. Dabei wird angenommen, dass die Richtcharakteristik über den gesamten Übertragungsbereich konstant ist. In der Praxis zeigt sich, dass die Richtwirkung von Druckgradientenempfänger in der Regel nicht frequenzstabil ist. Druckempfänger bilden zu hohen Frequenzen hin eine deutliche Richtwirkung aus (Wuttke, 2000). Die Ausprägung dieser Effekte ist stark abhängig von Mikrofontyp und Hersteller. Daher hat in der Praxis auch die Wahl der eingesetzten Schallwandler Einfluss auf den Verlauf der Kohärenzfunktionen.

Nach den obigen Ausführungen ist es nachvollziehbar, dass die errechneten Werte des DFI-P ebenfalls variieren und von den theoretischen Werten abweichen. Der DFI-P leitet sich aus der Kohärenzfunktion ab, weshalb auch hier idealisierte Annahmen zu Grunde liegen. Der Vergleich der DFI-P-Werte zweier Mikrofonsysteme hat gezeigt, dass sich insgesamt höhere Werte ergeben, wenn der vom *Schoeps Image Assistant* ermittelten Wert ebenfalls höher liegt. Dies lässt auf einen sinnvollen Zusammenhang zwischen Vorhersage und Berechnung realer DFI-P-Werte eines Mikrofonsystems schließen. Demnach eignen sich Mikrofonsysteme mit kleinen DFI-P-Werten besser, um Atmosphären mit breiter räumlicher Abbildung aufzuzeichnen. Letztlich stimmen dieser Überlegungen mit der allgemein bekannten Gesetzmäßigkeit überein, dass größere Mikrofonabstände und Öffnungswinkel zu einer breiteren räumlichen Abbildung führen. Im Zusammenhang mit Atmosphärenaufnahmen kann der DFI-P eine Orientierungshilfe bei der Planung von Mikrofonsystemen bieten, vermittelt dem erfahrenen Tonschaffenden jedoch nur geringen Mehrwert.

Die Kohärenzanalyse Ohrsignalen dass sich verschiedene von hat gezeigt, Lautsprecherkonfigurationen bei der Wiedergabe von dekorrelierten Signalen inverschiedenen Verläufen der Kohärenzfunktionen widerspiegeln. Die Methode scheint jedoch nicht dazu geeignet zu sein, feine Abstufungen bezüglich des Räumlichkeitseindrucks eindeutig darzustellen. So führten zwar zusätzliche Elevations-Lautsprecher zu abweichenden Kohärenzverläufen, jedoch kann anhand derer nicht eindeutig nachvollzogen werden, dass sich zusätzliche Höhenlautsprecher positiv auf den Räumlichkeits- und Umhüllungseindruck auswirken, wie in Kapitel 6.4 gezeigt wurde. Auffällig ist, dass die Kohärenzfunktionen der Ohrsignale im unteren Frequenzbereich durchweg hohe Werte

erzielen und somit keine Aussage zur Auswirkung dieses spektralen Bereichs auf den Räumlichkeitseindruck möglich ist. Diese Feststellung stimmt mit Aussagen von Morimoto und Maekawa (1988) überein, die zu dem Schluss kamen, dass der IACC den Räumlichkeitseindruck nicht eindeutig bestimmt und sich tiefe Frequenzkomponenten unabhängig vom IACC auf den Räumlichkeitseindruck auswirken. Auch der Vergleich von binauralisierten Mikrofonsystemen und Kunstkopfaufnahmen mit Hilfe einer Kohärenzanalyse lässt lediglich Tendenzen erkennen, kann aber keine eindeutigen Ergebnisse liefern.

9 Fazit und Ausblick

Sound recording is, above all, the skill of building a bridge between the technical and the artistic. Even though a great deal of understanding can also be gained from the scientific analysis of the microphone array design process, the training of the ear to interpret the different parameters of a sound image is probably the most important part of what is usually a long apprenticeship (Williams, 2004, S. XVI).

Im Rahmen dieser Thesis wurde das Thema Atmosphären in immersiven Audioproduktionen aus verschiedenen Blickwinkeln betrachtet. Hierbei wurden sowohl gestalterische als auch technische Überlegungen angestellt, um einen möglichst vollständigen Blick auf das Themenfeld zu ermöglichen. Dabei galt es den praktischen Nutzen und die Anwendbarkeit der Erkenntnisse im Rahmen einer audiovisuellen Medienproduktion zu berücksichtigen.

Die vorgestellte Abgrenzung und Definition der Begrifflichkeiten im Zusammenhang mit Atmosphären ermöglicht dabei eine präzise Kommunikation. Mit Hilfe verschiedener Kategorisierungsmodelle wird dem Tonschaffenden eine Möglichkeit geboten, sich systematisch mit Atmosphären auseinanderzusetzen. Dies ermöglicht es, bedeutsame Merkmale für die Planung und Durchführung von Atmosphärenaufnahmen und den Sounddesign-Prozess zu erkennen und zu definieren. Weiterhin wurde aufgezeigt, dass Atmosphären eine fundamentale Rolle im Sounddesign-Prozess spielen und dabei relevante Funktionen übernehmen. In diesem Zusammenhang wurden insbesondere die akustische Darstellung eines Ortes, die Etablierung eines auditiven Hintergrunds für eine komplexe Tonszene und die Wahrung von Kontinuität als wichtigste Funktionen identifiziert. Zudem sind Atmosphären die Schlüsselkomponente, um das Gefühl des akustischen "Vor Ort seins" zu vermitteln. Dabei wurde deutlich, dass Atmosphären idealerweise einen hintergründigen Charakter besitzen sollten, um dem Sounddesigner größtmögliche Kontrolle über die Ausgestaltung des finalen Soundscapes zu ermöglichen. Darüber hinaus besteht ein großer Bedarf an verschiedensten Atmosphären, mit denen sich vielfältige akustische Orte in einer Medienproduktion abwechslungsreich darstellen lassen. Es wurde dargelegt, dassAtmosphären in der Lage sind das immersive Potenzial einer Audioproduktion zu steigern. Hierbei konnte gezeigt werden, dass die Wahrnehmung von Umhüllung das bedeutendste

perzeptive Merkmal darstellt. In einem weiteren Schritt wurden die psychoakustischen Merkmale betrachtet, die bei der Wahrnehmung von Atmosphären relevant sind. Es wurde insbesondere dargestellt auf welchen Grundlagen die Wahrnehmung von Umhüllung beruht. Die detaillierte Betrachtung ausgewählter Wiedergabesysteme konnte aufzeigen, dass sich mit den aktuell verbreiteten Systemen immersive Audioproduktionen in hoher Qualität wiedergegeben lassen und es grundsätzlich möglich ist, den hohen Grad an Umhüllung darzustellen, der nötig ist, um die Anforderungen des Sounddesigns zu erfüllen. Diesbezüglich wurde die Anforderung an die Aufnahmetechnik abgeleitet, dass die Kanäle einer Atmosphärenaufnahme möglichst dekorreliert sein müssen, um einen hohen Grad an Umhüllung bei der Wiedergabe erreichen zu können. Im Hinblick auf die praktische Durchführung von Atmosphärenaufnahmen wurden zunächst ausgewählte technische Aspekte der Mikrofontechnik vorgestellt, die insbesondere im Kontext mobiler Tonaufzeichnung im Freien relevant sind. Ferner wurden spezifische Mikrofonsysteme zur Atmosphärenaufnahme präsentiert. Nach der Konzeption verschiedener Mikrofonsysteme wurden diese im Rahmen eines Feldversuchs praktisch erprobt. Im Anschluss erfolgte eine signaltechnische Analyse der erstellten Aufnahmen. Zunächst wurde überprüft, ob es möglich ist, anhand der theoretischen Diffusfeldkohärenz, wie sie vom Schoeps Image Assistant berechnet wird. auf die tatsächliche Kohärenz von diffusen Atmosphärenaufnahmen zu schließen. Dabei stellte sich heraus, dass die Schallfelder aller untersuchten Atmosphären zu sehr von einem idealen Diffusfeld abweichen und sich der Verlauf der theoretischen Diffusfeldkohärenz nicht anhand der berechneten Kohärenzfunktionen der Atmosphären nachvollziehen lässt. In einem weiteren Schritt wurde eine virtuelle Abhörumgebung in MAX entwickelt, um darin über eine Binauralsynthese Kunstkopfsignale zu generieren, welche die Wiedergabe ausgewählter Signale über die Lautsprecherkonfigurationen (4+5+0) und (4+7+0) simulieren. Über eine anschließende Kohärenzanalyse der Binauralsignale wurde versucht, Rückschlüsse der simulierten Wiedergabesituation auf die Ohrsignale zu ziehen. In einem Test mit dekorrelierten Rauschsignalen zeigte sich, dass diese Methode grundsätzlich in der Lage ist, Unterschiede im Verlauf der Kohärenzfunktion darzustellen. Allerdings wurde auch deutlich, dass der Vergleich der Kohärenzkurven nicht eindeutig interpretiert werden kann, wenn die Unterschiede der simulierten Wiedergabesituation gering ausfallen. Dies bestätigte sich beim Vergleich zwischen den binauralisierten Mikrofonsystemen und der synchron erstellten

Kunstkopfaufnahmen. Die Unterschiede der Mikrofonsysteme wurden nur ansatzweise deutlich.

Insgesamt kann festgehalten werden, dass die Aufnahme von Atmosphären für immersive Audioproduktionen ein komplexes aber lohnenswertes Vorhaben darstellt. Neben der Beschaffenheit der aufzunehmenden Atmosphären, den Merkmalen des Wiedergabesystems und den klanglichen und räumlichen Abbildungseigenschaften des eingesetzten Mikrofonsystems müssen vielfältige praktische Aspekte berücksichtigt werden. Die Kohärenzanalyse kann dazu beitragen, genauere Erkenntnisse über die Beschaffenheit der aufgezeichneten Signale zu erlangen. Der Tonschaffende muss mit technisch-gestalterischem Feingefühl und Durchhaltevermögen den Umweltbedingungen standhalten und praktische (Hör-)Erfahrungen sammeln, um Atmosphären mit dem größtmöglichen immersiven Potenzial aufnehmen zu können.

Nichtsdestotrotz bietet diese Herangehensweise die Möglichkeit, das Thema der immersiven Atmosphären weiter zu erforschen. Aus den Ergebnissen dieser Arbeit lassen sich Ansätze für weitere Untersuchungen ableiten. So könnte der Vergleich der erprobten Mikrofonsysteme mit dem parallel aufgezeichneten Ambisonics-Mikrofon weitere Erkenntnisse über die Abbildungsqualitäten der verschiedenen Mikrofonsysteme und ihre Eignung bezüglich der Aufnahme immersiver Atmosphären liefern. Des Weiteren wurde bei der Betrachtung der Wiedergabesysteme auf die Besonderheiten der Lautsprecherkonfiguration im Kino hingewiesen. Da sich 3D-Audio im Kino bereits fest etabliert hat, wäre es ein lohnenswertes Unterfangen, die Wiedergabe von immersiven Atmosphären im Kino systematisch zu untersuchen und Verbesserungen zur Wiedergabe über Lautsprechergruppen zu erarbeiten. Ferner hat sich im Rahmen der Signalanalyse dieser Arbeit gezeigt, dass die Kohärenzanalyse in der Lage ist, die Ahnlichkeit von Mikrofonsignalen frequenzbezogen darzustellen. Durch Echtzeitverarbeitung und grafische Aufbereitung könnte die Kohärenzanalyse zu einem Anzeigeinstrument weiterentwickelt werden und in der Postproduktion in Form eines Plug-Ins visuelle Unterstützung bei Sounddesign und Mischung immersiver Audioproduktionen bieten ("Immerse-O-Meter"). Durch das Aufkommen von 3D-Audio in Film- und Musik-Streaming-Diensten (Netflix, Tidal, Deezer etc.) sowie der technischen Implementierung von Spatial Audio Technologien

in Endgeräten, die der breiten Masse zur Verfügung stehen - wie beispielsweise den Apple AirPods Pro - wird der Verfügbarkeit von immersiven Atmosphärenaufnahmen weiter Bedeutung zuwachsen. Zukünftige weitere Forschungen im Bereich der immersiven Atmosphären werden zeigen, wie die Wiedergabe von Atmosphären über die neuen Systeme und Formate gelingen kann und inwieweit sich daraus weitere Anforderungen für die Mikrofonaufnahmetechnik ergeben.

