Bachelorarbeit im Studiengang Audiovisuelle Medien

Mehrwert von 3D-Audio in der Musikproduktion hinsichtlich Musikrichtung

vorgelegt von Yannik Weber
an der Hochschule der Medien Stuttgart
am 06.04.2019

zur Erlangung des akademischen Grades eines Bachelor of Engineering

Erst-Prüfer: Prof. Oliver Curdt

Zweit-Prüfer: Lasse Nipkow

Ehrenwörtliche Erklärung

Hiermit versichere ich, Yannik Weber, ehrenwörtlich, dass ich die vorliegende Bachelorarbeit mit dem Titel: "Mehrwert von 3D-Audio in der Musikproduktion hinsichtlich Musikrichtung" 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 (§ 24 Abs. 2 Bachelor-SPO (7 Semester) der HdM) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.

Stuttgart, den 06.04.2019	
	Yannik Weber

Danksagung

Zu Beginn möchte ich mich bei meinem Betreuer *Prof. Oliver Curdt* bedanken. Nicht nur für die Unterstützung bei Fragen zu dieser Arbeit, auch für die vielen lehrreichen Vorlesungen und Workshops während des Studiums. Sein leidenschaftliches Engagement trägt maßgeblich zur hohen Qualität der Lehrveranstaltungen und Produktionen in der Fachrichtung Ton bei. Auch möchte ich für die Möglichkeit danken, nahezu ohne Einschränkungen eigene Ideen und Semesterprojekte umsetzen zu können, wie dies bei *3Disco* der Fall war.

Im gleichen Atemzug möchte ich mich bei *Jörg Bauer*, dem technischen Mitarbeiter im Tonbereich, bedanken. Er ist bei technischen Problemen im Studio und Rückfragen stets zur Stelle. Ebenso großer Dank geht auch an *Prof. Jens-Helge Hergesell*, der in seinen Vorlesungen mit seiner Begeisterung für Tontechnik förmlich ansteckt.

Großer Dank gebührt auch meinem Zweit-Prüfer *Lasse Nipkow*, der mir während der gesamten Arbeit mit seinem enormen Know-how im Bereich 3D-Audio zuverlässig zur Seite stand. Seine Faszination und sein Engagement für 3D-Audio haben wesentlich zum erfolgreichen Abschluss der Arbeit beigetragen. Zudem hat er freundlicherweise das Klassik-Hörbeispiel zu Verfügung gestellt.

Weiter möchte ich mich bei allen beteiligten Tonschaffenden und Musikern bedanken. In Bezug auf das Jazz-Hörbeispiel besonders bei *Michael Thumm* und *Philipp Heck* der *Bauer Studios* für die Bereitstellung der Aufnahmen, *Jochen Neuffer* und der gesamten *Tobias Becker Bigband*. Die Umsetzung des Soul, Funk & Blues-Beispiels wäre ohne meine Mitmusiker *Lea Reister, Jan Bieletzki* und *Tobias Kroner* nicht möglich gewesen. Vielen Dank auch an *Lea* Reister für die Möglichkeit, *Change The World* als Pop-Hörbeispiel zu verwenden. Zudem möchte ich mich sowohl bei der gesamten Band *OXMO*, als auch bei *Jonas Müller* für die Bereitstellung der Musiktitel bedanken. Vielen Dank auch an alle *Probanden*, die sich für den Hörversuch Zeit genommen und diese Arbeit überhaupt erst ermöglicht haben.

Meinen Dank möchte ich ebenso *Johannes Frey* für die eigens angefertigten CAD-Modelle des 3D-Studios aussprechen. Danke auch an *Simon Hermentin* für das Bereitstellen des Filmtonstudio-Fotos, *rePhase*-Entwickler *Thomas Drugeon* für die ausgiebige Beratung in Sachen Raum-Entzerrung und *Ephraim Hahn* für das Bereitstellen seiner Arbeit.

Größter Dank geht an meine *Familie, Freunde, Kommilitonen* und zu guter Letzt meine *bessere Hälfte.*Sie sind in all meinen Vorhaben eine unglaubliche Stütze.

Zuletzt möchte ich Alfonso Bialetti für das Ermöglichen dieser Arbeit danken.

Kurzfassung

Die Aktualität des jungen Formates 3D-Audio ist in Anbetracht gegenwärtiger technischer Entwicklungen unumstritten. So findet 3D-Audio bereits vermehrt in Filmindustrie, Gamingindustrie und den "neuen Realitäten" wie Virtual Reality oder 360°-Videos Verwendung.

Aktuelle Forschungen zeigen einen, durch die gesteigerte Immersion hervorgerufenen, Anstieg der emotionalen Bewegung des Konsumenten durch 3D-Audio auf. Zum anderen zeichnet sich eindrücklich die Abhängigkeit der Wirkung von 3D vom verwendeten Beispiel ab.

Dies führt zur Forschungsfrage der vorliegenden Arbeit, die sich mit dem *Mehrwert von 3D-Audio in der Musikproduktion hinsichtlich Musikrichtung* beschäftigt. Dazu wurden mit der Produktion von Musiktiteln verschiedener musikalischer Genres in Stereo und *Auro-3D* die kreativen Möglichkeiten der Musikproduktion in 3D aufgezeigt. Die anschließende Evaluation erfolgte über einen Hörversuch.

Die Ergebnisse des Hörversuchs zeigen in den meisten Musikrichtungen vor allem emotional einen erheblichen Mehrwert durch 3D-Audio. Die Probanden sehen diesen zudem in Umhüllung und Transparenz des Klangs.

Abstract

In the light of current technical developments, the topicality of 3D audio is undisputed. The young audio format already found its popular use in the movie and gaming industry, as well as in the 'new realities' such as virtual reality or 360° videos.

Current research shows an increase in emotional movement through 3D audio caused by the enhanced immersion, and furthermore demonstrates a high-item dependency regarding the added value through 3D audio.

This leads to the research question of the present work, dealing with the evaluation of the *added* value of 3D audio in music production regarding different musical styles. Creative possibilities were pointed out with the production of songs of different musical styles in stereo and *Auro-3D*. The added value was evaluated afterwards through a listening test.

The results of the listening test show a significant increase in added value by 3D for most musical styles, especially emotionally by 3D audio. The immersion and transparency of 3D sound were remarked positively.

Inhaltsverzeichnis

Eł	nrenwö	irtlich	e Erklärung	3
D	anksag	ung		5
Κι	urzfassı	ung		7
A	bstract			7
In	haltsve	erzeicl	hnis	9
A	bkürzur	ngsve	rzeichnis	. 11
1.	Einl	eitun	g	. 13
2.	Vor	bereit	tung	. 15
	2.1	3D-A	Audio-Filmtonstudio der HdM	. 15
	2.2	Stud	lioausmessung und Kalibrierung	. 18
	2.2.	1	Pegelmessung	. 18
	2.2.	2	Frequenzgangmessung	. 19
	2.3	Wah	nl des 3D-Formates	. 27
	2.4	Proc	duktions methode	. 29
3.	Hör	beisp	iele	. 31
	3.1	Einte	eilung der Musikrichtungen	. 31
	3.2	Proc	duktion der Hörbeispiele	. 32
	3.2.	1	Klassik: PC-24, ein musikalischer Alpenflug	. 33
	3.2.	2	Jazz: Pennsylvania 6-5000	. 34
	3.2.	3	Soul, Funk & Blues: Love On Top	. 36
	3.2.	4	Rock: Rallye	. 37
	3.2.	5	Metal: Creeping Decay	. 39
	3.2.	6	Pop: Change The World	. 41
	3.2.	7	EDM: Rain On Your Skin	. 42
1	⊔är	vorcu	ch	17

	4.1	Konzeption des Hörversuchs		
	4.2	Fors	chungshypothesen	49
5.	Ausv	vertu	ing	50
	5.1	Gesa	amtanalyse	51
	5.2	Grup	ppenanalyse	54
	5.2.1	l	Hörtyp	54
	5.2.2	2	Alter	56
	5.2.3	3	Geschlecht	57
	5.3	Qua	litative Analyse und Interpretation	59
	5.3.1	L	Klassik	61
	5.3.2	2	Jazz	61
	5.3.3	3	Soul, Funk & Blues	62
	5.3.4	1	Rock	62
	5.3.5	5	Metal	63
	5.3.6	5	Pop	64
	5.3.7	7	EDM	64
6.	Fazit	und	Zukunft	65
Αı	nhang			67
	Daten-	CD		67
	Unterla	agen	zum Hörversuch	67
Αl	obildun	gsver	zeichnis	75
Ta	bellenv	erze	ichnis	77
Lit	teraturv	/erze	ichnis	78

Abkürzungsverzeichnis

Abkürzung	Bedeutung	
DAW	Digital Audio Workstation	
FDW	Frequency Dependent Window	
HdM	Hochschule der Medien	
HRTF	Head Related Transfer Function	
IEM	Institut für Elektronische Musik und Akustik Graz	
IR	Impulse Response	
MIDI	Musical Instrument Digital Interface	
REW	Room EQ Wizard	
S, F & B	Soul, Funk & Blues	
WDR	Westdeutscher Rundfunk	

Einheit	Bedeutung
dB	Dezibel
dBFS	Decibels relative to full scale
Hz	Hertz
m	Meter
cm	Zentimeter
S	Sekunden
ms	Millisekunden

Die Abkürzungen der 3D-Audiokanäle werden in Kapitel 2.1 erläutert.

1. Einleitung

Wirft man einen Blick auf die Geschichte der Schallaufzeichnung und -wiedergabe, so zeichnen sich mit der fortlaufenden Einführung neuer Formate auch tendenziell steigende Kanalanzahlen ab. War das 1887 erfundene Grammophon noch auf Mono-Wiedergabe beschränkt, konnten bereits auf der daraus hervorgehenden Vinyl-Schallplatte mit dem ±45°-Schriftverfahren stereophone Inhalte aufgenommen und wiedergegeben werden [1, pp. 609-610]. 1958 kamen derartige Stereo-Schallplatten in Deutschland erstmals auf den Markt, worauf der Hörfunksender *WDR 3* 1964 die ersten Sendungen in Stereo startete [2].

Die logische Weiterentwicklung zum Surround-Format wurde parallel von der Filmindustrie vorangetrieben, bis sich schließlich 1992 das bis heute gebräuchliche 5.1-System etablierte. Das System gab sein Debüt mit dem Film *Batman Returns* in *Dolby Digital*, das erstmals die Stereo-Anordnung um einen mittigen Center-, linken und rechten Surround- und einen LFE-Kanal für tieffrequenten Schall erweiterte. [1, pp. 781-782]

Wilfried van Baelen nahm diese Anordnung wiederum als Basis für fortführende Entwicklungen. 2005 experimentierte er erstmals mit dem Aufbau eines zweiten Höhen-Layers über dem Surround-Layer und schaffte damit eines der heute gängigen 3D-Audio-Formate (Auro-9.1). [3, p. 18]

Betrachtet man die Entwicklungen der Formate etwas abstrakter, ist eine stetige Weiterentwicklung der Dimensionen zu erkennen. Der Begriff Dimension beschreibt dabei die räumliche Ausdehnung des Klanges, was in Abbildung 1 veranschaulicht wird. Das noch junge Format 3D-Audio ist somit durch die dritte, über die Höhe abgebildete Dimension die konsequente Weiterentwicklung der Tonformate und gleichermaßen der "finale Schritt in der Klangreproduktion" [3, p. 18].

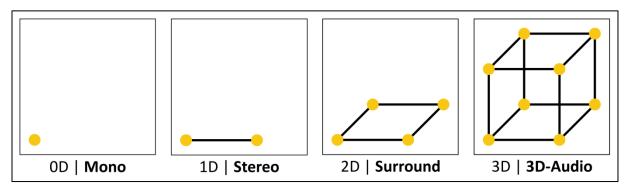


Abbildung 1: Dimensionen der verschiedenen Tonformate

Die Aktualität von 3D-Audio ist zurzeit nicht von der Hand zu weisen. Auf Fachtagungen wie der *Tonmeistertagung* in Köln oder der internationalen *AES Convention* gebührt 3D-Audio mittlerweile ein großer Teil der Vorträge. Von verschiedenen Distributionsformaten, über Mikrofontechniken und

Produktionstools, bis hin zu kreativen Methoden der Klanggestaltung, finden aktuell große Entwicklungen statt.

Auch beim Konsumenten kommt 3D-Audio zunehmend an. Neben der bereits vorangeschrittenen Implementierung von 3D-Audio-Systemen in Kinos [4] soll die Einführung der *AMBEO Soundbar* von *Sennheiser* im Mai 2019 der nächste große Schritt in Richtung Konsument sein. Diese verspricht mit 13 verbauten Lautsprecherchassis einen "vollkommene[n] 5.1.4-Sound mit einem einzigen Gerät". Erreicht wird dies über eine einmalige, selbstständige Raumkalibrierung und die anschließende Nutzung von Wandreflexionen. [5]

Sennheiser wirbt zudem damit, ein "vollkommen immersives, räumliches Audioerlebnis mit unglaublich hoher emotionaler Wirkung" erzeugen zu können [5]. Auch Auro-3D-Entwickler Baelen spricht von einem "großen emotionalen Mehrwert bei der Klangwiedergabe" durch 3D-Audio [3, p. 18]. So liegt es nahe, das Format auch für die Wiedergabe von Musik, ein Medium das unumstritten die Gefühlslage des Menschen beeinflussen kann [6–8], zu nutzen.

Ephraim Hahn zeigt mit seiner Bachelorarbeit "Evaluation des musikalischen Mehrwerts von 3D Audio in der klassischen Musikproduktion" bereits einen klaren Mehrwert durch 3D-Audio auf. Dabei beschränkt sich Hahn auf klassische Musik und stellt gleichermaßen die Notwendigkeit weiterer Forschungen in Bezug auf unterschiedliche musikalische Genres dar. [9]

Auch vergleichbare Arbeiten zeigen eine starke Abhängigkeit zwischen gespieltem Inhalt und empfundenem Mehrwert [10–12] auf. Dies lässt auf unterschiedliche Bewertungen für verschiedene Musikrichtungen schließen.

"[...] it can be concluded that there is a preference of the listeners for conditions with height reproduction and that this preference is, to some extent, item-dependent." [12, p. 339]

Ziel dieser Bachelorarbeit ist es, den Mehrwert von lautsprecherwiedergegebenem 3D-Audio in der Musikproduktion hinsichtlich verschiedener Musikrichtungen zu evaluieren. Kapitel 3.2 beschreibt die Auswahl und Produktion der verwendeten Hörbeispiele und beschäftigt sich zudem mit kreativen Methoden der Musikproduktion in 3D. Die sowohl eigens für die Bachelorarbeit angefertigten, als auch fremden Produktionen finden im anschließenden Hörversuch Verwendung (siehe Kapitel 4 und 5). Kapitel 6 schließt die Arbeit mit der Schlussbetrachtung ab, die unter anderem die Chancen von 3D-Audio in der Musikproduktion anhand der neu gewonnen Erkenntnisse einordnet. Doch zunächst werden in Kapitel 2 einige, für die spätere Arbeit essenzielle Vorbereitungen und Entscheidungen getroffen.

2. Vorbereitung

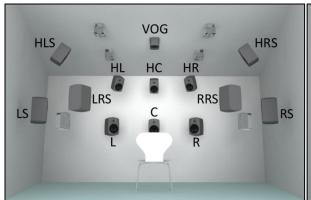
2.1 3D-Audio-Filmtonstudio der HdM

Das 3D-Audio-Filmtonstudio *U48* der *HdM* (siehe Abbildung 2) verfügt über eine 3D-Lautsprecheranordnung bestehend aus 20 Lautsprechern, also 19 Monitoren und einem Subwoofer. Diese sind so angeordnet, dass dort sowohl im Format *Auro-3D* (siehe Abbildung 3), als auch *Dolby Atmos* (siehe Abbildung 4) gearbeitet werden kann.



Abbildung 2: 3D-Audio-Filmtonstudio U48 der HdM Stuttgart (Foto: S. Hermentin)

Ein klassischer 9.1-Surround-Layer wird zum einen über den für *Auro-3D* üblichen, um 30° erhöhten 5-kanaligen Height-Layer und den Top-Speaker von oben erweitert. Zum anderen sind auf einer Ebene mit der zuletzt genannten *Voice of God* vier weitere Lautsprecher angebracht. Vorne und hinten sind je ein linker und ein rechter Lautsprecher angebracht, was gleichermaßen die Arbeit auf einer *Dolby Atmos*-Anordnung ermöglicht.



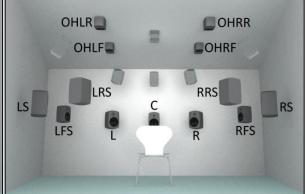


Abbildung 3: Auro-13.1 (Grafik: J. Frey & Y. Weber)

Abbildung 4: Atmos 9.1.4 (Grafik: J. Frey & Y. Weber)

Zur Beschreibung der Lautsprecherkonfiguration, werden bei *Auro-3D* die Kanäle aller Layer addiert. Die Zahl hinter dem Punkt beschreibt die Anzahl an LFE-Kanälen (z.B. 9.1). *Dolby Atmos* orientiert sich dagegen an der klassischen Surround-Namenskonvention (z.B. 5.1). Nach einem weiteren Punkt wird die Kanalanzahl des Overhead-Layers angegeben (bei vergleichbarem Setup 5.1.4). [13, 14]

Eine genauere Beschreibung, sowie der Vergleich beider Formate, sind in Kapitel 2.3 zu finden.

Wie in Abbildung 5 dargestellt, wird die Position der verschiedenen Lautsprecher über Winkel angegeben. Der erste Winkel bezieht sich auf die Transversalebene (horizontal), der zweite Winkel auf die Sagittalebene (vertikal).

Der 0°-Bezugspunkt fällt bei beiden Ebenen zusammen und ist in Blickrichtung vorne auf einer Höhe mit dem Kopf. Von diesem

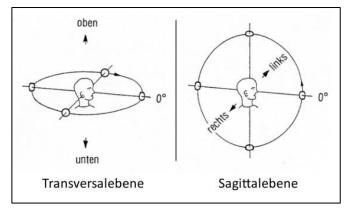


Abbildung 5: Ebenen zur Beschreibung der Lautsprecherpositionen (Grafik: Sengpiel)

Punkt aus bedeutet ein positiver Winkel eine Rotation nach rechts bzw. oben. Alle Lautsprecher auf der linken Seite werden mit den von 0° ausgehenden negativen Winkeln beschrieben.

In der folgenden Tabelle wird Auro-3D mit A3 und Dolby Atmos mit DA abgekürzt.

	Kurz	Kanalbezeichnung	Position	Layer	Lautsp.modell	Format [13, 14]
1	L	Left	-30°, 0°		Genelec 8040B	ab A3 9.1
2	R	Right	30°, 0°		Genelec 8040B	ab DA 5.1.2
3	С	Center	0°, 0°	1	Genelec 8040B	
4	LFE	Low Frequency Effect	-] 1	U48 Sub Machine*	
5	LS	Left Surround	-110°, 0°		Genelec 8040B	
6	RS	Right Surround	110°, 0°		Genelec 8040B	
7	HL	Height Left	-30°, 30°		Genelec 8030B	ab A3 9.1
8	HR	Height Right	30°, 30°		Genelec 8030B	
9	HC	Height Center	0°, 30°	0°, 30° 2 Genelec 8030B a		ab A3 11.1
10	HLS	Height Left Surround	-110°, 30°		Genelec 8030B	ab A3 9.1
11	HRS	Height Right Surround	110°, 30°		Genelec 8030B	
12	VOG, T	Voice Of God, Top	0°, 90°	3	Genelec 8020C	ab A3 10.1
13	LFS	Left Front Surround	-70°, 0°		Genelec 8030B	ab DA 9.1.2
14	RFS	Right Front Surround	70°, 0°		Genelec 8030B	
15	LRS	Left Rear Surround	-150°, 0°		Genelec 8030B	ab A3 11.1 (7+4)
16	RRS	Right Rear Surround	150°, 0°		Genelec 8030B	ab DA 7.1.2
17	OHLF	Overhead Left Front	Decke v.l.		Genelec 8020C	ab DA 5.1.2
18	OHRF	Overhead Right Front	Decke v.r.	3	Genelec 8020C	
19	OHLR	Overhead Left Rear	Decke h.l.	3	Genelec 8020C	ab DA 5.1.4
20	OHRR	Overhead Right Rear	Decke h.r.		Genelec 8020C	

^{*} in Tontechnik 4-Vorlesung (SS17) durch Udo Waldeck und Prof. Jens-Helge Hergesell konstruierter Subwoofer mit geschlossen-Hochpass-gefiltertem Gehäuse und 18"-Tieftöner mivoc AWX 184 MK II betrieben über eine Reckhorn A-409 Endstufe

Tabelle 1: Lautsprecherbestückung des 3D-Audio-Filmtonstudios der HdM

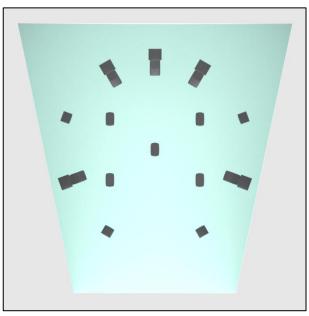


Abbildung 6: Grundfläche des Filmtonstudios (Grafik: J. Frey)

Als Audiointerface dient das im Studio integrierte *Penta 720 Modular I/O* von *NTP Technology*.

Die Grundfläche des Studios entspricht einem achsensymmetrischen Trapez (siehe Abbildung 6). Die hintere Wand des Studios, in der die Eingangstür liegt, hat eine Länge von 3,25 m. Die restlichen Wände haben eine Länge von 5,5 m. Damit hat der Raum eine Grundfläche von 23,6 m². Durch die Deckenhöhe von 2,7 m fasst der Raum 63,6 m³.

2.2 Studioausmessung und Kalibrierung

Die meisten 3D-Mischungen, sowie der Hörversuch dieser Arbeit, werden in diesem Studio realisiert. Dadurch ist der Klang des Studios von entscheidender Bedeutung. Zwar wurden die Lautsprecher nach Fertigstellung des Studios pegeltechnisch ausgemessen, doch der rege Betrieb im Studio, die frei zugänglichen Lautstärkeregler an den Lautsprechern, als auch Hörerfahrungen sind Grund für eine erneute Ausmessung und entsprechende Korrekturmaßnahmen.

2.2.1 Pegelmessung

Als Ausgangspunkt der Kalibrierung eines 3D-Studios kann die Messung einheitlicher Pegel bzw. der den Spezifikationen entsprechenden Pegel gesehen werden. Vor allem bei der Verwendung unterschiedlicher Lautsprechermodelle, wie es hier der Fall ist, ist dies unabdingbar.

Dazu wurde der linke Lautsprecher des gewöhnlichen Stereo-Setups als Referenz herangezogen und mit dem *Speaker Calibration Signal* der Software *Room EQ Wizard* bespielt. Die Gesamtlautstärke wurde so eingestellt, dass das SPL-Meter 80dB(C) S (C-gewichtet, langsame Zeitbewertung von 1 s) anzeigt. Der genaue Schalldruckpegel ist bei einer derartigen Messung allerdings zweitrangig. Es geht vielmehr um die Lautstärkerelationen der Lautsprecher untereinander.

Anschließend wurden alle weiteren 18 Monitore nacheinander mit dem gleichen Rauschen bespielt. Das Messmikrofon stand während der gesamten Messungen am Sweetspot, also dem idealen, zentralen Abhörpunkt zwischen den Lautsprechern. Bei Abweichungen wurde die Lautstärke direkt am Lautsprecher nachjustiert. Acht Lautsprecher wichen hierbei um ±1 dB oder mehr ab, die maximale Abweichung lag bei +3 dB.

Der Pegel des Subwoofers, der die LFE-Spur wiedergibt, wurde mit dem *Subwoofer Calibration Signal* gemessen und betrug richtigerweise 90dB(C) S, also ein um 10 dB größerer Schalldruckpegel als der, der Monitore.