Literaturverzeichnis

- Acoustical Society of America. (2020). Interaural cross correlation. ANSI/ASA S1.1 & S3.20 Standard Acoustical & Bioacoustical Terminology Database. https://asastandards.org/Terms/interaural-cross-correlation/
- Agrawal, S., Simon, A., Bech, S., Bærentsen, K., & Forchhammer, S. (2019, Oktober). Defining Immersion: Literature Review and Implications for Research on Immersive Audiovisual Experiences. *Audio Engineering Society Convention* 147. http://www.aes.org/e-lib/browse.cfm?elib=20648
- Andriessens, F., & Hoffmeister, M. (2017, September). A New 3D Microphone Array for Cinema, VR and Games – A Comparison. 4thInternational Conference on Spatial Audio. http://www.aes.org/e-lib/browse.cfm?elib=12272
- Audio Engineering Society. (2015). AES69-2015 Standard for file exchange—Spatial acoustic data file format.
- Auro Technologies. (2015). AURO-3D® HOME THEATER SETUP Installation Guidelines.
- Ballou, G., & Safari, an O. M. C. (2012). Electroacoustic Devices: Microphones and Loudspeakers.
- Bendat, J. S., & Piersol, A. G. (1980). Engineering applications of correlation and spectral analysis. Wiley.
- Berg, J. (2009, Mai). The Contrasting and Conflicting Definitions of Envelopment. Audio Engineering Society Convention 126. http://www.aes.org/elib/browse.cfm?elib=15004

- Berg, J., & Nyberg, D. (2008, Mai). Listener Envelopment What Has Been Done and What Future Research Is Needed? Audio Engineering Society Convention 124. http://www.aes.org/e-lib/browse.cfm?elib=14509
- Bernschütz, B. (2013). A Spherical Far Field HRIR/HRTF Compilation of the Neumann KU 100.
- Berzborn, M., Bomhardt, R., Klein, J., Richter, J.-G., & Vorländer, M. (2017, März 6). The ITA-Toolbox: An Open Source MATLAB Toolbox for Acoustic Measurements and Signal Processing. http://publications.rwth-aachen.de/record/687308
- Blauert, J. (1974). Räumliches Hören Komplettausgabe: Grundwerk, 1. Aufl. 1974 mit 1. Nachschrift 1985 (Ergebnisse und Trends seit 1972) und 2. Nachschrift 1997 (Ergebnisse und Trends seit 1982). S. Hirzel Verlag.
- Blauert, J., & Braasch, J. (2008). Räumliches Hören. In S. Weinzierl (Hrsg.), Handbuch der Audiotechnik (S. 87–121). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-34301-1 3
- Boré, G., & Peus, S. (1999). Microphones for Studio and Home-Recording Applications— Operation Principles and Type Examples. Georg Neumann GmbH.
- Borenius, J. (1985, Oktober). Perceptibility of Direction and Time Delay Errors in Subwoofer Reproduction. Audio Engineering Society Convention 79. http://www.aes.org/e-lib/browse.cfm?elib=11465
- Bregman, A. S. (1990). Auditory scene analysis: The perceptual organization of sound. MIT Press.
- Butzmann, F., & Martin, J. (2012). *Filmgeräusch: Wahrnehmungsfelder eines Mediums* (Erstausgabe). Wolke-Verl.

- Camerer, F. (2000, November). Mehrkanalton bei Aussenaufnahmen—Eine systematische Betrachtung. 21. Tonmeistertagung.
- Camerer, F. (2020, Oktober). Designing a 9-channel location sound microphone from scratch. Audio Engineering Society Convention 149. http://www.aes.org/elib/browse.cfm?elib=20908
- Carpentier, T. (2018). A New Implementation Of Spat In Max. https://doi.org/10.5281/ZENODO.1422552
- Carpentier, T., Noisternig, M., & Warusfel, O. (2015). Twenty Years of Ircam Spat: Looking Back, Looking Forward. 41st International Computer Music Conference (ICMC), 270–277. https://hal.archives-ouvertes.fr/hal-01247594
- Chernyak, R. I., & Dubrowsky, N. A. (1968). Pattern of the noise images and the binaural summation of loudness for the different interaural correlation of noise. Proc. 6th Int. Congr. Acoust., 3. https://ci.nii.ac.jp/naid/20000152886/en/
- Cherry, E. C. (1953). Some Experiments on the Recognition of Speech, with One and with Two Ears. The Journal of the Acoustical Society of America, 25(5), 975–979. https://doi.org/10.1121/1.1907229
- Chion, M., & Gorbman, C. (2019). *Audio-vision: Sound on screen* (Second edition). Columbia University Press.
- Cousins, M. P., Fazi, F. M., Bleeck, S., & Melchior, F. (2015, Oktober). Subjective Diffuseness in Layer-Based Loudspeaker Systems with Height. Audio Engineering Society Convention 139. http://www.aes.org/e-lib/browse.cfm?elib=17983
- Cremer, L., & Müller, H. A. (1976). Die wissenschaftlichen Grundlagen der Raumakustik (2., völlig neubearb. Aufl). Hirzel.

Damaske, P. (1967). Subjective Investigation of Sound Fields. Acustica, 19(4), 199–213.

DeLouise, A., & Ottenritter, C. (2020). Nonfiction sound and story for film and video: A practical guide for filmmakers and digital content creators.

Dickreiter, M. (1995). Mikrofon-Aufnahmetechnik (2., neu bearb. und erw. Aufl). Hirzel.

- Dickreiter, M., & Goeres-Petri, J. (2013). Schallwahrnehmung. In Handbuch der Tonstudiotechnik (S. 115–135). DE GRUYTER. https://doi.org/10.1515/9783110316506.115
- Djuric, P. M., & Kay, S. M. (2017). Spectrum Estimation and Modeling. In V. K. Madisetti (Hrsg.), The Digital Signal Processing Handbook. CRC Press.

Dolby Laboratories. (2018). Dolby Atmos (R) Home Theater Installation Guidelines.

- Elko, G. W. (2001). Spatial Coherence Functions for Differential Microphones in Isotropic Noise Fields. In M. Brandstein & D. Ward (Hrsg.), Microphone Arrays: Signal Processing Techniques and Applications (S. 61–85). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-04619-7 4
- Ellermeier, W., & Hellbrück, J. (2008). Hören Psychoakustik Audiologie. In S. Weinzierl (Hrsg.), *Handbuch der Audiotechnik* (S. 41–85). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-34301-1 2
- Fisher, J. P. (2005). *Instant surround sound*. CMP Books; Distributed in the U.S. by Publishers Group West.
- Flückiger, B. (2007). Sound design: Die virtuelle Klangwelt des Films (3. Auflage). Schüren.
- Friesecke, A. (2007). Metering: Studio-Anzeigen lesen und verstehen; vom VU-Meter bis zum Analyser; Technik, Formate, Normen (2., überarb. Aufl). PPVMedien.
- Friesecke, A. (2014). Die Audio-Enzyklopädie: Ein Nachschlagewerk f
 ür Tontechniker (2. Auflage). De Gruyter Saur.

- Furuya, H., Fujimoto, K., Nakano, Y., & Wakuda, A. (2005). The influence of total and directional energy of late sound on listener envelopment. Acoustical Science and Technology, 26, 208–211. https://doi.org/10.1250/ast.26.208
- Furuya, H., Fujimoto, K., Young Ji, C., & Higa, N. (2001). Arrival direction of late sound and listener envelopment. *Applied Acoustics*, 62(2), 125–136. https://doi.org/10.1016/S0003-682X(00)00052-9
- Georg Neumann GmbH Berlin. (2020). *Recording Tools*. App Store. https://apps.apple.com/us/app/recording-tools/id576702914
- Gernemann-Paulsen, A. (o. J.). Kohärenz und Korrelation in der Tonstudiotechnik. http://www.uni-koeln.de/phil-fak/muwi/ag/tec/kohkor.pdf
- Görne, T. (2015). Tontechnik. Hanser.
- Görne, T. (2017). Sounddesign: Klang, Wahrnehmung, Emotion. Hanser.
- Griesinger, D. (2001, Juni). The Psychoacoustics of Listening Area, Depth, and Envelopment in Surround Recordings, and their relationship to Microphone Technique. Audio Engineering Society Conference: 19th International Conference: Surround Sound - Techniques, Technology, and Perception. http://www.aes.org/elib/browse.cfm?elib=10076
- Griesinger, D. (2018, Oktober). Reproducing Low Frequency Spaciousness and Envelopment in Listening Rooms. Audio Engineering Society Convention 145. http://www.aes.org/e-lib/browse.cfm?elib=19789
- Hagelüken, A. (2009). Sag's doch mal nonverbal lautes Nachdenken über bedeutende Klang- und Geräuschereignisse im Radio. In Hörspiel ist schön! Beiträge aus sechs Jahren Hörspielsymposion am Nordkolleg Rendsburg (1. Aufl). Books on Demand.

- Hamasaki, K., & Hiyama, K. (2003, Juni). Reproducing Spatial Impression With Multichannel Audio. Audio Engineering Society Conference: 24th International Conference: Multichannel Audio, The New Reality. http://www.aes.org/elib/browse.cfm?elib=13880
- Hellbrück, J., & Ellermeier, W. (2004). Hören: Physiologie, Psychologie und Pathologie (2., aktualisierte und erw. Aufl). Hogrefe, Verl. für Psychologie.
- Hirst, J., Davies, W. J., & Philipson, P. (2006, Mai). The Adaptation of Concert Hall Measures of Spatial Impression to Reproduced Sound. Audio Engineering Society Convention 120. http://www.aes.org/e-lib/browse.cfm?elib=13536
- Hiyama, K., Komiyama, S., & Hamasaki, K. (2002, Oktober). The Minimum Number of Loudspeakers and its Arrangement for Reproducing the Spatial Impression of Diffuse Sound Field. Audio Engineering Society Convention 113. http://www.aes.org/elib/browse.cfm?elib=11272
- Holman, T. (2008). Surround sound: Up and running (2nd ed). Elsevier/Focal Press.
- Holman, T. (2010). Sound for film and television (3rd ed). Focal Press.
- Howard, D. M., & Angus, J. (2009). Acoustics and psychoacoustics (4th ed). Focal Press.
- Ircam. (o. J.). Spat / Ircam Forum. Abgerufen 2. Dezember 2020, von https://forum.ircam.fr/projects/detail/spat/
- Ircam. (2018). Spatialisateur User Manual.
- Ircam. (2020). Spatialisateur Reference Pages.
- ITU. (2012). RECOMMENDATION ITU-R BS.775-3: Multichannel stereophonic sound system with and without accompanying picture.
- ITU. (2018). RECOMMENDATION ITU-R BS.2051-2: Advanced sound system for programme production.

- Kamp, W. (2007). *AV-Mediengestaltung: Grundwissen* (2. Aufl., 1. Dr). Verl. Europa-Lehrmittel.
- Kim, S. (2018). Height Channels. In A. Roginska & P. Geluso (Hrsg.), Immersive sound: The art and science of binaural and multi-channel audio.
- Kock, M. (2019). Wie der Ton zum Bild passt: Wege zu effektivem Sounddesign. Fachverlag Schiele & Schön.
- Krause, B. L. (2016). Wild soundscapes: Discovering the voice of the natural world (Revised edition). Yale University Press.
- Kügler, C., & Theile, G. (1992, März). Loudspeaker Reproduction: Study on the Subwoofer Concept. Audio Engineering Society Convention 92. http://www.aes.org/elib/browse.cfm?elib=6798
- Kuttner, T. (2015). Praxiswissen Schwingungsmesstechnik. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-04638-5
- Lauterslager, H., & Nuyten, P. H. C. (1976, März). Ambience Related Transmission System (ARTS). Audio Engineering Society Convention 53. http://www.aes.org/elib/browse.cfm?elib=2322
- Lee, H. (2011, Oktober). The Relationship Between Interchannel Time and Level Differences in Vertical Sound Localization and Masking. Audio Engineering Society Convention 131. http://www.aes.org/e-lib/browse.cfm?elib=16082
- Lee, H. (2016, September). Capturing and Rendering 360° VR Audio Using Cardioid Microphones. Audio Engineering Society Conference: 2016 AES International Conference on Audio for Virtual and Augmented Reality. http://www.aes.org/elib/browse.cfm?elib=18511

- Lee, H. (2018, 06.04). Psychoacoustics of 3D sound recording techniques: Research and Practice [Schoeps Mikroforum 2018]. Schoeps Mikroforum 2018, Karlsruhe. https://www.youtube.com/watch?v=5SsQdfKAhVA
- Lensing, J. U. (2017). Sound-Design, Sound-Montage, Soundtrack-Komposition (3. Auflage). Fachverlag Schiele & Schön.
- Leonard, B. (2018). Applications of Extended Multichannel Technniques. In A. Roginska &P. Geluso (Hrsg.), Immersive sound: The art and science of binaural and multichannel audio.
- Licklider, J. C. R. (1948). The Influence of Interaural Phase Relations upon the Masking of Speech by White Noise. The Journal of the Acoustical Society of America, 20(2), 150–159. https://doi.org/10.1121/1.1906358
- Lund, T. (2020, Juni). Goodbye Stereo-Part I. AES Virtual Vienna Convention, Wien.
- M. Morimoto & Z. Maekawa. (1988). Effects of Low Frequency Components on Auditory Spaciousness. Acustica, 66(4), 190–196.
- Marks, A. (2017). Aaron Marks' complete guide to game audio: For composers, sound designers, musicians, and game developers (Third edition). CRC Press, Taylor & Francis Group.
- Martens, W. L. (1999, März). The Impact of Decorrelated Low-Frequency Reproduction on Auditory Spatial Imagery: Are Two Subwoofers Better Than One? Audio Engineering Society Conference: 16th International Conference: Spatial Sound Reproduction. http://www.aes.org/e-lib/browse.cfm?elib=8048
- Martens, W. L., Braasch, J., & Woszczyk, W. (2004, Oktober). Identification and Discrimination of Listener Envelopment Percepts Associated with Multiple Low-

Frequency Signals in Multichannel Sound Reproduction. Audio Engineering Society Convention 117. http://www.aes.org/e-lib/browse.cfm?elib=12886

- MathWorks Deutschland. (2020a). Cross power spectral density—MATLAB cpsd. Online Dokumentation. https://de.mathworks.com/help/signal/ref/cpsd.html
- MathWorks Deutschland. (2020b). Magnitude-squared coherence—MATLAB mscohere. Online Dokumentation. https://de.mathworks.com/help/signal/ref/mscohere.html?searchHighlight=mscoh ere&s_tid=srchtitle
- MathWorks Deutschland. (2020c). Welch's power spectral density estimate—MATLAB pwelch. Online Dokumentation. https://de.mathworks.com/help/signal/ref/pwelch.html
- Melchior, F. (2020). *3D-Audiosysteme* [Vorlesungsunterlagen, Hochschule der Medien Stuttgart].
- Morimoto, M. (1997, September). The Role of Rear Loudspeakers in Spatial Impression. *Audio Engineering Society Convention 103.* http://www.aes.org/elib/browse.cfm?elib=7225
- Murray, J. H. (1998). Hamlet on the holodeck: The future of narrative in cyberspace. MIT Press.
- Murray, L. (2019). Sound design theory and practice: Working with sound. Routledge.
- NEUMANN. (o. J.). Isobarendiagramm Neumann KH 310. Abgerufen 11. Dezember 2020, von https://de-de.neumann.com/kh-310-a

Peter, C. (2020). Grundlagen der Kinobeschallung. Schüren Verlag GmbH.

- Power, P., Davies, B., & Hirst, J. (2014, Oktober). Investigation into the Impact of 3D Surround Systems on Envelopment. Audio Engineering Society Convention 137. http://www.aes.org/e-lib/browse.cfm?elib=17431
- Pulkki, V., & Karjalainen, M. (2015). Communication acoustics: An introduction to speech, audio, and psychoacoustics. Wiley.
- Raffaseder, H. (2010). Audiodesign: Akustische Kommunikation, akustische Signale und Systeme, psychoakustische Grundlagen, Klangsysteme, Audioediting und Effektbearbeitung, Sounddesign, Bild-Ton-Beziehungen (2., aktualisierte und erw. Aufl). Hanser.
- Riekehof-Böhmer, H., Wittek, H., & Mores, R. (2010, November). Voraussage der wahrgenommenen räumlichen Breite stereofoner Mikrofonanordnungen. 26. Tonmeistertagung.
- Roginska, A. (2018). Binaural Audio Through Headphones. In A. Roginska & P. Geluso (Hrsg.), *Immersive sound: The art and science of binaural and multi-channel audio.*

Rumsey, F. (2001). Spatial audio. Focal Press.

- Rumsey, F. (2013). Cinema sound in the 3-D era. J. Audio Eng. Soc, 61(5), 340-344.
- Rumsey, F. (2020). Immersive audio—Defining and evaluating the experience. J. Audio Eng. Soc, 68(5), 388–392.
- Rumsey, F., & Lewis, W. (2002, April). Effect of Rear Microphone Spacing on Spatial Impression for Omnidirectional Surround Sound Microphone Arrays. Audio Engineering Society Convention 112. http://www.aes.org/elib/browse.cfm?elib=11421
- Sazdov, R., Paine, G., & Stevens, K. (2007, Juni). Perceptual Investigation into Envelopement, Spatial Clarity, and Engulfment in Reproduced Multi-Channel

Audio. Audio Engineering Society Conference: 31st International Conference: New Directions in High Resolution Audio. http://www.aes.org/elib/browse.cfm?elib=13961

- Schafer, R. M. (1993). The soundscape: Our sonic environment and the tuning of the world. Destiny Books.
- Schafer, R. M., & Breitsameter, S. (2010). Die Ordnung der Klänge: Eine Kulturgeschichte des Hörens. Schott.
- Schneider, M. (2008). Mikrofone. In S. Weinzierl (Hrsg.), Handbuch der Audiotechnik (S. 41–85). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-34301-1
- SCHOEPS Mikrofone. (o. J.-a). *Diffusfeld*. Abgerufen 6. Dezember 2020, von https://schoeps.de/wissen/knowledge-base/technische-grundlagen/diffusfeld.html
- SCHOEPS Mikrofone. (o. J.-b). *Image Assistant*. Abgerufen 2. August 2020, von https://schoeps.de/wissen/image-assistant.html
- SCHOEPS Mikrofone. (2020). Image Assistant. App Store. https://apps.apple.com/de/app/image-assistant/id1029051704
- Sengpiel, E. (2002). Gedanken zum Korrelationsgradmesser. http://www.sengpielaudio.com/GedankenZumKorrelationsgrad.pdf
- Slater, M., & Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. Presence: Teleoperators and Virtual Environments, 6(6), 603–616. https://doi.org/10.1162/pres.1997.6.6.603

Sofaconventions. (2020). SOFA API. SourceForge.

https://sourceforge.net/projects/sofacoustics/

- Sonnenschein, D. (2013). Sound Design: The expressive power of music, voice, and sound effects in cinema. Michael Wiese Productions.
- Soulodre, G. A., Lavoie, M. C., & Norcross, S. G. (2002, Oktober). Investigation of Listener Envelopment in Multichannel Surround Systems. Audio Engineering Society Convention 113. http://www.aes.org/e-lib/browse.cfm?elib=11283
- Soulodre, G. A., Lavoie, M. C., & Norcross, S. G. (2003). Objective Measures of Listener Envelopment in Multichannel Surround Systems. J. Audio Eng. Soc, 51(9), 826–840.
- Stevens, R., & Raybould, D. (2011). The game audio tutorial: A practical guide to sound and music for interactive games. Boston: Focal Press/Elsevier.
- Tellegen, A., & Atkinson, G. (1974). Openness to absorbing and self-altering experiences ("absorption"), a trait related to hypnotic susceptibility. Journal of Abnormal Psychology, 83(3), 268–277. https://doi.org/10.1037/h0036681
- The University of Huddersfield Enterprises Ltd. (2020). MARRS. App Store. https://apps.apple.com/us/app/marrs/id1295926126
- Theile, G. (2000, November). Microphone and mixing concepts for 5.1 music recordings. 21. Tonmeistertagung.
- Theile, G. (2001, Oktober). Multichannel Natural Music Recording Based on Psychoacoustic Principles. Nordic Sound Symposium.
- Theile, G., Dickreiter, M., Graul, W., Camerer, F., & Spikofski, G. (2014). Tonaufnahme und Tonwiedergabe. In *Handbuch der Tonstudiotechnik* (S. 217–369). De Gruyter Saur. https://doi.org/10.1515/9783110316506.217
- Theile, G., & Wittek, H. (2011a). Die dritte Dimension für Lautsprecher-Stereofonie. VdT-Magazin, 2011(2).

- Theile, G., & Wittek, H. (2011b, Mai). Principles in Surround Recordings with Height. *Audio Engineering Society Convention 130.* http://www.aes.org/elib/browse.cfm?elib=15870
- Tsingos, N., Chabanne, C., McCallus, M., & Robinson, C. (2010, November). Surround Sound with Height in Games Using Dolby Pro Logic IIz. Audio Engineering Society Convention 129. http://www.aes.org/e-lib/browse.cfm?elib=15670
- Viers, R. (2008). The sound effects bible: How to create and record Hollywood style sound effects. Michael Wiese Productions.
- Virostek, P. (2013). Field Recording: From research to wrap: an introduction to gathering sound effects (First edition). Airborne Publications.
- Wakuda, A., Furuya, H., Fujimoto, K., Isogai, K., & Anai, K. (2003). Effects of arrival direction of late sound on listener envelopment. Acoustical Science and Technology, 24(4), 179–185. https://doi.org/10.1250/ast.24.179
- Wallis, R., & Lee, H. (2015, Mai). Directional Bands Revisited. Audio Engineering Society Convention 138. http://www.aes.org/e-lib/browse.cfm?elib=17702
- Watson, N. (2013, Oktober). *Dolby Atmos erklärt*. Professional-Production. https://www.professionalproduction.de/fileadmin/user_upload/img/fachbeitrag/13/13-10/pdf/DolbyAtmos.pdf
- Weinzierl, S. (2008). Aufnahmeverfahren. In S. Weinzierl (Hrsg.), Handbuch der Audiotechnik (S. 551–607). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-34301-1_10
- Welch, P. (1967). The use of fast Fourier transform for the estimation of power spectra: A method based on time averaging over short, modified periodograms. *IEEE*

Transactions on Audio and Electroacoustics, 15(2), 70–73. https://doi.org/10.1109/TAU.1967.1161901

Wenzel, E. M., Begault, D. R., & Godfroy-Cooper, M. (2018). Perception of Spatial Sound. In A. Roginska & P. Geluso (Hrsg.), Immersive sound: The art and science of binaural and multi-channel audio.

Whittington, W. (2007). Sound design & science fiction (1st ed). University of Texas Press.

- Williams, M. (2004). Microphone arrays for stereo and multichannel sound recordings. Editrice Il Rostro.
- Williams, M. (2013). Microphone arrays for stereo and multichannel sound recording: Volume II. Editrice Il Rostro.
- Witmer, B. G., & Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. Presence: Teleoperators and Virtual Environments, 7(3), 225–240. https://doi.org/10.1162/105474698565686
- Wittek, H. (o. J.). ORTF-3D. Abgerufen 22. August 2020, von https://www.hauptmikrofon.de/3d/ortf-3d
- Wittek, H. (2012, November). Mikrofontechniken f
 ür Atmoaufnahme in 2.0 und 5.1 und deren Eigenschaften. 27. Tonmeistertagung.
- Wittek, H. (2015, September 20). "ORTF-3D": Eine Mikrofontechnik für Atmoaufnahmen in 3D-Audio und VR. https://www.hauptmikrofon.de/de/3d/ortf-3d
- Wittek, H., & Theile, G. (2016, November). Die Anwendung eines stereofonen Mehrkanalverfahrens für 3D-Audio und VR. 29th Tonmeistertagung, Köln.
- Wuttke, J. (2000). Mikrofonaufsätze (2. Aufl.). Schalltechnik Dr.-Ing. Schoeps GmbH.
- Wyatt, H., & Amyes, T. (2005). Audio post production for television and film: An introduction to technology and techniques (3rd ed). Focal Press.
- Yewdall, D. L. (2012). Practical art of motion picture sound (4th ed). Focal Press.
- Zdanowicz, G., & Bambrick, S. (2020). The game audio strategy guide: A practical course. Routledge, Taylor & Francis Group.