Diese Konvention entspringt den Dolby Spezifikationen für die Wiedergabe der LFE-Spur. Sie sorgt dafür, dass im Bassbereich genügend Energie aufgebracht werden kann, ohne dafür im Gegenzug Headroom auf allen anderen Spuren einzubüßen. [15]

Bei den zur Kalibrierung verwendeten Testsignalen handelt es sich um gefiltertes rosa Rauschen. Das *Speaker Calibration Signal* zur Messung der Monitore bedient eine Bandbreite von 500 Hz bis 2 kHz (Trennfrequenz bei -3 dB, Flankensteilheit von 12 dB/Oktave) und umfasst somit die beiden Oktaven um 1 kHz. Das *Subwoofer Calibration Signal* beschneidet das Rauschen gleichermaßen auf eine

Bandbreite von 30 Hz bis 80 Hz. Somit entsprechen beide Testsignale weitgehend den *THX*-Testsignalempfehlungen. [16]

Während die Leistung bei weißem Rauschen in absoluten Frequenzintervallen gleich ist, ist sie dies bei dem verwendeten rosa Rauschen in Bezug auf relative Frequenzintervalle. Pro Frequenzverdopplung (sprich Oktave) nimmt der Pegel um 3 dB ab, was in etwa den statistischen Eigenschaften von Musik entspricht. Daher wird rosa Rauschen gerne für derartige Messungen verwendet. [17, p. 3]

2.2.2 Frequenzgangmessung

Um eine möglichst natürliche Wiedergabe im Tonstudio zu erreichen, sind allerdings nicht nur einheitliche Pegel wichtig. Ebenso essenziell ist ein ausgeglichener Frequenzgang über die gesamte Bandbreite der Lautsprecher, also ein möglichst konstanter Schalldruckpegel aller vom Lautsprecher bedienten Frequenzen. Dieser weicht dagegen oft stark vom Ideal ab. Grund hierfür ist häufig weniger der im Labor gemessene, solide erscheinende Frequenzgang des Lautsprechers, sondern vielmehr dessen Positionierung im Raum und der Raum selbst. Allgemein bekannte Phänomene sind hierbei beispielsweise die Bassanhebung bei wandnaher Lautsprecheraufstellung, Kammfiltereffekte, Reflexionen und Raummoden. Beim Aufbau eines Stereo-Setups kann vielen der aufgezählten Punkte mit entsprechend geschickter Positionierung der Lautsprecher entgegengewirkt werden, da sich zwei Lautsprecher um einiges flexibler im Raum platzieren lassen. [18]

Bei einem 3D-Setup stellt sich dies als schwieriger heraus. Häufig müssen Lautsprecher aufgrund von Platzmangel in unmittelbarer Wandnähe angebracht werden, da die Größe des 3D-Lautsprecherkäfigs logischerweise direkten Einfluss auf die maximale Anzahl an Person hat, die dort gleichzeitig hören können.

Beim Bau eines Stereo-Abhörstudios wird häufig der "Spiegel-Trick" angewandt. Eine Person setzt sich an den Sweetspot, eine andere Person sucht mit einem Spiegel parallel zu Wand und Decke die Stellen, an denen der Hörer den Lautsprecher sieht. An diesen Stellen werden gerne Breitbandabsorber aufgehängt, um somit frühe Reflexionen zu dämpfen. [19, p. 281]

In einem 3D-Studio ist ein solches Vorgehen nur begrenzt realisierbar und es müssen auch hier Kompromisse eingegangen werden. Im Filmtonstudio der *HdM* sind Wände und Decke komplett mit gelochten Gipsplatten bedeckt. Dahinter befindet sich ein 10 cm starkes Fasermaterial. Weitere Breitbandabsorber oder Bassfallen sind nicht vorhanden. Die umgebenden Lautsprecher hängen in

einem Käfig, der aus den aus Bussen und Bahnen bekannten Rohrsystemen gefertigt wurde. Einige Lautsprecher hängen dicht an Wand und Decke.

Wenn wie hier die Möglichkeit wegfällt, akustische Verbesserungen im Raum vorzunehmen, bleibt noch die Korrektur im Abhörweg, also vor oder an dem Lautsprecher selbst. Klangeinstellungen am Lautsprecher eignen sich häufig gut, um breitbandig seinen Grundklang zu verändern, doch für die hier angestrebten Veränderungen sind diese zu unpräzise. In vielen Studios sind nach wie vor grafische Equalizer zur Entzerrung zu finden. Allerdings bringen diese nur mit erhöhtem Zeit- und Geldaufwand Verbesserung in ein 20-Lautsprecher-Setup. Immer größerer Beliebtheit erfreuen sich zurzeit Klangkalibrierungs-Plugins, die an letzter Stelle im Masterbus der Produktionssoftware eingefügt werden und beim Export deaktiviert werden. Ein Beispiel dafür ist die Software Reference 4 des Unternehmens Sonarworks, das bereits seit 2012 an dem System arbeitet [20]. Reference 4 ist allerdings bislang nur zur Entzerrung von Stereo-Systemen ausgelegt.

Es muss also eine andere Möglichkeit gefunden werden, ein solches Mehrkanal-System zu entzerren. Die Freeware REW, die ebenso zur Messung einheitlicher Schalldruckpegel beigetragen hat, zeichnet sich durch zahllose Funktionen im Bereich akustischer Messungen aus. Ziel war es, das System mithilfe REW (https://www.roomegwizard.com/ [Version 5.19]) von (https://rephase.org/ [Version 1.4.3]), einer ebenfalls freien Anwendung zur Erstellung endlicher Impulsantworten aus Filterkurven, zu entzerren. Umgesetzt wurde ein auf der rePhase-Homepage tutorial: verlinktes empfohlenes Verfahren (,SwissBear's REW rePhase', http://forums.melaudia.net/attachment.php?aid=22240) in modifizierter Form. Als kostenpflichtiges Pendant dieser Methode kann der Mehrkanal-Software Monitor Controller Acourate Convolver Pro gesehen werden, der ebenfalls digitale Lautsprecher- und Raumentzerrung bietet [19, pp. 270-272].

Als Messmikrofon wurde ein *Sonarworks XREF20* verwendet, dessen individuelle Kalibrationsdatei ebenfalls in REW geladen wurde. Diese entzerrt den Frequenzgang des Mikrofons bei einer Bespielung von 30° abseits der Mittelachse, da das Mikrofon bei Messungen mit *Sonarworks* immer zwischen die Lautsprecher des Stereo-Paares zeigen sollte. Wesentlich praktikabler ist es, bei geplanter Messung mit dem Mikrofon an jedem Punkt direkt auf den Lautsprecher zu zeigen. Abbildung 7 veranschaulicht die Messung des Frequenzgangs eines Lautsprechers bei auf den Lautsprecher zeigendem Mikrofon (0°) und seitlichem Bespielen (30°). Pro Winkel wurden in der gleichen Position drei Messungen gemacht und gemittelt. Bei einem Winkel von 0° steigt der Frequenzgang gegenüber der 30° Messung ab 5 kHz gleichmäßig bis 20 kHz um ca. 1 dB an, was sich später invers auf die Korrekturfilter abzeichnen wird. Dies muss im Hinterkopf behalten werden, hat

aber voraussichtlich im Endeffekt eine vernachlässigbare Auswirkung auf den Gesamtklang (leichter Höhenabfall).

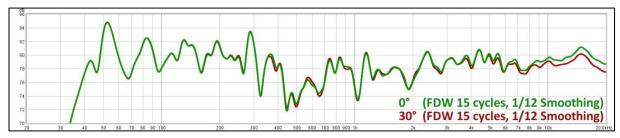


Abbildung 7: Frequenzgangmessung bei 0° und 30° Mikrofonausrichtung

Der Audio-In- und Output erfolgt über das USB-Audiointerface *Steinberg UR44*, für das ebenfalls eine Kalibrationsdatei vorliegt.

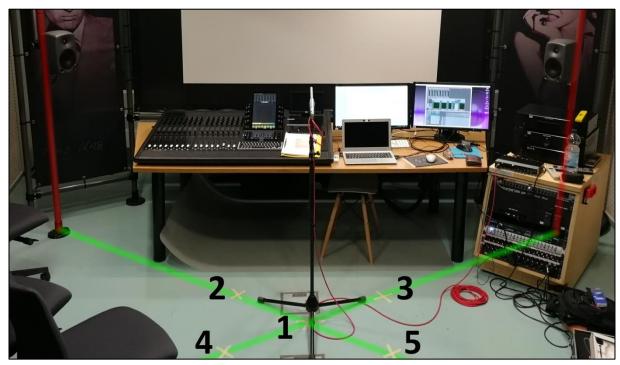


Abbildung 8: Messaufbau und -punkte

Startpunkt war die Festlegung und Markierung von fünf Messpunkten im Raum (siehe Abbildung 8). Mehrere Messpunkte sind insofern sinnvoll, da beim Messen an nur einer Stelle auch nur der Frequenzgang an exakt dieser Stelle beurteilt werden kann [21, pp. 136-137]. Steht das Messmikrofon an einem Punkt beispielsweise direkt in einer Raummode, kann der Frequenzgang im Bereich dieser Frequenz nach leichtem Verschieben des Mikrofons schon wieder völlig anders aussehen. Ebenso lassen sich bei in der Nähe von reflektierenden Flächen platzierten Lautsprechern häufig Kammfilter in den hohen Frequenzen beobachten. Auch diese können sich bei minimaler Bewegung des Mikrofons schon beachtlich verändern. Um diesen und weiteren möglichen

Korrekturfehlern entgegenzuwirken, wurde nach dem Messen eine Durchschnittskurve aus allen fünf Messungen berechnet und lokale Fehler somit ausgeglichen.

Als erster Messpunkt wurde der Sweetspot herangezogen. Die vier weiteren Punkte lagen auf zwei imaginäre Linien durch diesen und den LFS bzw. RFS (±70° von Mittelachse). Ausgehend vom Sweetspot wurden nun Punkte auf dieser Linie mit einem Abstand von 50 cm in beide Richtungen markiert. Die Höhe des Messmikrofons blieb bei der gesamten Messung bei 128 cm, da dies der Höhe der akustischen Achse der Lautsprecher entspricht, die idealerweise auf Ohrhöhe liegt. Die akustische Achse ist bei diesen Lautsprechern der Punkt zwischen Hoch- und Tieftönerkalotte [18, p. 15].

Die folgenden Abbildungen zeigen die Vorgehensweise exemplarisch anhand des rechten Lautsprechers R. Abbildung 9 zeigt die Kurven aller fünf Einzelmessungen, die teilweise stark voneinander abweichen. In Abbildung 10 liegen die Einzelmessung im Sweetspot und die aus allen fünf Einzelmessungen gemittelte Kurve übereinander. Einige Spitzen wurden durch die Miteinbeziehung der anderen Messungen besänftigt. Am eklatantesten ist der Einbruch um 600 Hz, der je nach Messposition wandert. Auch die in den Höhen aufgekommenen Kammfilter wurden durch die Mittelung in Summe geglättet.

Erwähnenswert hierbei ist, dass auf die Einzelmessungen angewandte Glättung (Smoothing) Einfluss auf die Berechnung der gemittelten Kurve hat. Alle gemittelten Kurven wurden aus ungeglätteten Einzelmessungen berechnet. Erst im Anschluss wurde eine 1/12-Oktavglättung auf alle Kurven angewandt. Dies schließt ebenfalls eine doppelte Glättung einer bereits aus geglätteten Einzelmessungen bestehenden, gemittelten Kurve aus. 1/12-Oktavglättung beschreibt die Aufteilung des Frequenzspektrums in, bei logarithmischer Betrachtung, zwölf gleichgroße Teile pro Oktave (im musikalischen Sinne Halbtöne) und deren Wertemittelung. Dadurch wird die Kurve glatter und übersichtlicher.

Die Messungen fanden mit den schon bei der Pegelmessung verwendeten Lautstärkeeinstellungen statt und das Messmikrofon wurde dementsprechend auf diesen Pegel kalibriert, um ebenso einen Überblick über die absoluten Schalldruckpegel zu bekommen.

In den Einstellungen von REW wurde zudem vor den Messungen ein Frequency Dependent Window (kurz FDW) mit einer Breite von 15 Zyklen gesetzt. Dies bedeutet, dass bei den Messungen frequenzabhängig lediglich die ersten 15 Schwingungsperioden zur Messung verwendet werden. [22]

Dadurch wird versucht, den Direktschall möglichst von Reflexionen und Nachhall zu trennen, um bei der späteren Filterung auch nur auf diesen und nicht auf eventuelle Resonanzen oder ähnliches einzugehen. Die Filterung einer überhöhten, späten Resonanz würde einen Einbruch der Frequenz im Direktschall bedeuten. Die Resonanz wird zwar abgeschwächt, bleibt aber trotzdem im Raum, und das Klangbild des Direktschalls leidet darunter.

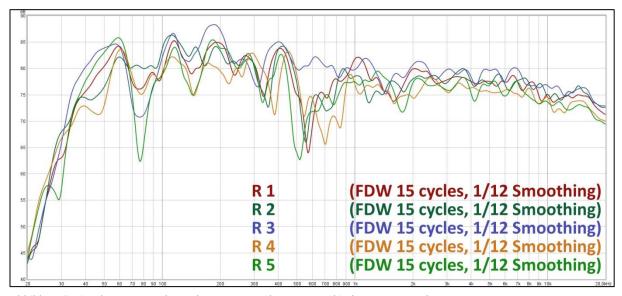


Abbildung 9: Einzelmessungen des rechten Lautsprechers an verschiedenen Messpunkten

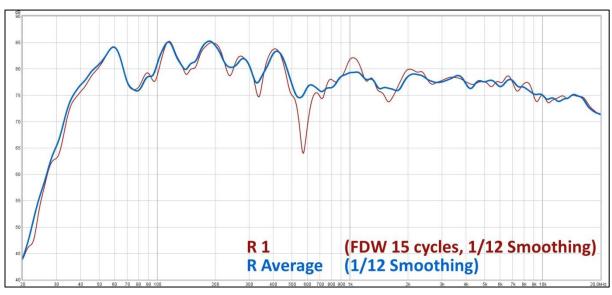


Abbildung 10: Einzelmessung am Sweetspot und gemittelte Kurve aus allen Messungen

Für die eigentliche Messung wurde *REWs Measurement Tool*, mit einem Sweep von 0 bis 20 kHz bei -12 dBFS (Kalibrierung ebenfalls mit -12 dbFS) und einer Länge von 256000 Samples (ca. 5,9 s bei einer Samplerate von 44,1 kHz) verwendet.