Anhang

A1 MATLAB Live-Skripte

A1.1 SOFA-Analyse

SOFA-Analyse

Matlab Live-Skript zur Auswertung von HRTF-Datensätzen im SOFA-Format nach AES 69-2015. Zur Ausführung wird die ita-toolbox und die Sofa API benötigt.

Auslesen der Metadaten

```
% starting SOFA API and loading SOFA-File
 SOFAstart;
 warning('off','SOFA:Upgrade');
 MySofaObject = SOFAload('SOFA\HRIR FULL2DEG.sofa')
MySofaObject = struct with fields:
                GLOBAL_Conventions: 'SOFA'
    GLOBAL_Version: '1.0'
GLOBAL_SOFAConventions: 'SimpleFreeFieldHRIR'
GLOBAL_SOFAConventionsVersion: '1.0'
                     GLOBAL_APIName: 'ARI SOFA API for Matlab/Octave'
                  GLOBAL APIVersion: '0.4.0'
            GLOBAL_ApplicationName: ''
         GLOBAL_ApplicationVersion: ''
              GLOBAL_AuthorContact: 'mail benjamin.bernschuetz@fh-koeln.de, GSM
+49 171 4176069'
                     GLOBAL_Comment: 'Converted from the original file format by
Piotr Majdak (piotr@majdak.com), Acoustics Research Institute, OeAW'
                    GLOBAL_DataType: 'FIR'
                     GLOBAL_History: 'Converted from the miro file format Upgraded
from SOFA 0.6'
               GLOBAL_License: 'CC 3.0 BY-SA'
GLOBAL_Organization: 'Fachhochschule K@ln, Germany'
GLOBAL_References: 'Bernsch@tz, B. (2013). "A Spherical Far Field
HRIR/HRTF Compilation of the Neumann KU 100", proceedings of the AIA/DAGA, Meran,
Italy'
                    GLOBAL RoomType: 'free field'
                      GLOBAL_Origin: 'http://www.audiogroup.web.fh-
koeln.de/ku100hrir.html'
                GLOBAL_DateCreated: '2012-08-23 10:56:10'
GLOBAL_DateModified: '2014-03-20 16:35:28'
                       GLOBAL_Title: 'HRTF'
               GLOBAL_DatabaseName: 'FHK'
          GLOBAL_ListenerShortName: 'HRIR_FULL2DEG'
                      GLOBAL_Author: 'Benjamin Bernsch@tz'
        GLOBAL_ListenerDescription: 'Neumann KU100'
        GLOBAL_ReceiverDescription: 'Neumann KU100; Internal RME Fireface UCX
Preamps
          GLOBAL_SourceDescription: 'Genelec 8260'
            GLOBAL_RoomDescription: 'FHK ZW 8-4 (Anechoic Chamber)'
                                  API: [1×1 struct]
                   ListenerPosition: [0 0 0]
             ListenerPosition_Type: 'cartesian'
            ListenerPosition_Units: 'metre'
```

```
ReceiverPosition: [2×3 double]
ReceiverPosition_Type: 'cartesian'
ReceiverPosition_Units: 'metre'
SourcePosition_Type: 'spherical'
SourcePosition_Units: 'degree, degree, metre'
EmitterPosition: [0 0 0]
EmitterPosition_Type: 'cartesian'
EmitterPosition_Units: 'metre'
ListenerUp: [0 0 1]
ListenerView_Type: 'cartesian'
ListenerView_Units: 'metre'
Data: [1×1 struct]
```

SOFAinfo(MySofaObject)

HRTF

====

Anechoic HRTF mesurement done by Fachhochschule K@ln, Germany.

Contact: mail benjamin.bernschuetz@fh-koeln.de, GSM +49 171 4176069 License: CC 3.0 BY-SA URL: http://www.audiogroup.web.fh-koeln.de/ku100hrir.html Reference: Bernsch@tz, B. (2013). "A Spherical Far Field HRIR/HRTF Compilation of the Neumann KU 100", proceedings of the AIA/DAGA, Meran, Italy

Measurement details:

Number of azimuth angles: 180 Number of elevation angles: 89 Number of radii: 1 Sampling Rate: 48000 hertz Dummy head: HRIR_FULL2DEG Loudspeaker: Genelec 8260 Listener at (0.0,0.0,0.0) metre 16020 source positions from (0.0,88.5,3.2) degree, degree, metre to (358.0,-88.5,3.2) degree, degree, metre

APV = SOFAcalculateAPV(MySofaObject); % extract source positions

3D-Plot aller Messpunkte

```
% plotting all mesured source positions of HRTF
f1 = figure('Name','HRTF Messpunkte');
HRTF_Pos =ita_SOFACoordinates2ItaHRTFCoordinates(MySofaObject);
plot(HRTF_Pos,'linestyle','none','Marker','.','Markersize',1)
hold on
plot3(0,0,0,'linestyle','none','Marker','.','Markersize',20,"Color",'red')
hold off
% plot style
title('HRTF Messpunkte')
xlabel('Abstand in m',"Rotation",23)
xticks([-3.2,0,3.2])
ylabel('Abstand in m',"Rotation",-33)
yticks([-3.2,0,3.2])
```



2D-Plot relevanter Messpunkte für (4+7+0) Lautsprecherkonfiguration

-3.2

-3.2

-3.2 3.2

Abstand in m

```
% creating vectors to separete azimuth and elevation measurement points
APV_azm = APV(:,1);
APV_elv = APV(:,2);
% calculationg needed length
n=0;
for i = 1:length(APV_azm)
    if ismember(APV_azm(i),[0;30;-30;90;-90;136;-136])
        n = n+1;
    end
end
% preallocating arrays for relevant measurements points
APV_azm_704=zeros(n,1);
APV_elv_704=zeros(n,1);
```

3.2

0 Abstand in m

```
% additional arrys for plot marking
 APV azm 704 mark=zeros(7,1);
 APV_elv_704_mark=zeros(7,1);
 APV elv 704 mark(:,:)=-2.50447806138254e-06;
 APV azm 704 mark h=zeros(4,1);
 APV_elv_704_mark_h=zeros(4,1);
APV elv 704 mark h(:,:)=46.256050109863280;
 % adding relevant datapoints to arrays
 ii=1;
 for k = 1:length(APV azm)
     if ismember(APV_azm(k),[0;30;-30;90;-90;136;-136])
         APV_azm_704(ii,1) = APV_azm(k,1);
         APV elv 704(ii,1) = APV elv(k,1);
         ii = ii+1;
     end
 end
% adding data points for plot marking
% elevated points
 APV_azm_704_mark_h(1,1)=APV_azm_704(90,1);
 APV_azm_704_mark_h(2,1)=APV_azm_704(268,1);
 APV_azm_704_mark_h(3,1)=APV_azm_704(375,1);
 APV_azm_704_mark_h(4,1)=APV_azm_704(535,1);
 % non-elevated points
 APV_azm_704_mark(1,1)=APV_azm_704(1,1);
 APV_azm_704_mark(2,1)=APV_azm_704(90,1);
 APV_azm_704_mark(3,1)=APV_azm_704(179,1);
 APV_azm_704_mark(4,1)=APV_azm_704(286,1);
 APV_azm_704_mark(5,1)=APV_azm_704(357,1);
 APV azm 704 mark(6,1)=APV azm 704(446,1);
APV_azm_704_mark(7,1)=APV_azm_704(535,1);
% creating plot for relevant HRIR measurements
f2 = figure('Name', 'Auswahl von HRTF-Messpunkten');
 plot(APV_azm_704, APV_elv_704, 'LineStyle', 'none', 'Marker', '.', "MarkerSize",
10)
hold on
% highlight needed speaker positions
 plot(APV_azm_704_mark,APV_elv_704_mark,'LineStyle','none','Marker','square
',"MarkerSize",15,"Color",'red')
plot(APV_azm_704_mark_h,APV_elv_704_mark_h,'LineStyle','none','Marker','sq
uare', "MarkerSize", 15, "Color", 'red')
% plot style
title('Auswahl von HRTF-Messpunkten für (4+7+0)
Lautsprecherkonfiguration')
 xlabel('Azimutwinkel in Grad')
```

```
ylabel('Elevationswinkel in Grad')
xlim([-180,180])
ylim([-10,90])
grid on
xticks([-180,-136,-90,-30,0,30,90,136,180])
yticks([0,30,46.256050109863280,60,90])
hold off
```



2D-Plot relevanter Messpunkte für (4+5+0) Lautsprecherkonfiguration

```
% calculationg needed length
k=0;
for l = 1:length(APV_azm)
    if ismember(APV_azm(l),[0;30;-30;110;-110])
        k = k+1;
    end
end
% preallocating arrays for relevant measurements points
APV_azm_504=zeros(k,1);
APV_elv_504=zeros(k,1);
% additional arrys for plot marking
APV_azm_504_mark=zeros(5,1);
APV_elv_504_mark=zeros(5,1);
APV_elv_504_mark(:,:)=-2.50447806138254e-06;
```

```
APV azm 504 mark h=zeros(4,1);
 APV elv 504 mark h=zeros(4,1);
APV_elv_504_mark_h(:,:)=46.256050109863280;
% adding relevant datapoints to arrays
 ii=1;
 for k = 1:length(APV azm)
     if ismember(APV azm(k),[0;30;-30;110;-110])
         APV_azm_504(ii,1) = APV_azm(k,1);
         APV_elv_504(ii,1) = APV_elv(k,1);
         ii = ii+1;
     end
 end
% adding data points for plot marking
% elevated points
 APV_azm_504_mark_h(1,1)=APV_azm_504(90,1);
 APV_azm_504_mark_h(2,1)=APV_azm_504(179,1);
 APV_azm_504_mark_h(3,1)=APV_azm_504(268,1);
 APV azm 504 mark h(4,1)=APV azm 504(357,1);
 % non-elevated points
 APV_azm_504_mark(1,1)=APV_azm_504(1,1);
 APV_azm_504_mark(2,1)=APV_azm_504(90,1);
 APV_azm_504_mark(3,1)=APV_azm_504(179,1);
 APV_azm_504_mark(4,1)=APV_azm_504(268,1);
APV_azm_504_mark(5,1)=APV_azm_504(357,1);
% creating plot for relevant HRIR measurements
f3 = figure('Name', 'Auswahl von HRTF-Messpunkten');
 plot(APV_azm_504, APV_elv_504, 'LineStyle', 'none', 'Marker', '.', "MarkerSize",
10)
hold on
% highlight needed speaker positions
plot(APV_azm_504_mark,APV_elv_504_mark,'LineStyle','none','Marker','square
', "MarkerSize", 15, "Color", 'red')
plot(APV_azm_504_mark_h,APV_elv_504_mark_h,'LineStyle', 'none', 'Marker', 'sq
uare', "MarkerSize", 15, "Color", 'red')
% plot style
title('Auswahl von HRTF-Messpunkten für (4+5+0)
Lautsprecherkonfiguration')
xlabel('Azimutwinkel in Grad')
ylabel('Elevationswinkel in Grad')
 xlim([-180,180])
ylim([-10,90])
 grid on
xticks([-180,-110,-30,0,30,110,180])
yticks([0,30,46.256050109863280,60,90])
```

hold off



A1.2 Kohärenz-Auswertung

Auswertung der Kohärenz von Mikrofonpaaren

Skript zu Analyse der Kohärenzfunktion (msc) von Mikrofonpaaren eines Mikrofonarrays. Die Darstellung der Funktionswerte erfolgt dabei gemittelt in drittel-oktav Abständen. Zur besseren Übersichtlichkeit werden die Funktionsgraphen nach den Layern des Mikrofonarrays aufgeteilt. Die Analyse kann für einen beliebigen Zeitabschnitt erfolgen.