Die Messung wurde mit allen 20 Lautsprechern an allen fünf Messpunkten durchgeführt. Aus den insgesamt 100 Messungen wurden 20 Durchschnittskurven, also eine pro Lautsprecher, berechnet und im Anschluss mit 1/12-Oktavglättung versehen. Diese Glättung beeinflusst die spätere Filterberechnung und verhindert eine Überkorrektur des Frequenzgangs mit zu vielen Filtern.

Daraufhin wurde jede Messung mit dem EQ-Modul in REW geöffnet. In der *Equaliser*-Einstellung wurde *rePhase* angewählt. Bei *Target Settings* der *Speaker Type ,None'* ausgewählt und ein geradliniger und linearer Zielfrequenzgang eingestellt. In Anbetracht aller Frequenzgangmessungen und dessen, dass *rePhase* lediglich mit absenkender Filterung arbeitet, hat sich eine Ziellautstärke von 75 dB, also 5 dB unter der ursprünglichen Messlautstärke, als sinnvoll herausgestellt (dementsprechend 85dB bei dem LFE-Subwoofer). Durch die Verwendung einer einheitlichen Ziellautstärke wird sichergestellt, dass auch nach der Filterung alle Lautsprecher mit demselben Pegel spielen. Die *Filter Task*-Einstellungen waren wie folgt: *Match Range*: 20 to 20.000 Hz, *Individual Max Boost* und *Overall Max Boost*: 0 dB, *Flatness Target*: 2dB und die Checkbox *Allow narrow filters below 200 Hz* wurde nicht angewählt. Durch einen Klick auf *Match response to target* werden die Filter berechnet und ebenso der vorhergesagte Frequenzgang abgebildet (siehe Abbildung 11).

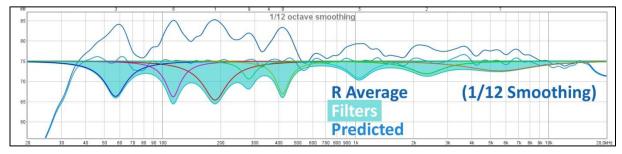


Abbildung 11: von REW erstellte Korrekturfilter

Über die Funktion Save filter settings to file lassen sich die Filtereinstellungen als eine in rePhase öffenbare Datei abspeichern.

Nach dem Öffnen von rePhase kann diese unter Paragraphic Gain EQ → Tools → import REW filter settings... importiert werden. Wie in Abbildung 12 zu sehen, wird dabei bei centering der Wert 0 eingetragen und unter optimization 'moderate' angewählt. Als Zielformat wird '24 bits LPCM mono ('wav)' eingestellt. Unter generate erstellt rePhase eine Impulsantwort mit vordefinierter Länge.

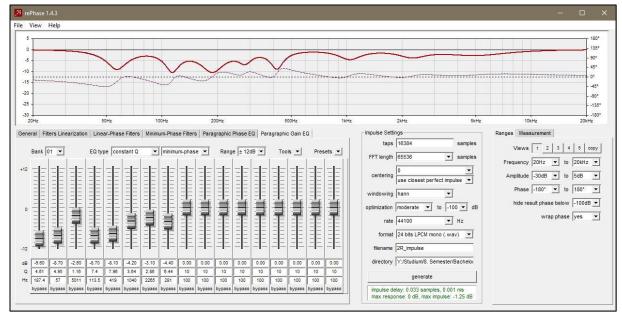


Abbildung 12: Filterkurve des rechten Lautsprechers in rePhase

Unter *EQ type* lässt sich neben der Default-Einstellung "minimum-phase" auch "linear-phase" anwählen. Ein testweiser Vergleich mit Analyse beider Einstellungen zeigt bei linearer Einstellung zwar einen linearen Phasengang, aber gleichzeitig auch einen enormen Einschwingvorgang gerade bei tiefen Frequenzen. Dies ist seinem Funktionsprinzip geschuldet. Die Impulsantwort eines linearphasigen Filters verläuft grundsätzlich symmetrisch, was in einer schlechteren klanglichen Abbildung resultieren kann [1, p. 577]. Die minimalphasige Filtereinstellung stellt dagegen einen guten Kompromiss zwischen Einschwingvorgang und Phasenlage dar.

Die oben beschriebenen Einstellungen haben sich nach eigenen Untersuchungen in Anbetracht von Latenz, Frequenz- und Phasengang des Filters als die für die Anwendung optimalen Einstellungen erwiesen. Ein *centering* von 0 legt die Spitze der Impulsantwort um das erste Sample fest. Dadurch wird der bei minimalphasigen Filtern relativ kurze Einschwingvorgang abgeschnitten und damit die Latenz reduziert, was sich im Frequenzgang niederschlägt. Die Option *,optimization'* gleicht dies wieder aus. Großer Dank geht hier an den *rePhase*-Entwickler Thomas Drugeon, der mit zahlreichen Anmerkungen und Tipps weiterhalf.

Sind die Impulsantworten für alle Lautsprecher erstellt, werden diese in den freien Audio-Editor *Audacity* (https://www.audacityteam.org/ [Version 2.3.0]) geladen. Durch extreme Vergrößerung der Wellenform (siehe Abbildung 13) kann beurteilt werden, an welcher Sampleposition die Letzte der Impulsantworten ein Sample mit einem Wert ungleich 0 besitzt. Das Ende aller Impulsantworten wird an der Sampleposition nach der letzten Stelle ungleich 0 der entsprechenden Impulsantwort beschnitten. In diesem Fall schwingt die Impulsantwort des LFE am längsten aus. Durch dieses Vorgehen können eventuelle unnötige Berechnungen vermieden und somit Rechenressourcen bei

der späteren Anwendung eingespart werden, da alle Impulsantworten eine vordefinierte Länge haben. Dies ist allerdings kein zwingend notwendiger Schritt.

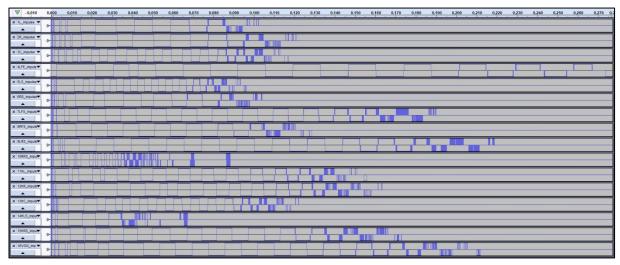


Abbildung 13: stark vergrößerte, bereits beschnittene IR-Waveforms der Lautsprecher 1-16

Für alle Kanäle, die über einen Stereo-Return beschickt werden (alle Kanäle, die paarweise links und rechts auftreten), wird direkt bei dieser Gelegenheit eine entsprechende Stereo-Impulsantwort erstellt. Alle anderen werden mono exportiert.

2.3 Wahl des 3D-Formates

Um die Musikstücke untereinander vergleichen zu können, sollte sich auf ein einheitliches Wiedergabeformat festgelegt werden. Zu den populärsten 3D-Audio Formaten zählen *Auro-3D* und *Dolby Atmos*.

Neben der Möglichkeit, Lautsprecher wie bei einer Surround-Produktion üblich kanalbasiert zu beschicken, können in *Dolby Atmos* über zusätzliche Spuren ebenso Objekte definiert werden. Diese werden mit Koordinaten versehen. Bei der Wiedergabe berechnet dann ein Prozessor, der die Lautsprecheranordnung kennt, über welche Lautsprecher die Objekte entsprechend ausgegeben werden. So kann die Lautsprecheranordnung flexibel skaliert werden. [1, p. 799]

Auro-3D ist zwar in seiner ursprünglichen Form kanalbasiert, wurde aber später um die objektbasierte Technologie *AuroMax* erweitert. Diese bringt bei der *Auro-*Lautsprecheranordnung aber erst bei großen Kinos einen Mehrwert gegenüber der kanalbasierten Methode. [3, p. 21]

Entscheidend bei der Wahl des Formates sollten allerdings nicht die Technik, sondern der Klang und die Möglichkeiten des Formates sein. Diese unterscheiden sich im Wesentlichen in den Methoden der Höhendarstellung.

Zur Darstellung der Höhe werden bei *Dolby Atmos* die an der Decke angebrachten, sogenannten *overhead speakers* (siehe Abbildung 14) verwendet. Optional finden auch gerne *Dolby Atmos enabled*

speakers Verwendung. Diese werfen den Schall gerichtet an die Decke. Von dort aus wird er zur Hörposition reflektiert. [14, pp. 7-13]

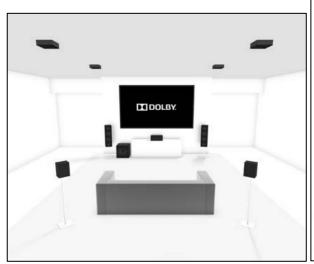








Abbildung 15: Auro-9.1 (oben) & 10.1 (unten) (Grafik: Auro)

Im vergleichbaren 9.1-Setup von *Auro-3D* (siehe Abbildung 15 oben) wird der gewöhnliche 5.1-Layer durch vier um 30° erhöhte Lautsprecher über L, R, LS und RS um den sogenannten Height-Layer erweitert. *Wilfried van Baelen*, Gründer und Entwickler von *Auro-3D*, begründet dies damit, dass die größte Immersion durch das vertikale Stereofeld entsteht. Ist der Winkel wie bei *Dolby Atmos* größer als 35°, geht die vertikale Kohärenz, also die Verbindung bzw. der akustische Zusammenhang zwischen Surround-Layer und Height-Layer, verloren. Dadurch werden die Natürlichkeit und Räumlichkeit des Klangerlebnisses gemindert, die für die Musikwiedergabe in 3D eine große Rolle spielen. Zudem ist das menschliche Gehör unempfindlicher für Reize von oben.

Im 10.1 Setup (siehe Abbildung 15 unten) wird die Anordnung durch den bei *Auro-3D* einzigartigen dritten Layer komplettiert. Der direkt über dem Hörer angebrachte Top Speaker, auch *Voice of God* genannt, vervollständigt letztlich die Abbildung einer Halbkugel um den Kopf des Hörers (siehe Abbildung 16). [3, p. 18]

Auch eigene Erfahrungen mit den Formaten können die Aussagen *Baelens*

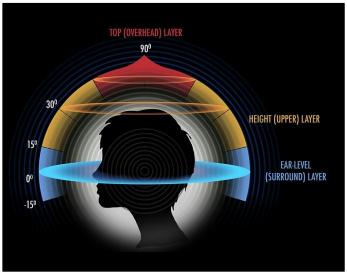


Abbildung 16: die verschiedenen Layer bei Auro-3D (Grafik: Auro)

nur bestätigen. Aufgrund dessen und der Tatsache, dass in *Auro-3D* bereits einige geeignete Produktionen vorhanden waren, wurde das Format *Auro-3D* in der 9.0-, 9.1- bzw. wenn es sich anbietet 10.1-Anordnung gewählt. Zudem stellt sich *Auro-3D* aktuell als sehr beliebtes und gängiges Format heraus, wenn es um reine Musikproduktionen in 3D geht [23].

2.4 Produktionsmethode

Bei der kreativen Musikproduktion sollte in einer gewohnten und damit intuitiven Produktionsumgebung gearbeitet werden, was in meinem Fall der Stereo-DAW *Ableton Live 10* entspricht. Zudem bietet es sich an, gerade bei eigens für 3D produzierten Titeln vom ersten Ton an direkt in 3D zu arbeiten, zu hören, wie sich die Sounds aus den entsprechenden Richtungen anhören und somit den Raum unmittelbar bei der Klanggestaltung miteinzubeziehen. Es ist also keine Option, die Multitracks zu exportieren und nachträglich auf dem *Protools*-Rechner der Hochschule mit den proprietären *Auro-3D*-Tools zu mischen. Es muss ein Weg gefunden werden, mehrdimensional mit *Ableton* arbeiten zu können.

Die Plugin-Lösung Spatial Audio Designer von New Audio Technology verspricht beispielsweise, in jeder DAW immersive Produktionen (Surround und 3D) durchführen zu können. Ein weiterer Vorteil des Spatial Audio Designers ist Möglichkeit, auch noch nachträglich das gewünschte Endformat zu bestimmen.