Load audio data

```
% Samplerate
 fs = 48000;
 % Duration
 d = ([1+0*fs,120*fs]);
 % Get Infos
 audioinfo("06_IRT2/Atmo01_IRT_L.wav")
ans = struct with fields:
            Filename: '/Users/ElloAir/Documents/MATLAB/06_IRT2/Atmo01_IRT_L.wav'
   CompressionMethod: 'Uncompressed'
         NumChannels: 1
          SampleRate: 48000
        TotalSamples: 5800448
            Duration: 120.8427
               Title: []
             Comment: []
              Artist: []
       BitsPerSample: 24
 % Titel
 T_main = sprintf('Ort 1, Aufnahme 6, Doppel-IRT - Ausschnitt: Sekunde
%0.0f bis %0.0f',(d(1,1)/fs),(d(1,2))/fs);
 Tgraph_1 = ('Mikrofonpaare Bottom-Layer: Nieren 25 cm Kapselabstand');
 Tgraph_2 = ('Mikorofonpaare Height-Layer: Nieren 25 cm Kapselabstand');
 Tgraph_3 = ('Mikorofonpaare Bottom & Height-Layer: Nieren 25 cm
Kapselabstand');
 %Load Audio
 au_L = audioread("06_IRT2/Atmo01_IRT_L.wav",d);
 au R = audioread("06 IRT2/Atmo01 IRT R.wav",d);
 au_Ls = audioread("06_IRT2/Atmo01_IRT_Ls.wav",d);
 au_Rs = audioread("06_IRT2/Atmo01_IRT_Rs.wav",d);
 %au_Lss = audioread();
 %au_Rss = audioread();
 au_Ltf = audioread("06_IRT2/Atmo01_IRT_Ltf.wav",d);
 au_Rtf = audioread("06_IRT2/Atmo01_IRT_Rtf.wav",d);
 au_Ltr = audioread("06_IRT2/Atmo01_IRT_Ltr.wav",d);
 au_Rtr = audioread("06_IRT2/Atmo01_IRT_Rtr.wav",d);
```

Kohärenz (quadriert)

% nfft Size bestimmt die Frequenzauflösung % window Size bestimmt untere Grenzfrequenz

%calculate coherency

```
[Cohere_L_R,f]= mscohere(au_L,au_R,hann(16384),[],16384,fs);
[Cohere_L_Ls,~] = mscohere(au_L,au_Ls,hann(16384),[],16384,fs);
[Cohere_R_Rs,~] = mscohere(au_R,au_Rs,hann(16384),[],16384,fs);
[Cohere_Ls_Rs,~] = mscohere(au_Ls,au_Rs,hann(16384),[],16384,fs);
```

```
[Cohere_Ltf_Rtf,~]= mscohere(au_Ltf,au_Rtf,hann(16384),[],16384,fs);
[Cohere_Ltf_Ltr,~] = mscohere(au_Ltf,au_Ltr,hann(16384),[],16384,fs);
[Cohere_Rtf_Rtr,~] = mscohere(au_Rtf,au_Rtr,hann(16384),[],16384,fs);
[Cohere_Ltr_Rtr,~] = mscohere(au_Ltr,au_Rtr,hann(16384),[],16384,fs);
```

```
[Cohere_L_Ltf,~]= mscohere(au_L,au_Ltf,hann(16384),[],16384,fs);
[Cohere_R_Rtf,~] = mscohere(au_R,au_Rtf,hann(16384),[],16384,fs);
[Cohere_Ls_Ltr,~] = mscohere(au_Ls,au_Ltr,hann(16384),[],16384,fs);
[Cohere_Rs_Rtr,~] = mscohere(au_Rs,au_Rtr,hann(16384),[],16384,fs);
```

```
% Steps of 1/3 octave for datareduction/smoothing
```

```
MidFreq=[16 20 25 31.5 40 50 63 80 100 125 160 200 250 315 400 500 630 800 1000 1250 1600 2000 2500 3150 4000 5000 6300 8000 10000 12500 16000 20000];
FreqBands=[14.1 17.8 22.4 28.2 35.5 44.7 56.2 70.8 89.1 112 141 178 224 282 355 447 562 708 891 1122 1413 1778 2239 2818 3548 4467 5623 7079 8913 11220 14130 17780 22390];
```

%preallocate arrays

```
F_Arry_Cohere_L_R = [f,Cohere_L_R];
 F_thirdoctave_Cohere_L_R= nan(length(Cohere_L_R),length(FreqBands)-1);
 F_Arry_Cohere_L_Ls = [f,Cohere_L_Ls];
 F_thirdoctave_Cohere_L_Ls= nan(length(Cohere_L_Ls),length(FreqBands)-1);
 F_Arry_Cohere_R_Rs = [f,Cohere_R_Rs];
 F_thirdoctave_Cohere_R_Rs= nan(length(Cohere_R_Rs),length(FreqBands)-1);
 F_Arry_Cohere_Ls_Rs = [f,Cohere_Ls_Rs];
 F_thirdoctave_Cohere_Ls_Rs= nan(length(Cohere_Ls_Rs),length(FreqBands)-1);
 F_Arry_Cohere_Ltf_Rtf = [f,Cohere_Ltf_Rtf];
 F_thirdoctave_Cohere_Ltf_Rtf=
nan(length(Cohere_Ltf_Rtf),length(FreqBands)-1);
 F_Arry_Cohere_Ltf_Ltr = [f,Cohere_Ltf_Ltr];
 F_thirdoctave_Cohere_Ltf_Ltr=
nan(length(Cohere_Ltf_Ltr),length(FreqBands)-1);
 F_Arry_Cohere_Rtf_Rtr = [f,Cohere_Rtf_Rtr];
 F_thirdoctave_Cohere_Rtf_Rtr=
nan(length(Cohere_Rtf_Rtr),length(FreqBands)-1);
 F_Arry_Cohere_Ltr_Rtr = [f,Cohere_Ltr_Rtr];
```

```
F thirdoctave Cohere Ltr Rtr=
nan(length(Cohere Ltr Rtr),length(FreqBands)-1);
 F Arry Cohere L Ltf = [f,Cohere L Ltf];
 F thirdoctave Cohere L Ltf= nan(length(Cohere L Ltf),length(FreqBands)-1);
 F_Arry_Cohere_R_Rtf = [f,Cohere_R_Rtf];
 F thirdoctave Cohere R Rtf= nan(length(Cohere R Rtf),length(FreqBands)-1);
 F Arry Cohere Ls Ltr = [f,Cohere Ls Ltr];
 F_thirdoctave_Cohere_Ls_Ltr= nan(length(Cohere_Ls_Ltr),length(FreqBands)-
1);
 F Arry Cohere Rs Rtr = [f,Cohere Rs Rtr];
 F_thirdoctave_Cohere_Rs_Rtr= nan(length(Cohere_Rs_Rtr),length(FreqBands)-
1);
%calculate mean coherence per band
for i = 1:length(FreqBands)-1
     make_range = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Arry_Cohere_L_R)
         if F_Arry_Cohere_L_R(n,1) >= make_range(1) &&
F Arry Cohere L R(n,1) < make range(2)
             F_thirdoctave_Cohere_L_R(n,i) = F_Arry_Cohere_L_R(n,2);
         end
     end
 end
for i = 1:length(FreqBands)-1
     make_range = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Arry_Cohere_L_Ls)
         if F_Arry_Cohere_L_Ls(n,1) >= make_range(1) &&
F_Arry_Cohere_L_Ls(n,1) < make_range(2)</pre>
             F_thirdoctave_Cohere_L_Ls(n,i) = F_Arry_Cohere_L_Ls(n,2);
         end
     end
 end
for i = 1:length(FreqBands)-1
     make_range = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F Arry Cohere R Rs)
         if F_Arry_Cohere_R_Rs(n,1) >= make_range(1) &&
F_Arry_Cohere_R_Rs(n,1) < make_range(2)</pre>
             F_thirdoctave_Cohere_R_Rs(n,i) = F_Arry_Cohere_R_Rs(n,2);
         end
     end
 end
 for i = 1:length(FreqBands)-1
     make_range = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Arry_Cohere_Ls_Rs)
```

```
F Arry Cohere Ls Rs(n,1) < make range(2)
             F_thirdoctave_Cohere_Ls_Rs(n,i) = F_Arry_Cohere_Ls_Rs(n,2);
         end
     end
 end
 for i = 1:length(FreqBands)-1
     make range = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Arry_Cohere_Ltf_Rtf)
         if F_Arry_Cohere_Ltf_Rtf(n,1) >= make_range(1) &&
F Arry Cohere Ltf Rtf(n,1) < make range(2)
             F_thirdoctave_Cohere_Ltf_Rtf(n,i) =
F_Arry_Cohere_Ltf_Rtf(n,2);
         end
     end
 end
 for i = 1:length(FreqBands)-1
     make_range = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Arry_Cohere_Ltf_Ltr)
         if F_Arry_Cohere_Ltf_Ltr(n,1) >= make_range(1) &&
F Arry Cohere Ltf Ltr(n,1) < make range(2)
             F_thirdoctave_Cohere_Ltf_Ltr(n,i) =
F_Arry_Cohere_Ltf_Ltr(n,2);
         end
     end
 end
for i = 1:length(FreqBands)-1
     make_range = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Arry_Cohere_Rtf_Rtr)
         if F_Arry_Cohere_Rtf_Rtr(n,1) >= make_range(1) &&
F_Arry_Cohere_Rtf_Rtr(n,1) < make_range(2)</pre>
             F_thirdoctave_Cohere_Rtf_Rtr(n,i) =
F_Arry_Cohere_Rtf_Rtr(n,2);
         end
     end
 end
 for i = 1:length(FreqBands)-1
     make_range = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F Arry Cohere Ltr Rtr)
         if F_Arry_Cohere_Ltr_Rtr(n,1) >= make_range(1) &&
F_Arry_Cohere_Ltr_Rtr(n,1) < make_range(2)</pre>
             F_thirdoctave_Cohere_Ltr_Rtr(n,i) =
F_Arry_Cohere_Ltr_Rtr(n,2);
         end
     end
 end
 for i = 1:length(FreqBands)-1
     make_range = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Arry_Cohere_L_Ltf)
```

if F Arry Cohere Ls Rs(n,1) >= make range(1) &&

```
if F Arry Cohere L Ltf(n,1) >= make range(1) &&
F Arry Cohere L Ltf(n,1) < make range(2)
             F_thirdoctave_Cohere_L_Ltf(n,i) = F_Arry_Cohere_L_Ltf(n,2);
         end
     end
 end
for i = 1:length(FreqBands)-1
     make_range = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Arry_Cohere_R_Rtf)
         if F Arry Cohere R Rtf(n,1) >= make range(1) &&
F_Arry_Cohere_R_Rtf(n,1) < make_range(2)</pre>
             F_thirdoctave_Cohere_R_Rtf(n,i) = F_Arry_Cohere_R_Rtf(n,2);
         end
     end
 end
for i = 1:length(FreqBands)-1
     make_range = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Arry_Cohere_Ls_Ltr)
         if F Arry Cohere Ls Ltr(n,1) >= make range(1) \&\&
F_Arry_Cohere_Ls_Ltr(n,1) < make_range(2)</pre>
             F_thirdoctave_Cohere_Ls_Ltr(n,i) = F_Arry_Cohere_Ls_Ltr(n,2);
         end
     end
 end
for i = 1:length(FreqBands)-1
     make_range = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Arry_Cohere_Rs_Rtr)
         if F_Arry_Cohere_Rs_Rtr(n,1) >= make_range(1) &&
F Arry Cohere Rs Rtr(n,1) < make range(2)
             F_thirdoctave_Cohere_Rs_Rtr(n,i) = F_Arry_Cohere_Rs_Rtr(n,2);
         end
     end
 end
% mean numbers for smoothed curve
x = MidFreq;
y1 = mean(F_thirdoctave_Cohere_L_R, 'omitnan');
y2 = mean(F_thirdoctave_Cohere_L_Ls, 'omitnan');
y3 = mean(F_thirdoctave_Cohere_R_Rs, 'omitnan');
y4 = mean(F_thirdoctave_Cohere_Ls_Rs, 'omitnan');
y5 = mean(F_thirdoctave_Cohere_Ltf_Rtf, 'omitnan');
 y6 = mean(F_thirdoctave_Cohere_Ltf_Ltr, 'omitnan');
y7 = mean(F_thirdoctave_Cohere_Rtf_Rtr, 'omitnan');
y8 = mean(F_thirdoctave_Cohere_Ltr_Rtr, 'omitnan');
```

```
y9 = mean(F_thirdoctave_Cohere_L_Ltf, 'omitnan');
y10 = mean(F_thirdoctave_Cohere_R_Rtf, 'omitnan');
y11 = mean(F_thirdoctave_Cohere_Ls_Ltr, 'omitnan');
y12 = mean(F_thirdoctave_Cohere_Rs_Rtr, 'omitnan');
```

Kohärenz plotten Tiled-Layout

```
% plotting coherence over frequency
f3 = figure('Name','Kohärenz');
t = tiledlayout(3,1,'TileSpacing','compact',"Padding","compact");
title(t,T_main,'Fontsize',14,'FontWeight','bold')
 xlabel(t, 'Frequenz (Hz)')
ylabel(t, "Kohärenz")
 nexttile
 semilogx(x,y1,'Color',[0 0.4470 0.7410],'LineWidth',2,'Marker',"^")
 hold on
 semilogx(x,y2,'Color',[0.8500 0.3250 0.0980],'LineWidth',2,'Marker',"*")
 semilogx(x,y3,'Color',[0.9290 0.6940 0.1250],'LineWidth',2,'Marker',"+")
 semilogx(x,y4,'Color',[0.4660 0.6740 0.1880],'LineWidth',2,"Marker","0")
 title(Tgraph_1, 'Fontsize', 12, 'FontWeight', 'normal')
 ylim([-0.1,1.1])
 xlim([10,20000])
 xticks([10 20 40 100 200 400 1000 2000 4000 8000 16000])
 xticklabels({'10','20','40','100','200','400','1k','2k','4k','8k', '16k'})
 legend ('L - R', 'L - Ls', 'R - Rs', 'Ls - Rs', "Location", "northeast")
hold off
 grid on
 nexttile
 semilogx(x,y5,'Color',[0 0.4470 0.7410],'LineWidth',2,'Marker',"^")
 hold on
 semilogx(x,y6,'Color',[0.8500 0.3250 0.0980],'LineWidth',2,'Marker',"*")
 semilogx(x,y7,'Color',[0.9290 0.6940 0.1250],'LineWidth',2,'Marker',"+")
 semilogx(x,y8,'Color',[0.4660 0.6740 0.1880],'LineWidth',2,"Marker","0")
 title(Tgraph_2, 'Fontsize', 12, 'FontWeight', 'normal')
 ylim([-0.1,1.1])
 xlim([10,20000])
 xticks([10 20 40 100 200 400 1000 2000 4000 8000 16000])
xticklabels({'10','20','40','100','200','400','1k','2k','4k','8k', '16k'})
 legend ('Ltf - Rtf','Ltf - Ltr','Rtf - Rtr','Ltr -
Rtr', "Location", "northeast")
hold off
 grid on
 nexttile
```

```
semilogx(x,y9,'Color',[0 0.4470 0.7410],'LineWidth',2,'Marker',"^")
hold on
semilogx(x,y10,'Color',[0.8500 0.3250 0.0980],'LineWidth',2,'Marker',"*")
semilogx(x,y11,'Color',[0.9290 0.6940 0.1250],'LineWidth',2,'Marker',"+")
semilogx(x,y12,'Color',[0.4660 0.6740 0.1880],'LineWidth',2,"Marker","o")
title(Tgraph_3, 'Fontsize', 12, 'FontWeight', 'normal')
ylim([-0.1,1.1])
xlim([10,20000])
xticks([10 20 40 100 200 400 1000 2000 4000 8000 16000])
xticklabels({'10','20','40','100','200','400','1k','2k','4k','8k', '16k'})
legend ('L - Ltf', 'R - Rtf', 'Ls - Ltr', 'Rs - Rtr', "Location", "northeast")
grid on
hold off
```





Frequenz (Hz)

A1.3 DFI-P Berechnung

Berechnung des Diffusefield Image Predictors (DFI-P)

Skript zur Berechnung des DFI-P einer zweikanaligen Mikrofonanordnung. Die Berechnungen basieren auf den Ausführungen von Riekehof-Böhmer et al. (Voraussage der wahrgenommenen räumlichen Breite stereofoner Mikrofonanordnungen, 2010).

Load audio data

```
%samplerate
fs = 48000;
% Duration
d = ([1+0*fs,120*fs]);
%load audio data
audio_DFI_1 = audioread("06_IRT2/Atmo01_IRT_L.wav",d);
audio_DFI_2 = audioread("06_IRT2/Atmo01_IRT_R.wav",d);
%transpose arrays of audio data
audio_DFI_1_neu = audio_DFI_1';
audio_DFI_2_neu = audio_DFI_2';
```

Kohärenz (nicht quadriert)

```
pxx = pwelch(audio_DFI_1_neu,hann(16384),[],16384,fs);
pyy = pwelch(audio_DFI_2_neu,hann(16384),[],16384,fs);
[pxy,f_welch] =
cpsd(audio_DFI_1_neu,audio_DFI_2_neu,hann(16384),[],16384,fs);
my_msc = (abs(pxy)).^2./((pxx.*pyy));
my_mnsc = pxy./((pxx.^0.5).*(pyy.^0.5));
```

Testplot der quadrierten und nicht-quadrierten Kohärenzfunktion

```
% plotting
f1 = figure('Name','Kohärenz nichtquadriert');
warning('off')
semilogx(f_welch,my_mnsc)
hold on
semilogx(f_welch,my_msc)
hold off
xlim([10,20000])
```

```
ylim([-1,1])
xticks([20 40 60 100 200 400 1000 2000 4000 6000 10000 20000])
xticklabels({'20','40','60','100','200','400','1k','2k','4k','6k','10k','2
0k'})
grid on
legend('Nicht quadrierte Kohärenz (realanteil)','Quadrierte Kohärenz
(MSC)')
```



DFI Predictor

```
%Wichtungsfunktion
```

```
%gain factors over whole spectrum
f_gains = zeros(length(f_welch),3);
f_gains(:,1) = f_welch;
f_gains(:,2) = -3*log2(f_welch/400); % gain factors in dB;
f_gains(:,3) = 10.^(f_gains(:,2)/20); % gain factors absolute;
%reduce gain factors for definition area
for i = 1:length(f_gains(:,1))
    if f_gains(i,1) < 40
        f_gains (i,2) = intmin;
        f_gains (i,3) = 0;
    elseif f_gains(i,1) > 1500
        f_gains(i,2) = intmin;
        f_gains(i,2) = intmin;
        f_gains(i,2) = intmin;
        f_gains(i,3) = 0;
    end
```

end

```
% plotting Wichtungsfunktion
 f2= figure('Name', 'Wichtungsfunktion');
 semilogx(f_gains(:,1),f_gains(:,2))
 xlim([10,10000])
ylim([-12,12])
xticks([20 40 60 100 200 400 1000 2000 4000 6000 10000 20000])
 xticklabels({'20','40','60','100','200','400','1k','2k','4k','6k','10k','2
0k'})
yticks([-15 -12 -9 -6 -3 0 3 6 9 12 15])
yticklabels({'-15','-12','-9','-6','-3','0','3','6','9','12','15'})
xlabel('Frequenz (Hz)')
ylabel('Amplitude in dB')
% second axis for absolute gain factors
yyaxis right
 semilogx(f_gains(:,1),f_gains(:,2),'Color','none')
ylim([-12,12])
yticks([-12 -9 -6 -3 0 3 6 9 12])
yticklabels(compose('%1.3f',string(10.^([-12 -9 -6 -3 0 3 6 9 12]./20))))
ylabel('Gain-faktor absolut')
 ax = gca; % Get handle to current axes.
 ax.YColor = 'k'; % Red
 title('Wichtungsfunktion')
 grid on
```



```
% Anzahl der FFT-Stützstellen an denen die Wichtungsfunktion definiert ist
n_DFI = 0;
for i = 1:length(f_gains)
    if f_gains(i,3) ~= 0
        n_DFI = n_DFI+1;
    end
end
%print DFI value
my_dfi = 1/n_DFI * (sum((real(my_mnsc).*f_gains(:,3)).^2));
fprintf('DFI-Prädiktor: %.3f\n',my_dfi)
```

DFI-Prädiktor: 0.253

A1.4 Binaural-Analyse Wiedergabesystem 4+5+0

Binaural-Analyse Wiedergabesystem (4+5+0)

Skript zur Analyse synthetischer Binauralsignale. In einer virtuellen Abhörumgebung in MAX wurden dekorrelierte Rauschsignale über verschiedene Lautsprecherkonstellationen des 4+5+0-Wiedergabesystemsabgespielt und binauralisiert. Die Binauralsignale werden in diesem Skript auf die Kohärenz der Ohrsignale hin betrachtet und miteinander verglichen. Es soll untersucht werden, ob sich die Anzahl der Lautsprecher, welche dekorreliertes Rauschen wiedergaben, in der Kohärenz der Ohrsignale niederschlagen.

Laden der Audiodaten

```
% Get Infos
 audioinfo("Noise_SpeakerTest_450PN/Noise_SpeakerTest_450PN_C.wav");
% Samplerate
fs = 48000;
% Duration
 d = ([1+1*fs,60*fs]);
%Titel
MyTitel=({'Vergleich synthetischer Binauralsignale bei Wiedergabe von
dekorreliertem rosa Rauschen', "über verschiedene Lautsprecher des (4+5+0)
Wiedergabesystems",sprintf("Ausschnitt: Sekunde %0.0f bis
%0.0f",(d(1,1)/fs),(d(1,2))/fs)});
%Load Audio
Bi1 ST =
audioread("Noise SpeakerTest 450PN/Noise SpeakerTest 450PN C.wav",d);
 Bi1_L = Bi1_ST(:,1); % linker Kanal
Bi1_R = Bi1_ST(:,2); % rechter Kanal
 Bi2_ST =
audioread("Noise SpeakerTest 450PN/Noise SpeakerTest 450PN L R.wav",d);
 Bi2_L = Bi2_ST(:,1);
Bi2_R = Bi2_ST(:,2);
 Bi3 ST =
audioread("Noise_SpeakerTest_450/Noise_SpeakerTest_450_L_R_Ls_Rs.wav",d);
 Bi3_L = Bi3_ST(:,1);
Bi3_R = Bi3_ST(:,2);
Bi4 ST =
audioread("Noise_SpeakerTest_450PN/Noise_SpeakerTest_450PN_L_R_Ls_Rs_Height
.wav",d);
Bi4_L = Bi4_ST(:,1);
Bi4_R = Bi4_ST(:,2);
```