Die über den 3D-Panner

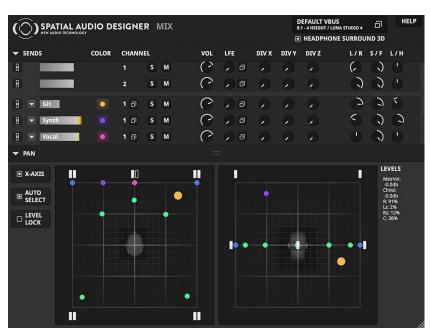


Abbildung 17: Spatial Audio Designer von New Audio Technology (Screenshot: NAT)

festgelegten Positionen der Spuren bestimmen letztendlich, über welche Lautsprecher die entsprechenden Signale wiedergegeben werden. Durch diverse Probleme und Instabilitäten des Plugins in *Ableton* kann das Tool allerdings nicht zur Produktion verwendet werden.

Die finale Lösung ist, die Kanäle über entsprechende Mono- und Stereo-Sends zu beschicken, deren Returns auf die Ausgänge des Interfaces geroutet sind. Auf den Return-Spuren liegt zudem ein Plugin zur Faltung mit den in Kapitel 2.2.2 erstellten Impulsantworten. Hierbei ist auf eine neutrale Einstellung, also ausschließliche Faltung, zu achten. Zur Faltung wird der *Convolution Reverb* von *Max for Live* verwendet.

Beim Panning über Sends handelt es sich um eine Methode mit der sich tadellos arbeiten lässt, die in allen eigenen 3D-Mischungen (siehe Kapitel 3.2) Anwendung findet und ebenso Vorteile mit sich bringt. Mit Panning-Tools werden gerne unbedacht Positionen zwischen Lautsprechern gewählt.

Dabei sollte man sich immer im Klaren sein, dass Realschallquellen (Positionierung des Elements auf Lautsprecher) die maximal mögliche Lokalisationsschärfe bieten, zumal das Gehör seitliche Phantomschallquellen nicht zuverlässig orten kann [17, p. 242, 24, p. 23]. Dies wird auch bei nachfolgenden Mischungen berücksichtigt, um einen möglichst transparenten Gesamtklang zu erhalten. Zudem verleitet die grafische Darstellung des Tools leicht dazu, Objekte innerhalb des Raums zu platzieren. Mit einem derartigen Setup können allerdings keine Klangquellen im Raum dargestellt werden, die Phantomschallquelle erfährt lediglich eine größere räumliche Ausdehnung. Bei Verteilung der Instrumente sollte man sich die klanglichen Möglichkeiten stets über die Projektion des Klanges auf der Außenfläche des 3D-Setups verdeutlichen. [24, p. 23]



Komplexere Panning-Fahrten werden über die *Max for Live-*Erweiterung *Surround Panner* realisiert (siehe Abbildung 18).

Abbildung 18: Max for Lives Surround Panner im Rock-Beispiel

Die Erzeugung von Hall und Effekten wird in den Produktionsbeschreibungen erläutert (siehe Kapitel 3.2).

Schnittstelle zwischen Rechner und Lautsprecher ist das im Studio verbaute Interface *Penta 720 Modular I/O* von *NTP Technology*. Die auf dem Rechner installierte *Dante Virtual Soundcard* bespielt dabei den *Dante-*Eingang des Interfaces, der auf dessen analoge Outputs durchgeschaltet ist.

3. Hörbeispiele

3.1 Einteilung der Musikrichtungen

Um eine Beurteilung des Mehrwertes von 3D-Audio in Bezug auf die Musikrichtung vornehmen zu können, muss zwingend eine Genreeinteilung stattfinden. Eine allgemein gültige Einteilung ist in der Wissenschaft allerdings nicht zu finden, da die Grenzen zwischen den Genres oft fließend verlaufen. Unterschiedliche Ansätze teilen die Musikrichtungen unterschiedlich tiefgehend ein. Auch Einflüsse durch lokale musikalische Einschläge sind meist nicht von der Hand zu weisen.

Es muss also eine passende Einteilung gefunden werden und diese auch in Bezug auf das spätere Forschungsziel angeglichen werden. Bei der Kategorisierung sollen viel mehr Produktionspraktiken und musikalische Eigenschaften wie Instrumentierung, Klang und Struktur, als historische oder lokale Hintergründe betrachtet werden.

Ein für den Hörversuch sinnvoll scheinender Startpunk ist die Einteilung des *STOMP* (*Short Test Of Music Preferences*). Zur Feststellung musikalischer Präferenzen teilt dieser die Musik in die folgenden 14 verschiedenen Musikrichtungen auf:

Blues – Jazz – Classical – Folk – Rock – Alternative – Heavy Metal – Country – Sound Tracks – Religious – Pop – Rap & Hip-Hop – Soul & Funk – Electronica & Dance [25, p. 1241]

Sound Tracks und Religious werden bei dieser Arbeit nicht betrachtet, da sich diese weniger auf den musikalischen Stil, sondern mehr auf die Funktion der Musik beziehen. Rock und Alternative werden aufgrund ihrer musikalischen Ähnlichkeit zu Rock zusammengefasst. Ebenso werden Soul, Funk & Blues vereinheitlicht. Folk und Country werden aufgrund ihres stark lokal abhängigen Aufkommens und ihrer musikalischen Ähnlichkeit mit dem Blues nicht eigenständig aufgeführt. Die Ergebnisse der Untersuchungen können aber mit Einschränkungen von der Kategorie Soul, Funk & Blues auf Folk und Country übertragen werden. Gleichermaßen können aus dem Pop hervorgehende Erkenntnisse unter Vorbehalt auf Rap & Hip-Hop übertragen werden, da die Musikrichtungen häufig starke Gemeinsamkeiten in den Produktionspraktiken aufzeigen. Darüber hinaus wird Heavy Metal durch Metal ersetzt, da Heavy Metal als eine Stilrichtung des Metal gesehen werden kann [26]. Auch Electronica & Dance wird in das allgemein gebräuchliche EDM, also Electronic Dance Music (dt. elektronische Tanzmusik), umbenannt.

Somit stehen für den Hörversuch folgende sieben Musikrichtungen fest:

Klassik – Jazz – Soul, Funk & Blues – Rock – Metal – Pop – EDM

3.2 Produktion der Hörbeispiele

Im nächsten Schritt wird jeder Musikrichtung ein möglichst charakteristischer Musiktitel zugewiesen (siehe Tabelle 2).

	Künstler	Musiktitel	Musikrichtung	übertragbar auf	Format
1	Festival Strings Lucerne	PC-24, ein musi-	Klassik		9.0
		kalischer Alpenflug			
2	Tobias Becker Bigband	Pennsylvania	Jazz		9.1
	& Jochen Neuffer	6-5000			
3	Yannik Weber	Love On Top	Soul, Funk & Blues	Folk; Country	10.1
	& Lea Reister				
4	OXMO	Rallye	Rock		9.1
5	Jonas Müller	Creeping Decay	Metal		10.1
6	Lea Reister	Change the World	Рор	Rap & Hip-Hop	9.1
7	Yannik Weber	Rain On Your Skin	EDM		10.1 [9.1]

Tabelle 2: für den Hörversuch verwendete Musiktitel; []: Format des Hörversuchsausschnitts

Grundlegend ist die Bestrebung, sowohl klanglich als auch musikalisch ein einheitliches, hohes Niveau zwischen den Stücken zu wahren. Zudem werden Stücke ausgewählt, die für die Mischung in 3D genügend Optionen, also eine gewisse Spielwiese, bieten. Natürlich lassen sich die Stücke nicht uneingeschränkt auf das gesamte Genre und seine Sub-Genres übertragen. Dennoch wird mit der Auswahl versucht, die Musikrichtungen möglichst exemplarisch zu treffen.

Die Stereo- und 3D-Mischungen werden unabhängig voneinander erstellt. Ziel ist es, diese nicht strikt aneinander anzugleichen, sondern sich bei den Stereo-Mischungen auch klanglich am aktuellen Markt zu orientieren. Das 3D-Format wird nach Bedarf zwischen 9.0, 9.1 und 10.1 (inkl. VOG) gewählt.

Aus Gründen der Machbarkeit und Belastbarkeit der Probanden wird pro Stück ein einminütiger Ausschnitt gewählt.

Im Folgenden werden die Stücke, Produktionsmethoden und Mischansätze, vorgestellt. Die Produktionsbeschreibungen sind ausdrücklich als Erfahrungsbericht zu sehen. In Anbetracht des Umfangs dieser Arbeit, wird das Augenmerk hierbei auf den vorgespielten Ausschnitt gelegt. Dennoch ist es das Ziel, auch die Gesamtheit der Produktionen mit wenigen Worten zu beleuchten.

3.2.1 Klassik: PC-24, ein musikalischer Alpenflug

Musikrichtung	Klassik
Titel	PC-24, ein musikalischer Alpenflug
Komposition	Stephan Hodel
Dirigent	Daniel Dodds
Interpret	Festival Strings Lucerne
Aufnahme	Lasse Nipkow
Stereo-Mix	Lasse Nipkow
3D-Mix	Lasse Nipkow
Ausschnitt	0:00 - 1:00

Bei *PC-24, ein musikalischer Alpenflug,* dem Hörbeispiel des Genres Klassik, handelt es sich sich um das Werk des Schweizer Komponisten *Stephan Hodel.* Initiator des Projektes war *Daniel Dodds,* Leiter der *Festival Strings Lucerne*. Er war von der Schönheit und Eleganz des Düsenjets *PC-24* der *Pilatus Flugzeugwerke*

inspiriert und fand mit *Hodel* einen für seine Vision geeigneten Komponisten. "Im ersten Teil etwa, wo das Flugzeug quasi produziert wird, steht der musikalische Drive für die Energie und den lang anhaltenden Schwung, den ein solches Projekt erfordert." führt Dodd aus. [27] Durch die romantisch, sehnsüchtig und beflügelnd erscheinende musikalische Darstellung gleich zu Beginn, bietet sich die erste Minute des Stückes auch als Ausschnitt für den Hörversuch an. Gleichzeitig entfaltet das Orchester an dieser Stelle seine volle Klangfülle.

Das Stück wurde am 17. Juni 2018 im *Konzertsaal* des *Kultur- und Kongresszentrums Luzern* durch die *Festival Strings Lucerne* in mittlerer sinfonischer Besetzung uraufgeführt [27]. Die am selben Abend entstandenen Aufnahmen hat *Lasse Nipkow*, Gründer und Geschäftsführer der *Silent Work GmbH*, durchgeführt und freundlicherweise für diese Arbeit zur Verfügung gestellt. *Nipkow* zählt zu den Pionieren im Bereich 3D-Audio und publiziert seine Erkenntnisse zu 3D-Audio und der damit einhergehenden angewandten Psychoakustik regelmäßig auf internationalen Tagungen und in Papers [28].

Das Hauptmikrofonsystem bildet bei dieser Produktion klassisch die Basis der Orchesteraufnahme. Das Kugelflächenmikrofon *Schoeps KFM 6* wurde hierbei mit zwei zur Bühne gerichteten *Schoeps CCM 41V* kombiniert. Als Stützen dienten diverse Klein-AB-Stereo-Paare. Über einen *Illusonic LCR-Upmix* wurde zur Wiedergabe der somit bestehenden Stereo-Grundlage zusätzlich der Center miteinbezogen. Zudem fand mit besagtem Material ein *Illusonic Room Sound-Upmix* auf die umgebenden Lautsprecher statt. Dies wird von den beiden, der Bühne abgewandten Raummikrofonen *Schoeps CCM 4* hinter dem *KFM 6* [LS-RS] und den, an den entsprechenden Stellen im Saal angebrachten Kugelmikrofonen im Groß-AB [HL-HR, HLS-HRS], unterstützt. Der LFE wurde bei dieser Produktion nicht verwendet.

Ziel der Produktion war es, den Höreindruck im Saal über die 3D-Aufnahme abzubilden.