Berechnung und plotten der Kohärenzfunktionen

```
% nfft Size bestimmt die Frequnezauflösung
% window Size bestimmt untere Grenzfrequenz
 %Kohärenzberechnung
 [Cohere_BiL_BiR_1,f] = mscohere(Bi1_L,Bi1_R,hann(16384),[],16384,fs);
 [Cohere_BiL_BiR_2,~] = mscohere(Bi2_L,Bi2_R,hann(16384),[],16384,fs);
 [Cohere_BiL_BiR_3,~] = mscohere(Bi3_L,Bi3_R,hann(16384),[],16384,fs);
 [Cohere_BiL_BiR_4,~] = mscohere(Bi4_L,Bi4_R,hann(16384),[],16384,fs);
% Steps of 1/3 octave for datareduction/smoothing
MidFreq=[16 20 25 31.5 40 50 63 80 100 125 160 200 250 315 400 500 630 800
1000 1250 1600 2000 2500 3150 4000 5000 6300 8000 10000 12500 16000 20000];
 FregBands=[14.1 17.8 22.4 28.2 35.5 44.7 56.2 70.8 89.1 112 141 178 224
282 355 447 562 708 891 1122 1413 1778 2239 2818 3548 4467 5623 7079 8913
11220 14130 17780 22390];
%preallocate arrays
 F_Band_Array_Bi_1 = [f,Cohere_BiL_BiR_1];
 F thirdoctave Cohere Bi 1 =
nan(length(Cohere_BiL_BiR_1),length(FreqBands)-1);
 F_Band_Array_Bi_2 = [f,Cohere_BiL_BiR_2];
 F_thirdoctave_Cohere_Bi_2 =
nan(length(Cohere_BiL_BiR_2),length(FreqBands)-1);
 F_Band_Array_Bi_3 = [f,Cohere_BiL_BiR_3];
 F_thirdoctave_Cohere_Bi_3 =
nan(length(Cohere BiL BiR 3),length(FreqBands)-1);
 F Band Array Bi 4 = [f,Cohere BiL BiR 4];
 F_thirdoctave_Cohere_Bi_4 =
nan(length(Cohere_BiL_BiR_4),length(FreqBands)-1);
%sort coherence value per band
for i = 1:length(FreqBands)-1
     make range Bi = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Band_Array_Bi_1)
         if F_Band_Array_Bi_1(n,1) >= make_range_Bi(1) &&
F_Band_Array_Bi_1(n,1) < make_range_Bi(2)</pre>
             F_thirdoctave_Cohere_Bi_1(n,i) = F_Band_Array_Bi_1(n,2);
         end
     end
 end
 for i = 1:length(FreqBands)-1
     make_range_Bi = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F Band Array Bi 2)
         if F_Band_Array_Bi_2(n,1) >= make_range_Bi(1) &&
F_Band_Array_Bi_2(n,1) < make_range_Bi(2)</pre>
             F_thirdoctave_Cohere_Bi_2(n,i) = F_Band_Array_Bi_2(n,2);
         end
```

```
end
 end
for i = 1:length(FreqBands)-1
     make range Bi = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Band_Array_Bi_3)
         if F Band Array Bi 3(n,1) >= make range Bi(1) &&
F_Band_Array_Bi_3(n,1) < make_range_Bi(2)</pre>
             F_thirdoctave_Cohere_Bi_3(n,i) = F_Band_Array_Bi_3(n,2);
         end
     end
 end
for i = 1:length(FreqBands)-1
     make_range_Bi = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Band_Array_Bi_4)
         if F_Band_Array_Bi_4(n,1) >= make_range_Bi(1) &&
F_Band_Array_Bi_4(n,1) < make_range_Bi(2)</pre>
             F_thirdoctave_Cohere_Bi_4(n,i) = F_Band_Array_Bi_4(n,2);
         end
     end
end
% mean values for smoothed curve
 x = MidFreq;
y_Bi_1 = mean(F_thirdoctave_Cohere_Bi_1, 'omitnan');
y_Bi_2 = mean(F_thirdoctave_Cohere_Bi_2, 'omitnan');
y_Bi_3 = mean(F_thirdoctave_Cohere_Bi_3, 'omitnan');
y_Bi_4 = mean(F_thirdoctave_Cohere_Bi_4, 'omitnan');
% plotting coherence over frequency
 f2 = figure('Name','Binaural');
 semilogx(x,y_Bi_1,'LineWidth',2,'Marker',"square",'LineStyle',"-")
 hold on
 semilogx(x,y_Bi_2,'LineWidth',2,'Marker',"o",'LineStyle',"-")
 semilogx(x,y_Bi_3,'LineWidth',2,'Marker',"*",'LineStyle',"-")
 semilogx(x,y_Bi_4,'LineWidth',2,'Marker',"^",'LineStyle',"-")
 hold off
 legend ('C', 'L, R', "L, R, Ls, Rs", "L, R, Ls, Rs, Ltf, Rtf, Ltr,
Rtr", "Location", "best")
ylim([-0.1,1.1])
xlim([10,20000])
xticks([10 20 40 100 200 400 1000 2000 4000 8000 12000 16000])
xticklabels({'10','20','40','100','200','400','1k','2k','4k','8k','12k','1
6k'})
xlabel('Frequenz (Hz)')
ylabel("Kohärenz")
title(MyTitel, 'Fontsize', 14, 'FontWeight', 'bold')
```

grid <mark>on</mark>





A1.5 Binaural-Analyse Wiedergabesystem 4+7+0

Binaural-Analyse Speakertest (4+7+0)

Skript zur Analyse synthetischer Binauralsignale. In einer virtuellen Abhörumgebung in MAX wurden dekorrelierte Rauschsignale über verschiedene Lautsprecherkonstellationen des 4+7+0-Wiedergabesystemsabgespielt und binauralisiert. Die Binauralsignale werden in diesem Skript auf die Kohärenz der Ohrsignale hin untersucht und miteinander verglichen. Es soll untersucht werden, ob sich die Anzahl der Lautsprecher, welche dekorreliertes Rauschen wiedergaben, in der Kohärenz der Ohrsignale niederschlägt.

Laden der Audiodaten

```
% Get Infos
 audioinfo("Noise SpeakerTest 470PN/Noise SpeakerTest 470PN C.wav");
% Samplerate
fs = 48000;
% Duration
 d = ([1+1*fs,60*fs]);
%Titel
MyTitel=({'Vergleich synthetischer Binauralsignale bei Wiedergabe von
dekorreliertem rosa Rauschen', "über verschiedene Lautsprecher des (4+7+0)
Wiedergabesystems",sprintf("Ausschnitt: Sekunde %0.0f bis
%0.0f",(d(1,1)/fs),(d(1,2))/fs)});
%Load Audio
 Bi1 ST =
audioread("Noise SpeakerTest 470PN/Noise SpeakerTest 470PN C.wav",d);
 Bi1_L = Bi1_ST(:,1); % linker Kanal
Bi1_R = Bi1_ST(:,2); % rechter Kanal
 Bi2 ST =
audioread("Noise_SpeakerTest_470PN/Noise_SpeakerTest_470PN_L_R.wav",d);
 Bi2_L = Bi2_ST(:,1);
Bi2 R = Bi2 ST(:,2);
Bi3 ST =
audioread("Noise_SpeakerTest_470PN/Noise_SpeakerTest_470PN_L_R_Ls_Rs.wav",d
);
Bi3_L = Bi3_ST(:,1);
Bi3_R = Bi3_ST(:,2);
Bi4 ST =
audioread("Noise SpeakerTest 470PN/Noise SpeakerTest 470PN L R Ls Rs Lss Rs
s.wav",d);
Bi4_L = Bi4_ST(:,1);
Bi4_R = Bi4_ST(:,2);
```

```
Bi5_ST =
audioread("Noise_SpeakerTest_470PN/Noise_SpeakerTest_470PN_L_R_Ls_Rs_Height
.wav",d);
Bi5_L = Bi5_ST(:,1);
Bi5_R = Bi5_ST(:,2);
Bi6_ST =
audioread("Noise_SpeakerTest_470PN/Noise_SpeakerTest_470PN_L_R_Ls_Rs_Lss_Rs
s_Height.wav",d);
Bi6_L = Bi6_ST(:,1);
Bi6_R = Bi6_ST(:,2);
```

Berechnung und plotten der Kohärenzfunktionen

```
% nfft Size bestimmt die Frequnezauflösung
% window Size bestimmt untere Grenzfrequenz
```

%Kohärenzberechnung

<pre>[Cohere_BiL_BiR_1,f] = mscohere(Bi1_L,Bi1_R,hann(16384),[],16384,fs);</pre>
<pre>[Cohere_BiL_BiR_2,~] = mscohere(Bi2_L,Bi2_R,hann(16384),[],16384,fs);</pre>
<pre>[Cohere_BiL_BiR_3,~] = mscohere(Bi3_L,Bi3_R,hann(16384),[],16384,fs);</pre>
<pre>[Cohere_BiL_BiR_4,~] = mscohere(Bi4_L,Bi4_R,hann(16384),[],16384,fs);</pre>
<pre>[Cohere_BiL_BiR_5,~] = mscohere(Bi5_L,Bi5_R,hann(16384),[],16384,fs);</pre>
<pre>[Cohere_BiL_BiR_6,~] = mscohere(Bi6_L,Bi6_R,hann(16384),[],16384,fs);</pre>

% Steps of 1/3 octave for datareduction/smoothing

MidFreq=[16 20 25 31.5 40 50 63 80 100 125 160 200 250 315 400 500 630 800 1000 1250 1600 2000 2500 3150 4000 5000 6300 8000 10000 12500 16000 20000]; FreqBands=[14.1 17.8 22.4 28.2 35.5 44.7 56.2 70.8 89.1 112 141 178 224 282 355 447 562 708 891 1122 1413 1778 2239 2818 3548 4467 5623 7079 8913 11220 14130 17780 22390];

```
%preallocate arrays
 F_Band_Array_Bi_1 = [f,Cohere_BiL_BiR_1];
 F thirdoctave Cohere Bi 1 =
nan(length(Cohere_BiL_BiR_1),length(FreqBands)-1);
 F_Band_Array_Bi_2 = [f,Cohere_BiL_BiR_2];
 F thirdoctave Cohere Bi 2 =
nan(length(Cohere_BiL_BiR_2),length(FreqBands)-1);
F_Band_Array_Bi_3 = [f,Cohere_BiL_BiR_3];
 F thirdoctave Cohere Bi 3 =
nan(length(Cohere BiL BiR 3),length(FreqBands)-1);
 F_Band_Array_Bi_4 = [f,Cohere_BiL_BiR_4];
 F thirdoctave Cohere Bi 4 =
nan(length(Cohere BiL BiR 4),length(FreqBands)-1);
 F_Band_Array_Bi_5 = [f,Cohere_BiL_BiR_5];
 F_thirdoctave_Cohere_Bi_5 =
nan(length(Cohere BiL BiR 5),length(FreqBands)-1);
```

```
F Band Array Bi 6 = [f,Cohere BiL BiR 6];
 F thirdoctave Cohere Bi 6 =
nan(length(Cohere_BiL_BiR_6),length(FreqBands)-1);
%sort coherence value per band
for i = 1:length(FreqBands)-1
     make range Bi = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Band_Array_Bi_1)
         if F_Band_Array_Bi_1(n,1) >= make_range_Bi(1) &&
F Band Array Bi 1(n,1) < make range Bi(2)
             F_thirdoctave_Cohere_Bi_1(n,i) = F_Band_Array_Bi_1(n,2);
         end
     end
 end
 for i = 1:length(FreqBands)-1
     make_range_Bi = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Band_Array_Bi_2)
         if F_Band_Array_Bi_2(n,1) >= make_range_Bi(1) &&
F Band Array Bi 2(n,1) < make range Bi(2)
             F_thirdoctave_Cohere_Bi_2(n,i) = F_Band_Array_Bi_2(n,2);
         end
     end
 end
for i = 1:length(FreqBands)-1
     make_range_Bi = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Band_Array_Bi_3)
         if F_Band_Array_Bi_3(n,1) >= make_range_Bi(1) &&
F_Band_Array_Bi_3(n,1) < make_range_Bi(2)</pre>
             F_thirdoctave_Cohere_Bi_3(n,i) = F_Band_Array_Bi_3(n,2);
         end
     end
 end
for i = 1:length(FreqBands)-1
     make_range_Bi = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F Band Array Bi 4)
         if F_Band_Array_Bi_4(n,1) >= make_range_Bi(1) &&
F_Band_Array_Bi_4(n,1) < make_range_Bi(2)</pre>
             F_thirdoctave_Cohere_Bi_4(n,i) = F_Band_Array_Bi_4(n,2);
         end
     end
 end
 for i = 1:length(FreqBands)-1
     make_range_Bi = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Band_Array_Bi_5)
```

```
if F Band Array Bi 5(n,1) >= make range Bi(1) &&
F Band Array Bi 5(n,1) < make range Bi(2)
             F_thirdoctave_Cohere_Bi_5(n,i) = F_Band_Array_Bi_5(n,2);
         end
     end
 end
for i = 1:length(FreqBands)-1
     make_range_Bi = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F_Band_Array_Bi_6)
         if F Band Array Bi 6(n,1) >= make range Bi(1) &&
F_Band_Array_Bi_6(n,1) < make_range_Bi(2)</pre>
             F_thirdoctave_Cohere_Bi_6(n,i) = F_Band_Array_Bi_6(n,2);
         end
     end
 end
% mean values for smoothed curve
x = MidFreq;
y_Bi_1 = mean(F_thirdoctave_Cohere_Bi_1, 'omitnan');
y Bi 2 = mean(F thirdoctave Cohere Bi 2, 'omitnan');
y_Bi_3 = mean(F_thirdoctave_Cohere_Bi_3, 'omitnan');
y_Bi_4 = mean(F_thirdoctave_Cohere_Bi_4, 'omitnan');
y_Bi_5 = mean(F_thirdoctave_Cohere_Bi_5, 'omitnan');
y_Bi_6 = mean(F_thirdoctave_Cohere_Bi_6, 'omitnan');
% plotting coherence over frequency
 f2 = figure('Name','Binaural');
 semilogx(x,y_Bi_1,'LineWidth',2,'Marker',"square",'LineStyle',"-")
 hold on
 semilogx(x,y Bi 2,'LineWidth',2,'Marker',"o",'LineStyle',"-")
 semilogx(x,y_Bi_3,'LineWidth',2,'Marker',"*",'LineStyle',"-")
 semilogx(x,y_Bi_4,'LineWidth',2,'Marker',"^",'LineStyle',"-")
 semilogx(x,y_Bi_5,'LineWidth',2,'Marker',"v",'LineStyle',"-")
 semilogx(x,y Bi 6,'LineWidth',2,'Marker',"+",'LineStyle',"-")
hold off
legend ('C', 'L, R', "L, R, Lrs, Rrs", "L, R, Lss, Rss, Lrs, Rrs", "L, R, Lrs,
Rrs, Ltf, Rtf, Ltr, Rtr","L, R, Lss, Rss, Lrs, Rrs, Ltf, Rtf, Ltr,
Rtr", "Location", "best")
ylim([-0.1,1.1])
xlim([10,20000])
xticks([10 20 40 100 200 400 1000 2000 4000 8000 12000 16000])
 xticklabels({'10','20','40','100','200','400','1k','2k','4k','8k','12k',
'16k'})
xlabel('Frequenz (Hz)')
ylabel("Kohärenz")
title(MyTitel, 'Fontsize', 14, 'FontWeight', 'bold')
 grid on
```