Funktion	Mikrofon	Richtcharakteristik	Position	Processing	
Haupt-	Schoeps KFM 6	Kugel	L-R	Illusonic	Illusonic Room
mikrofone	2 x Schoeps CCM 41V	Superniere		LCR-	Sound Upmix
Stütz-	Stützung über versch.	divers		Upmix	[auf LS-RS, HL-
mikrofone	Stereo-Paare (Klein-AB)			[auf L-C-R]	HR, HLS-HRS]
Raum-	2 x Schoeps CCM 4	Niere	LS-RS		
mikrofone	hinter KFM 6				
	diverse (Groß-AB)	Kugel	HL-HR,		
			HLS-HRS		

Tabelle 3: Mikrofonierung von PC-24

3.2.2 Jazz: Pennsylvania 6-5000

Musikrichtung	Jazz
Titel	Pennsylvania 6-5000
Komposition	Jerry Gray
Arrangement	Jochen Neuffer
Dirigent	Jochen Neuffer
Interpret	Tobias Becker Bigband
Produktion	Jochen Neuffer & Henk Heijink
Aufnahme	Philipp Heck
Stereo-Mix	Philipp Heck
3D-Mix	Yannik Weber
Ausschnitt	1:28 – 2:28

Während meines Praxissemesters bei den Bauer Studios in Ludwigsburg durfte ich bei der Produktion des Albums Augmented Reality von Jochen Neuffer und der Tobias Becker Bigband assistieren. Schon im Verlauf der Aufnahme kristallisierte sich die "erfrischend moderne Umsetzung des Klassikers «Pennsylvania 6-5000»"
[29] als persönlicher Lieblingstitel des Albums

heraus. Ein Stück, das sich schon durch sein dynamisches Arrangement, seine klangliche Vielfalt und einzigartig intime und mystische Stimmung, für eine 3D-Mischung anbietet. Freundlicherweise stellten mir die *Bauer Studios* und *Jochen Neuffer* die einzelnen Spuren für den 3D-Mix zur Verfügung. Als Stereo-Version wurde die auf CD veröffentlichte Mischung von *Philipp Heck* verwendet. Der für den Hörversuch ausgewählte Ausschnitt erstreckt sich vom ersten B-Teil über den A-Teil bis zum Beginn des Saxophon-Solos.

Die grundlegende Idee der 3D-Mischung von *Pennsylvania 6-5000* war, den Fokus und die Basis wie auch bei den anderen Stücken in der vorderen Ebene zu belassen und durch entsprechende Positionierung und Auffächerung der Instrumente ein transparentes, aber dennoch nicht auseinanderfallendes Klangbild zu erhalten (siehe Abbildung 19).

Somit wurde zu Beginn der Mischung das Schlagzeug in der vorderen Ebene untergebracht. Die Overheads fanden Platz auf HL und HR, eine Etage tiefer wurden die Toms entsprechend im Stereo-Panorama verteilt. Snare und Bassdrum wurden, wie der E-Bass, dem Center zugeordnet. Die beiden letztgenannten Spuren zudem dem LFE. Während der Flügel direkt in Stereo aufgenommen wurde [L-R], wurden die beiden Monoquellen Rhodes-Piano [HL-HR] und elektrische Jazz-Gitarre [LS-RS] über einen dezenten Stereo-Detune-Effekt verbreitert. Die Verbreiterung des Signals hat hierbei eine angenehme Wirkung. Beim Einschalten des Effektes lösen sich die Phantomschallquellen eindrücklich

zwischen den Lautsprecherpaaren auf, und die flächigen Sounds verlieren ihren Fokus, was den recht direkten Bläsern in die Karten spielt.

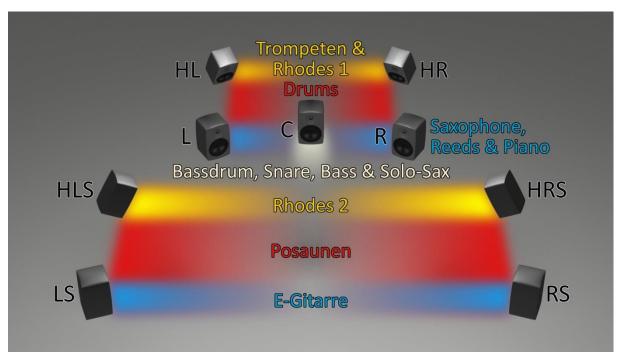


Abbildung 19: Verteilung der Instrumente im Jazz-Beispiel (Grafik: J. Frey & Y. Weber)

Die untere Basis derer bilden hierbei die Saxophone und Reeds [L-R]. Eine Ebene höher wurden die Trompeten verteilt [HL-HR]. Bei der Verteilung dieser Bläsersätze wurde hartes LCR-Panning angewandt, eine Methode, bei der Mono-Elemente ohne Zwischenschritte entweder extrem einer Seite oder der Mitte zugewiesen werden. Der Center wurde bei den Bläsern einzig für das solierende Saxophon verwendet. Das Arrangement und der unaufdringliche Klang der Posaunen boten bei diesem Stück an, die vier Spuren je einem der hinteren Lautsprecher zuzuordnen [LS, RS, HLS, HRS]. Gerade im letzten Aufbau des Stücks unterstützt die Beschallung von hinten die etwas bedrückende, aber gleichermaßen erhabene Stimmung des Titels. An einigen Stellen des Stücks entsprechen die beiden verschiedenen Stimmen der Tenor-Saxophone denen der Posaunen. Eine Umverteilung der gleichstimmigen Instrumente auf die jeweils gegenüberliegende Seite bewirkte dagegen einen deutlich ausgewogeneren Klang.

Über vier H-Reverb-Stereo-Instanzen, mit leicht differierenden, an der Realität orientierten Setups, wurde ein natürlicher, umhüllender Raumklang generiert.

Zwei entscheidende Dinge sind mir beim Mischen des Stücks in 3D und dem Vergleich zur Stereo-Version klar geworden. Zum einen kann in 3D klanglich viel natürlicher gearbeitet werden, da sich die verschiedenen Elemente durch ihre räumliche Verteilung nicht gegenseitig den Platz im Frequenzspektrum wegnehmen. Ich kann beispielsweise im Solo-Modus auf einen ästhetischen Klang des Flügels hinarbeiten und diesen dann an die gewünschte Position setzen, ohne dass er im Gesamtklang untergeht. In Stereo überlagert sich der Klang eines Instruments dagegen sehr oft stark mit anderen Instrumenten. Übertragen auf dieses Beispiel wird eventuell der Gesamtklang zu mittig oder es muss, wenn eine differenzierte Darstellung der Instrumente gewünscht ist, mit entsprechender Filterung nachgeholfen werden. Zum anderen gilt selbiges für die Dynamik der Stücke. Es ist sowohl weniger Kompression als auch Automation nötig, um einzelne Instrumente sauber wahrnehmen zu können. Der *Cocktailparty-Effekt* ist eine mögliche Erklärung dafür. Das Gehör kann auch bei vielen gleichzeitig ablaufenden Schallereignissen in gleicher Lautstärke je nach Position, also Richtung der Schallquelle, Reize herausfiltern und wahrnehmen [1, p. 137]. Es kann also je nach Interesse seine eigene Automation fahren. Die Möglichkeit, die Aufmerksamkeit des Hörers durch Automation auf bestimmte Elemente zu leiten, besteht natürlich auch im 3D-Mix. Bei dem Vergleich beider Versionen fällt die wesentlich höhere Lautstärke der Soli in Stereo auf.

3.2.3 Soul, Funk & Blues: Love On Top

Musikrichtung	Soul, Funk & Blues
Titel	Love On Top
Songwriting	Yannik Weber
Interpret	Yannik Weber & Lea Reister
Produktion	Yannik Weber
Aufnahme	Yannik Weber
Stereo-Mix	Yannik Weber
3D-Mix	Yannik Weber
Ausschnitt	2:48 – 3:48

Bei Love On Top handelt es sich um einen selbstgeschriebenen, -eingespielten und -produzierten Song, der dem Soul, Funk & Blues zugeordnet werden kann. Großer Dank geht an meine Mitmusiker Lea Reister (Gesang), Jan Bieletzki (Saxophon) und Tobias Kroner (Trompete).

In dem im Sommersemester 2018 an der *HdM* durchgeführten, eigens initiierten Projekt *3Disco* ging es darum, tanzbare Musik in 3D-Audio zu produzieren. Nach der Produktion des elektronischen Titels *Rain On Your Skin* (siehe Kapitel 3.2.7) war es das Ziel, einen zweiten Titel zu produzieren, der mit Groove und Instrumentierung an die soulbehafteten Disco-Songs der 70er erinnern sollte. Textlich beschreibt der Song - getarnt als gewöhnliches Liebeslied - die Beziehung zwischen dem Tonschaffenden (männliche Stimme) und dem Surround- bzw. 3D-Format (weibliche Stimme).

Der gesamte Titel wurde in *Ableton Live 10* produziert. Als Audiointerface diente ein *Steinberg UR44*. Über *Addictive Drums 2*, ein Drumsampler des Herstellers *XLN Audio*, wurde die Schlagzeugspur für das Stück programmiert. Von Bass und Gitarre wurde das DI-Signal aufgezeichnet und entsprechende Verstärkersimulationen über *Native Instruments Guitar Rig 5* realisiert. Über den *Sampler Kontakt 5* des gleichen Herstellers wurden zudem ein Rhodes Piano und eine Hammond Orgel eingespielt. Der in *Ableton* integrierte Sampler wurde für die Erzeugung der für die Ära üblichen Streichersounds

verwendet. Das Saxophon wurde über ein *AKG C 414 B-ULS*, die Trompete über ein *Neumann TLM 170 R* abgenommen. Für den Gesang wurde ein *Rode NT2-A* eingesetzt.

Bei *Love On Top* wurde das Schlagzeug klassisch in der unteren Ebene gemischt [L-R], dessen kurzer Nachhall (Predelay: 20 ms, Decay: 600 ms, 100% wet) dagegen über HL und HR wiedergegeben wird. Dadurch geht der Klang subtil und natürlich in die Höhe auf, ohne indirekter zu wirken. Gitarren und Bass bilden das Fundament des Stückes in der vorderen Ebene, bis die synkopierte Bläserlinie, eine Hammond und Streicher, das Klangbild im Refrain öffnen. Für die Main Vocals wurde bis zur Bridge der Center verwendet. Dort teilen sich die beiden Stimmen dann aber auf L und R auf (Ausschnitt des Hörbeispiels).

Thematisch finden an dieser Stelle männliche und weibliche Stimme zusammen. Auch im Mix kommen sich beide Stimmen näher und wandern dabei nach oben, bevor sie im letzten Refrain "die Liebe obendrauf setzen". Unterstützt wird die besondere Stimmung durch Tiefpassfilterung der Rhythmusgruppe und die lange Vocal-Hallfahne, die an der Decke zu schwimmen scheint. Auch die Backing Vocals konnten unaufdringlich im Height-Layer verteilt werden. Schließlich löst sich der zuvor nur umspielte Akkord des Refrains zu einer komplettierenden Akkordfolge auf. Wie bereits oben angeschnitten, werden die textlichen Metaphern des Stückes über das Panning im Raum unterstützt. So dreht sich der Nachhall von "Keep on spinning me around" dezent um den Hörer. Auf "Surround me as before" folgt ein über die Surrounds wiedergegebenes "uh uh uh".

Neben der Hall-Nachbildung eines kleinen Raumes wie im Jazz-Beispiel (siehe Kapitel 3.2.2), wurde das Format zudem zur Erzeugung eines räumlichen Slapback-Delays für den Gesang verwendet. Der Slapback-Effekt auf den Surrounds [LS-RS] ist um 40 ms zu dem auf den vorderen Lautsprechern [L-R] verzögert. Dies öffnet nicht nur den Raum, sondern gibt dem Titel auch einen leichten Vintage-Beigeschmack.

Nach Fertigstellung der 3D-Mischung wurde die Stereo-Mischung angegangen.

3.2.4 Rock: Rallye

Musikrichtung	Rock
Titel	Rallye
Songwriting	Fabian Moser
Interpret	OXMO
Produktion	Yannik Weber
Aufnahme	Yannik Weber
Stereo-Mix	Yannik Weber
3D-Mix	Yannik Weber
Ausschnitt	0:20 - 1:20

Der Song *Rallye* der Band *OXMO* wurde als Repräsentant der Musikrichtung Rock gewählt. Durch die Tatsache, selbst als einer der beiden Gitarristen Bandmitglied zu sein, und der Mitarbeit am Stück als Musiker, Produzent und Engineer, stellte mir die Band den Ende 2018 aufgenommenen Titel gerne für meine Arbeit zur Verfügung. Als Besonderheit des Punkrock-Songs kann sowohl die

Trompete, als auch der Sprechgesang genannt werden. Die Kommentarfelder des Hörversuchs zeigen eine Assoziation des Titels mit der Band *Kraftklub*. Inhaltlich besingt der Sänger die Fahnenschwingerin eines Autorennens in der Ich-Perspektive und will diese mit seinem Fahrrad von ihrem Leid befreien. Der Ausschnitt des Rock-Hörbeispiels beginnt kurz vor Einsatz der ersten Strophe und erstreckt sich über Refrain und Zwischenspiel in die zweite Strophe hinein.

Wie bei vielen aktuellen Produktionen üblich, wurde auch hier die Schlagzeugspur im Vorfeld über MIDI programmiert und vor dem Start der Aufnahmen mithilfe von Addictive Drums 2 (siehe Kapitel 3.2.3) in die Ableton-Session integriert. Zur Aufnahme der Audiosignale wurde ein RME Fireface UC verwendet. Die Bass- und Gitarrenspuren des Titels wurden nacheinander direkt über einen Kemper Profiling Amp, ein Gerät zur Klangsimulation analoger Verstärker, eingespielt. Gesang und Trompete wurden mit einem Shure SM 7 B abgenommen.

Der ursprünglich für Stereo produzierte Titel sollte mit verzerrtem Gitarrensound ohne großartige Effekte, schlichter Instrumentierung und einem etwas raueren Klang, eine entschiedene Energie und Härte übermitteln. Mit recht strikt durchgezogenem LCR-Panning der Gitarren und des Gesangs orientierte man sich an charakteristischen Produktionen. Zudem wurden die Schlagzeug-Gruppe und ihr Raum für einen punchigen, lauten Sound stark komprimiert. Generell wurde mit Dopplungen gearbeitet, eine viel verwendete Methode, bei der beispielsweise die gleichen Gitarrenstimmen mehrmals eingespielt werden. Durch extremes Panning auf die beiden Seiten, wird der Sound durch die verschiedenen Signale unheimlich breit, fällt aber durch den zusammenhängenden Klang nicht auseinander. Das Verfahren wurde ebenso bei den Backing Vocals eingesetzt.

Beim 3D-Mix wurde das Schlagzeug des Titels mit der bereits beim Jazz-Beispiel (siehe Kapitel 3.2.2) angewandten Vorgehensweise gemischt. Zusätzlich sollte aber die beim Stereo-Mix auf die Drum-Gruppe angewandte Kompression übernommen werden. Da Effekte auf der Gruppe aber bei direkter Verteilung der Einzelspuren auf die entsprechenden Kanäle keine Wirkung haben, wurde die Gruppe dupliziert. Jede Einzelspur der Drums wurde mit dem ursprünglichen Bus-Kompressor belegt, als dessen Steuersignal nun die duplizierte Drum-Gruppe eingestellt wurde.

Eine Besonderheit des 3D-Mixes stellt mit Sicherheit die Mischung der Gitarren dar. Die gedoppelte Hauptgitarre [L-R] wurde dupliziert und mit den frühen Reflexionen eines Hall-Plugins versehen. Der veränderte, nicht merklich räumlichere aber weniger korrelierte Sound, wurde dann den vorderen Höhenlautsprechern [HL-HR] zugewiesen. Der Effekt wird lediglich in den Refrains und der Bridge eingesetzt, wenn wortwörtlich die berühmte *Wall Of Guitars* gewünscht ist. In diesen Teilen spielt

zudem die zweite, ebenfalls gedoppelte Gitarre [LS-RS] in einer etwas anderen Lage, auf die dieselbe Methode mit den entsprechenden Höhenlautsprechern [HLS-HRS] angewendet wurde.

Ein ähnlicher Effekt wurde auf die beiden verschiedenstimmigen Trompeten [HL-HR] angewandt. Für die Wiedergabe ihres Slapback-Delays (100% wet) wurden allerdings die horizontal benachbarten Lautsprecher gewählt [HLS-HRS]. An einigen Stellen wurde das Format gleichermaßen als Effekt verwendet. So kreist das anschwellende Feedback einer Gitarre vor ihrem eigentlichen Part um den Kopf des Hörers. Im energetischen Zwischenteil des Stücks fährt zudem der Reverse-Reverb-Effekt der Stimme auf dem Surround-Layer von hinten nach vorne. Dieser kann durch das Einfrieren und rückwärtige Abspielen eines Hall-Effekts generiert werden. Im folgenden *Comedian Harmonists*-Teil vor dem letzten Refrain wird dagegen das ganze Signal mit dem Klang eines Grammophons aus dem Center wiedergegeben. Dies löst sich über den Sound einer verrutschenden Plattennadel auf, der pro Lautsprecherpaar unterschiedlich gepitcht aus allen Richtungen [L-R, LS-RS, HL-HR, HLS-HRS] kommt. Wie bei der Jazz-Produktion (siehe Kapitel 3.2.2) beschrieben, wurde aus Stereo-Hallinstanzen ein kleiner Raum generiert.

3.2.5 Metal: Creeping Decay

Musikrichtung	Metal
Titel	Creeping Decay
Songwriting	Jonas Müller
Interpret	Jonas Müller
Produktion	Jonas Müller
Aufnahme	Jonas Müller
Stereo-Mix	Jonas Müller
3D-Mix	Yannik Weber
Ausschnitt	3:27 – 4:27

Bei *Creeping Decay* handelt es sich um eine Metal-Produktion meines Kommilitonen *Jonas Müller*, der sich in seiner Bachelorarbeit mit der Produktion seines progressiven Metal-Konzeptalbums *Moebius* beschäftigt. Dies handelt von einer Person mit dissoziativer Identitätsstörung, was in dem für den Hörversuch ausgewählten Titel (zweiter Titel des Albums) zum ersten Mal eingeführt und dargelegt wird. Der Ausschnitt des

Hörversuchs startet wenige Sekunden vor der unten beschrieben Soundeffekt-Stelle.

Jonas Müller produzierte den Titel in der DAW Cubase des Herstellers Steinberg und verwendete zur Programmierung der Drums Toontracks Superior Drummer 3. BIAS Amp 2 von Positive Grid wurde für die Verstärker-Simulation der von ihm selbst eingespielten Gitarren- und Bass-Spuren verwendet. Neben Müllers gutturalem Gesang (Aufnahme mit Shure SM 7 B) durfte ich die "konventionellen" Gesangsparts (Aufnahme mit Rode NT2-A) übernehmen. Durch den ausproduzierten, vollen Klang mit vielen kleineren Soundeffekten und Spielereien bietet sich Jonas Müllers Sound bestens für die Mischung in 3D an. Nach Fertigstellung des Stereo-Mixes stellte mir Müller diesen und die vorgemischten Einzelspuren, inklusive separater Effekt-Returns, zur Verfügung.

Als zentrales Element des Titels können die verzerrten E-Gitarren gesehen werden. Das Fundament bildet die gedoppelte Rhythmusgitarre. Ziel war es, diese so mächtig und bedrückend wie möglich klingen zu lassen, also den Hörer möglichst in den Mittelpunkt des Hörereignisses zu stellen. Allerdings war, wie eingangs erwähnt, nur eine einzelne, gedoppelte Rhythmusgitarre vorhanden. Abhilfe schuf abermals das Duplizieren der Stereo-Spur. Daraufhin wurde diese um 10 ms verzögert und mit dem *Ableton-*Effekt *Frequency Shifter* eine Frequenzverschiebung von -4 Hz hervorgerufen.

Bei gewöhnlichem Pitch-Shifting werden die Frequenzen des Signals mit einem bestimmten Wert multipliziert. Durch einen Faktor von zwei klingt das Signal eine Oktave höher und die harmonischen Relationen werden beibehalten. *Frequency Shifter* verschiebt dagegen die Frequenzen des Signals durch Addition oder Subtraktion eines Wertes in Hz, wodurch sich die harmonischen Relationen verändern. [30, pp. 365-367]

Der bei starkem Einsatz metallisch klingende Effekt ist im A-B-Vergleich klanglich auf den Gitarren kaum merkbar und zeigt in Kombination mit der zeitlichen Verzögerung große Wirkung. Die auf HL-HR gepannte, originale Stereo-Spur, wirkt in Zusammenspiel mit der eben prozessierten Spur auf HLS-HRS extrem breit. Noch mächtiger wird der Klangeindruck durch die bereits beim Rock-Beispiel angewandte Methode – in diesem Beispiel allerdings in umgekehrter Form. Das Signal der oberen Lautsprecher speist das Hall-Plugin der jeweils darunterliegenden Lautsprecher, und dessen frühe Reflexionen ziehen den Klangeindruck zusätzlich in die Höhe auf. Weitere Lead-Gitarren wurden nach Geschmack auf den Surrounds verteilt. In dem Height-Layer fand der flächige Synthesizer Platz. Generell fällt auf, dass sich gerade flächige Elemente wie Pads, aber beispielsweise auch das Rhodes-Piano in der Jazz-Produktion, unheimlich unaufdringlich in dem Height-Layer positionieren lassen. Sie lassen das Klangbild groß wirken, ohne zu viel Aufmerksamkeit auf sich zu ziehen.

Die bereits erwähnten Soundeffekte spielen im Ausschnitt des Hörversuchs eine prägnante Rolle. Im Zwischenteil brechen alle Instrumente weg, lediglich Gesang und Effekte sind zu hören. Die besagte FX-Spur beinhaltet eine Leadgitarre und Snare-Fills, die von einem Schallplattenknistern überlagert werden. Im Verlauf des etwa 8 Sekunden langen Abschnitts öffnet sich ein Tiefpass-Filter auf der Spur. Hierbei bot es sich an, die FX-Spur lediglich auf die ansonsten selten verwendete VOG zu legen. Konträr zum zuvor so mächtigen Klangbild, geht die klangliche Ausdehnung auf einen Schlag extrem zurück. Das Ende des Swooshes (anschwellendes Rauschen) am Ende des Abschnitts leitet die Überblendung der Spur von VOG hin zu HL-HR ein. Dies bildet einen runden Übergang zu den danach mit voller Energie wiedereinsetzenden Gitarren. Die von *Müller* verwendeten sehr tieffrequenten Booms (Bass-Effektschläge) bieten zudem die Ansteuerung des LFEs an.

Die Stimmen wurden ihrer angestrebten Wirkung entsprechend verteilt. Die gescreamte Hauptstimme mittig zwischen HL und HR, die gedoppelten Backings je links und rechts daneben. Zudem wurden die Backing Vocals dupliziert [HLS-HRS], mit dem oben erwähnten Frequency Shifter um 200 Hz abgesenkt und anschließend zusätzlich mit frühen Reflexionen versehen. Der Fokus der Stimme bleibt durch die zeitliche Verzögerung in der Front, ein weiteres Mal geht der Klang aber in der Raumtiefe auf, und durch die Frequenzverschiebung entsteht im Gesamtklang ein subtiler ,Monster-Effekt'. Die ,konventionell' gesungene Hauptstimme fand Platz auf dem Center. Die zugehörigen Backings wurden den Abschnitten entsprechend auf L-R, HL-HR und HLS-HRS verteilt. Der Stereo-Stem der vorgemischten Drums beinhaltete bereits einen beachtlichen Hall-Anteil. Da der Nachhall der Drums nicht noch größer werden sollte, diese aber dennoch räumlich wirken sollten, wurden dementsprechend leicht zueinander veränderte Hall-Instanzen mit einem Decay von 1 s auf die L-R umgebenden Lautsprecher [HL-HR, LS-RS, HLS-HRS] gelegt. Auch die weiteren Hall- und FX-Returns sollten genutzt werden. Die Verdopplung einer Spur und Verzögerung dieser um 30 ms funktionierte auf derartig diffusem, indirektem Material, enorm gut. Spannend klang ebenso die etwas unkonventionelle Methode, den Nachhall des Gescreamten nur auf dem Height-Layer und der des Gesungenen nur auf dem Surround-Layer wiederzugeben. Ein Effekt, der beim Hören der gesamten Mischung nicht wirklich heraussticht, dennoch aber, wie so oft unterbewusst, seine Wirkung zeigt.

3.2.6 Pop: Change The World

Musikrichtung	Pop
Titel	Change The World
Songwriting	Lea Reister
Interpret	Lea Reister
Produktion	Yannik Weber & Lea Reister
Aufnahme	Yannik Weber
Stereo-Mix	Yannik Weber
3D-Mix	Yannik Weber
Ausschnitt	1:34 – 2:34

Change The World ist ein Popsong der Singer-Songwriterin Lea Reister. Durch eigene Mitarbeit am Song als Produzent, stellte sie mir den Titel für die Arbeit gerne zur Verfügung. Das Stück Change The World erzählt von den Missständen in der Welt, kontrastiert diese mit den Unternehmungen wohlhabenderer Gesellschaften und ruft dazu auf,

die Welt zu verändern. Für den Hörversuch wurde ein Ausschnitt von der zweiten Strophe bis zum Outro gewählt.

Der in *Ableton* produzierte Stereo-Titel sollte mit Beat und Groove an Hip-Hop erinnern und dennoch, nach den eher gefühlsbetonten Strophen, im Refrain eingängig und poptypisch aufgehen. Der auf L-R verteilte Beat wurde auch hier durch einen über HL-HR wiedergegebenen, Deckenreflektionen simulierenden Nachhall, erweitert. In Zusammenspiel mit den verschiedenen, auf L-R verteilten Basselementen, stellen die beiden Komponenten das Fundament des Stücks dar. Im

Refrain wird eines der Basselemente zudem durch einen breiten Staccato-Bläsersatz unterstützt. Durch dessen Anordnung in der Höhe [HL-HR] und Verbreitung auf HLS-HRS über die abermals erwähnte Early Reflections-Methode, geht das Stück im Refrain luftig in die Höhe auf. Ein darauffolgender Synthesizer auf den zweiten Schlag des Taktes steht in musikalischer Verbindung zum Bläsersatz. Dank der vorausgegangenen Bläser mit Fokus in der Front, wird trotz extremen Pannings [LS-RS] das Gleichgewicht der Mischung gewahrt.