rgleich synthetischer Binauralsignale bei Wiedergabe von dekorreliertem rosa Raus über verschiedene Lautsprecher des 4+7+0 Wiedergabesystems Ausschnitt: Sekunde 1 bis 60

A1.6 Binaural-Analyse zum Vergleich der Wiedergabesysteme

Binaural-Analyse Speakertest (4+5+0) vs (4+7+0)

Skript zur Analyse synthetischer Binauralsignale. In einer virtuellen Abhörumgebung in MAX wurden dekorrelierte Rauschsignale – rosa und weißes Rauschen – Wiedergabesysteme (4+5+0) und (4+7+0) abgespielt und binauralisiert. Die Kanäle Center und LFE wurden nicht genutzt. Die Binauralsignale werden in diesem Skript auf die Kohärenz der Ohrsignale hin untersucht und miteinander verglichen. Es soll untersucht werden, ob sich die Anzahl der Lautsprecher, welche dekorreliertes Rauschen wiedergaben, in der Kohärenz der Ohrsignale niederschlägt.

Laden der Audiodaten

```
% Get Infos
 audioinfo("Noise SpeakerTest 450PN/Noise SpeakerTest 450PN L R Ls Rs Heigh
t.wav");
% Samplerate
fs = 48000;
% Duration
 d = ([1+1*fs,60*fs]);
%Titel
MyTitel=({'Vergleich synthetischer Binauralsignale bei Wiedergabe von
dekorreliertem Rauschen', "über die Wiedergabesystem (4+5+0) und
4+7+0", sprintf("Ausschnitt: Sekunde %0.0f bis
%0.0f",(d(1,1)/fs),(d(1,2))/fs)});
%Load Audio
Bi1 ST =
audioread("Noise SpeakerTest 450PN/Noise SpeakerTest 450PN L R Ls Rs Height
.wav",d);
Bi1 L = Bi1 ST(:,1); % linker Kanal
Bi1 R = Bi1 ST(:,2); % rechter Kanal
Bi2 ST =
audioread("Noise_SpeakerTest_450/Noise_SpeakerTest_450_L_R_Ls_Rs_Height.wav
",d);
Bi2_L = Bi2_ST(:,1);
Bi2_R = Bi2_ST(:,2);
Bi3 ST =
audioread("Noise SpeakerTest 470PN/Noise SpeakerTest 470PN L R Ls Rs Lss Rs
s_Height.wav",d);
Bi3 L = Bi3 ST(:,1);
Bi3 R = Bi3 ST(:,2);
Bi4 ST =
audioread("Noise SpeakerTest 470/Noise SpeakerTest 470 L R Ls Rs Lss Rss He
ight.wav",d);
Bi4 L = Bi4 ST(:,1);
```

Bi4_R = Bi4_ST(:,2);

Berechnung und plotten der Kohärenzfunktionen

```
% nfft Size bestimmt die Frequnezauflösung
% window Size bestimmt untere Grenzfrequenz
```

```
%Kohärenzberechnung
```

```
[Cohere_BiL_BiR_1,f] = mscohere(Bi1_L,Bi1_R,hann(16384),[],16384,fs);
[Cohere_BiL_BiR_2,~] = mscohere(Bi2_L,Bi2_R,hann(16384),[],16384,fs);
[Cohere_BiL_BiR_3,~] = mscohere(Bi3_L,Bi3_R,hann(16384),[],16384,fs);
[Cohere_BiL_BiR_4,~] = mscohere(Bi4_L,Bi4_R,hann(16384),[],16384,fs);
```

```
% Steps of 1/3 octave for datareduction/smoothing
```

MidFreq=[16 20 25 31.5 40 50 63 80 100 125 160 200 250 315 400 500 630 800 1000 1250 1600 2000 2500 3150 4000 5000 6300 8000 10000 12500 16000 20000]; FreqBands=[14.1 17.8 22.4 28.2 35.5 44.7 56.2 70.8 89.1 112 141 178 224 282 355 447 562 708 891 1122 1413 1778 2239 2818 3548 4467 5623 7079 8913 11220 14130 17780 22390];

```
%preallocate arrays
 F Band Array Bi 1 = [f,Cohere BiL BiR 1];
 F_thirdoctave_Cohere_Bi_1 =
nan(length(Cohere_BiL_BiR_1),length(FreqBands)-1);
 F Band Array Bi 2 = [f,Cohere BiL BiR 2];
 F thirdoctave Cohere Bi 2 =
nan(length(Cohere_BiL_BiR_2),length(FreqBands)-1);
 F_Band_Array_Bi_3 = [f,Cohere_BiL_BiR_3];
 F thirdoctave Cohere Bi 3 =
nan(length(Cohere_BiL_BiR_3),length(FreqBands)-1);
 F_Band_Array_Bi_4 = [f,Cohere_BiL_BiR_4];
 F thirdoctave Cohere Bi 4 =
nan(length(Cohere_BiL_BiR_4),length(FreqBands)-1);
%sort coherence value per band
for i = 1:length(FreqBands)-1
     make_range_Bi = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F Band Array Bi 1)
         if F_Band_Array_Bi_1(n,1) >= make_range_Bi(1) &&
F_Band_Array_Bi_1(n,1) < make_range_Bi(2)</pre>
             F_thirdoctave_Cohere_Bi_1(n,i) = F_Band_Array_Bi_1(n,2);
         end
     end
```

end

```
for i = 1:length(FreqBands)-1
    make_range_Bi = [FreqBands(1,i),FreqBands(1,i+1)];
```

```
for n = 1 : length(F Band Array Bi 2)
         if F_Band_Array_Bi_2(n,1) >= make_range_Bi(1) &&
F_Band_Array_Bi_2(n,1) < make_range_Bi(2)</pre>
             F_thirdoctave_Cohere_Bi_2(n,i) = F_Band_Array_Bi_2(n,2);
         end
     end
 end
 for i = 1:length(FreqBands)-1
     make_range_Bi = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F Band Array Bi 3)
         if F_Band_Array_Bi_3(n,1) >= make_range_Bi(1) &&
F_Band_Array_Bi_3(n,1) < make_range_Bi(2)</pre>
             F thirdoctave Cohere Bi 3(n,i) = F Band Array Bi 3(n,2);
         end
     end
 end
for i = 1:length(FreqBands)-1
     make_range_Bi = [FreqBands(1,i),FreqBands(1,i+1)];
     for n = 1 : length(F Band Array Bi 4)
         if F_Band_Array_Bi_4(n,1) >= make_range_Bi(1) &&
F_Band_Array_Bi_4(n,1) < make_range_Bi(2)</pre>
             F_thirdoctave_Cohere_Bi_4(n,i) = F_Band_Array_Bi_4(n,2);
         end
     end
end
% mean values for smoothed curve
 x = MidFreq;
y_Bi_1 = mean(F_thirdoctave_Cohere_Bi_1, 'omitnan');
y Bi 2 = mean(F thirdoctave Cohere Bi 2, 'omitnan');
y Bi 3 = mean(F thirdoctave Cohere Bi 3, 'omitnan');
y_Bi_4 = mean(F_thirdoctave_Cohere_Bi_4, 'omitnan');
% plotting coherence over frequency
 f2 = figure('Name','Binaural');
 semilogx(x,y Bi 1,'LineWidth',2,'Marker',"square",'LineStyle',"-")
 hold on
 semilogx(x,y_Bi_2,'LineWidth',2,'Marker',"o",'LineStyle',"-")
 semilogx(x,y_Bi_3,'LineWidth',2,'Marker',"*",'LineStyle',"-")
 semilogx(x,y_Bi_4,'LineWidth',2,'Marker',"^",'LineStyle',"-")
 hold off
%legend ('C', 'L, R', "L, R, Ls, Rs", "L, R, Ls, Rs, Ltf, Rtf, Ltr,
Rtr", "Location", "best")
 legend ("Pink Noise (4+5+0) - L, R, Ls, Rs, Ltf, Rtf, Ltr, Rtr", ...
     "White Noise (4+5+0) - L, R, Ls, Rs, Ltf, Rtf, Ltr, Rtr", ...
     "Pink Noise (4+7+0) - L, R, Lss, Rss,Lrs,Rrs, Ltf, Rtf, Ltr, Rtr", ...
```

Vergleich synthetischer Binauralsignale bei Wiedergabe von dekorreliertem Rausch über die Wiedergabesystem 4+5+0 und 4+7+0

Ausschnitt: Sekunde 1 bis 60



A2 Kohärenzfunktionen der Mikrofonsysteme



A2.1 "Hamasaki+Height"





* Anmerkung: Aufgrund von Störgeräuschen musste der Gesamttake auf 114 Sekunden gekürzt werden.



Ort 3, Aufnahme 1, "Hamasaki+Height" – Ausschnitt: Sekunde 0 bis 120

A2.2 "Kugel-Array"





Frequenz (Hz)

Ort 2, Aufnahme 2, "Kugel-Array" - Ausschnitt: Sekunde 0 bis 120



A2.3 "Grenzflächen-Array"

* Anmerkung: Aufgrund von Störgeräuschen musste der Gesamttake auf 73 Sekunden gekürzt werden.


Ort 3, Aufnahme 3, "Grenzflächen-Array" – Ausschnitt: Sekunde 0 bis 120

A2.4 "Nieren-Array I"





Ort 2, Aufnahme 4, "Nieren-Array I" - Ausschnitt: Sekunde 0 bis 120



A2.5 "Nieren-Array II"

* Anmerkung: Aufgrund eines technischen Defekts konnte am Aufnahmeort 1 (Habichtswald) das "Nieren-Array II" nicht aufgezeichnet werden.



Ort 2, Aufnahme 5, "Nieren-Array II" - Ausschnitt: Sekunde 0 bis 120







A2.6 "Doppel-IRT"





A3 Kohärenzfunktionen der Kunstkopfaufnahmen und der binauralisierten Mikrofonsysteme im Vergleich



A3.1 "Hamasaki+Height"

A3.2 "Kugel-Array"

