Generell kommen im Refrain zu den ohnehin schon zahlreichen Elementen noch einige weitere, füllende Spuren hinzu. Dies lud bei der Verteilung im 3D-Raum zum unbefangenen, intuitiven Experimentieren ein. Bei nachträglicher Betrachtung fällt auch angesichts anderer hier aufgeführter Produktionen auf, dass perkussive Elemente eher in der Front, dort aber gerne auch in der Höhe ihren Platz fanden. Flächigere Elemente konnten für gewöhnlich frei verteilt werden, um das Klangbild voll und breit wirken zu lassen. Aber auch diese brauchen wie so oft ihren Gegenspieler im vorderen Bereich. Ist dieser vorhanden, sind auch bei perkussiveren Elementen Pannings im hinteren Teil des Raumes möglich. Hierbei spielt ein weiteres Mal das Arrangement des Stückes eine sehr große Rolle.

Die Stimme wurde klassisch auf den Center gelegt. Weitere stimmliche Elemente wie Backing Vocals oder Verzierungen wurden auf L-R bzw. HL-HR verteilt. Der Delay-Effekt der Backing Vocals im Refrain [L-R] wurde der darüber liegenden Ebene [HL-HR] zugewiesen. So konnte auch ohne Panning-Automation Bewegung in den Titel kommen. An passenden Stellen wurde diese dennoch angewandt. So boten sich beispielsweise anschwellende Swooshes für Fahrten durch den Raum an. Das einleitende Piano zu Beginn [L-R] wird nach zwei Takten durch die Kindheitsassoziationen hervorrufende Spieluhr [HL-HR] untermalt, die ein zweites Mal am Ende des Titels auftaucht und diesen abschließt. Die Hallerzeugung erfolgte auch hier über verschiedene Stereo-Instanzen.

3.2.7 EDM: Rain On Your Skin

Musikrichtung	EDM
Titel	Rain On Your Skin
Songwriting	Yannik Weber
Interpret	Yannik Weber
Produktion	Yannik Weber
Aufnahme	Yannik Weber
Stereo-Mix	Yannik Weber
3D-Mix	Yannik Weber
Ausschnitt	2:06 – 3:06

Der elektronische, tanzbare Titel *Rain On Your Skin* entstand wie *Love On Top* (siehe Kapitel 3.2.3) ebenfalls im Projekt *3Disco*. Die Neugier, Klang und Möglichkeiten elektronischer Tanzmusik in 3D-Audio auszureizen, war die ursprüngliche Motivation und Initialzündung zum Projekt, das ich im Frühjahr 2018 mit einigen Kommilitonen startete (siehe Abbildung 20). Jeder Student nahm sich dabei der Produktion

eines tanzbaren Titels (anfangs nur elektronische Tanzmusik) in 3D an. Nach dem Erfolg von Rain On

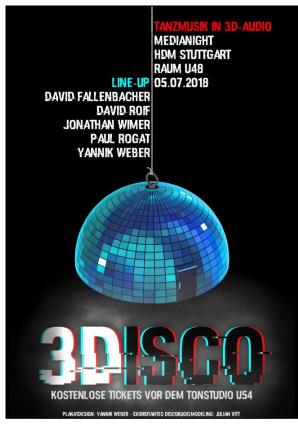


Abbildung 20: Plakat von 3Disco (Grafik: J. Vitt & Y. Weber)

Your Skin in der 2nd Student 3D Audio Production Competition des IEMs Graz auf der Tonmeistertagung 2018 (Gold in der Kategorie Music Recording/Studio Production), entwickelte sich die Leitfrage hin zu: "Wie kann 3D in der Musikproduktion mit einem Mehrwert für verschiedene Genres genutzt werden?" Somit war diese Produktion auch gewissermaßen Auslöser dieser Arbeit.

Ziel der Produktion war es, die Möglichkeiten von 3D-Audio im Stile der aktuellen elektronischen Tanzmusik mit den klassischen Build-Ups (Aufbau und Steigerung zum Drop hin), Drops (charakteristischer Moment bevor der Beat einsetzt) und Breakdowns (oft minimalistischer Zwischenteil) auszureizen. Der energetische House-Track sollte das Potenzial des 3D-Formats

nach einem recht schmal gehaltenen, aus den Clubs bekannten Stereo-Build-Up zu Beginn aufzeigen. Für den Hörversuch wurde allerdings ein Ausschnitt im Mittelteil gewählt. Der Beginn des Stücks stellt durch den musikalisch umgesetzten Wechsel von Stereo auf 3D einen direkten Vergleich beider Formate dar. Dies ist zwar zum einen ein sehr wirksames Stillmittel, würde aber zum anderen durch ebendiesen Überraschungseffekt möglicherweise eine Beeinflussung des Ergebnisses bedeuten.

Das Stück wurde in *Ableton* von Anfang an in 3D produziert. Als Format wurde *Auro-13.1* [13, p. 14] verwendet, das zusätzlich über die beiden Front-Surrounds LFS und RFS zu einem inoffiziellen 15.1-Setup erweitert wurde. Um kompatibel und vergleichbar mit den anderen Hörbeispielen zu bleiben, fand nachträglich ein manueller 10.1 Downmix statt. Bei diesem wurde versucht, die Positionen der Elemente wenn möglich über Phantomschallquellen abzubilden. Elemente auf dem HC wurden dafür beispielsweise mit gleicher Intensität auf HL und HR verteilt und im Pegel um 3 dB abgesenkt [1, p. 807]. Klänge auf den Rear-Surrounds wurden den Surrounds zugewiesen.

Zur Klangerzeugung dienten der Software-Synthesizer *Serum, Abletons* integrierte Synthesizer und Sampler, und der Hardware-Synthesizer *Korg Minilogue*. Über ein *Steinberg UR44* und ein *Rode NT2-A* wurden der Gesang, sowie einige perkussive Elemente wie Shaker, Bongos und Schellenkranz, aufgenommen. Wie oben erwähnt, startet das Stück in Stereo. Im Drop werden dann die

umgebenden Lautsprecher sinnbildlich mit dementsprechenden Sounds eingeschaltet und eine sich aufbauende Fläche breitet sich von den Lautsprechern L und R auf die umgebenden aus. Dies wurde über eine Automation der Sends erreicht.

Der ursprüngliche Startpunkt der Produktion war allerdings der knackige, raue Basssound des Synthesizers *Serum*, dessen Subbassanteile auf den LFE gelegt wurden. Der höhenreiche Anteil des Sounds sollte dagegen so mächtig wie möglich klingen und den Hörer komplett umgeben. Um dies zu erreichen, sind Signale mit einer Korrelation um Null auf den umgebenden Kanälen erforderlich. Deshalb wurde die Stereo-Spur einige Male dupliziert und Synthesizer-Parameter wie der Grad der Verzerrung, die Grenzfrequenz eines Filters, und die Wellenposition des Wavetable-Synthesizers, auf den verschiedenen Spuren unterschiedlich automatisiert. Auf der Höhenebene wurde zusätzlich eine Quinte hinzugefügt.

Ebenso erwähnenswert sind die gerne als Stabs bezeichneten, kurzen Synthesizer-Impulse in den Hauptteilen des Stücks. Gepannt wurden diese auf die beiden vorderen Höhenlautsprecher. Die darauffolgende musikalische Antwort jeweils einen Takt später (ebenfalls mit einem Synthesizer umgesetzt) wurde dagegen komplett auf LS und RS gelegt. Die Gefahr beim Panning entsprechend dominanter Klänge auf umgebende Lautsprecher ist ein Ungleichgewicht im Gesamtmix. Der Hörer verliert die Orientierung, die beim Hören normalen Stereo-Materials für gewöhnlich gegeben ist. Doch durch den vorausgehenden verwandten Klang aus HL und HR wird das Gleichgewicht gewahrt. So kann die räumliche Verteilung der Elemente ebenso das musikalische Zusammenspiel dieser beiden Motive unterstützen.

In Bezug auf Struktur des Stückes war die Vision, den Titel ohne wesentliche Akkordfolgen recht kühl, dunkel und monoton starten zu lassen. Zudem wurde in einem Zwischenteil (1:32-1:47) angestrebt, durch das Weglassen jeglicher Raumanteile und Klangflächen einen bedrohlich direkten Klangeindruck zu schaffen. Im darauffolgenden und den Hörversuch startenden ruhigen, atmosphärischen Mittelteil (2:02-2:37) wird mit gegenteiligen Mitteln gearbeitet. Die Wiedergabe und Automatiserung von Klangflächen und Pads über die Höhenebene bewirken eine wirksame Umhüllung. Glockenähnliche, an verschiedene Stellen im Raum gepannte Sounds, kontrastieren diese Umhüllung und leiten gleichermaßen die melodische Entwicklung des Stückes ein. Der vorhandene Höhenanteil in diesen Sounds steigert dabei die Ortbarkeit erheblich. Der darauffolgende Hauptteil beinhaltet die bereits aus den vorherigen Teilen bekannten Elemente. Die Flächen bleiben weiterhin bestehen und die Akkordfolge zeichnet sich bei Beateinsatz auch im eingangs erwähnten, rauen Basssound ab. Später setzt erneut der Gesang ein, der, wie nicht selten in der Musikrichtung, recht eingängig und einfach gehalten, von der Idee des Projektes handelt.

Nachdem die Hauptstimme beim ersten Teil des Stückes aus dem VOG kam, kommt sie in diesem Teil aus dem HC. Die gedoppelte zweite Stimme wurde auf HL und HR verteilt. Ebenso wurden zwei energetischer gesungene Dopplungen auf HLS und HRS verteilt. Ziel der Verlagerung der Stimmen in die Höhe ist eine, wie die Bezeichnung VOG schon anspielt, angestrebte "überirdische" Wirkung.

Zur Hallerzeugung während des Produzierens wurde ein Stereo-Hall verwendet, der auf die umgebenden Lautsprecher verteilt wurde. Vor Abschluss der Produktion wurde dann die trockene Stereo-Return-Spur exportiert und über die DAW *Protools* und das Plugin *Altiverb 7* zwei vierkanalige Hallspuren erstellt. Eine für L, R, LS und RS und eine für HL, HR, HLS und HRS. Letztere mit leicht längerem Predelay, Nachhall und einer etwas stärkeren Höhendämpfung, um auch hier unkorrelierte Signale zu erhalten und mögliche Eigenschaften eines realen Raumes nachzubilden.

Nach Fertigstellung des 3D-Mixes wurde die Erstellung des Stereo-Mixes angegangen. Die größtenteils stereophonen Elemente des Titels wurden anstatt über Sends auf die Lautsprecherfeeds direkt mit den auch bei der Stereo-Mischung verwendeten Pegeln auf den Stereo-Master geschickt.

Sowohl Klang- und Lautstärkeverhältnisse, als auch Kompressionseinstellungen der einzelnen Spuren, mussten beim Stereo-Mix wesentlich akribischer eingestellt werden, um nicht im Gesamtbild unterzugehen und eine druckvolle Mischung zu erhalten. Dies lässt sich mit dem Prinzip der frequenzabhängigen Verdeckung erklären. Laute Elemente verdecken unmittelbar leisere, die sich im selben Frequenzbereich abspielen [1, pp. 130-131]. Wie bereits in Abschnitt 3.2.2 erläutert, können Elemente an verschiedenen Positionen dagegen richtungsbasiert gefiltert werden.

Zudem existiert beim Mischen in Stereo unmittelbar ein Klangideal, das sich jeder Musikrichtung entsprechend über Jahre hinweg entwickelt hat. Bei diesem Titel machte sich dies in dem Bedürfnis bemerkbar, einen nicht unbeachtlichen Anteil an Summenkompression anzuwenden, um den Titel dicht klingen und im klanglichen Sinne pumpen zu lassen.

Eine weitere, eher persönliche Erfahrung waren die Bedenken, bedeutende Elemente im Stereo-Panorama vollständig auf eine Seite zu pannen. Die kreative Arbeit mit 3D-Audio lässt hingegen durch Unbefangenheit angesichts noch nicht etablierter Klangideale und Hörgewohnheiten Spielraum. Darüber hinaus besteht bei extremem Stereo-Panning stets die Sorge der noch extremeren Stereo-Wahrnehmung über Kopfhörer. 3D-Inhalte können dagegen nur über Umwege auf Kopfhörern wiedergegeben werden. Eine gängige Methode ist die Binauralisierung über HRTFs. Eine einzelne HRTF beschreibt dabei die Übertragungsfunktion eines Schallereignisses auf ein Ohr mit definierter Position in Bezug auf den Kopf [31]. Wird ein Signal aus einer bestimmten Richtung mit den HRTFs beider Ohren für die entsprechende Richtung gefaltet (also die Übertragungsfunktion mit

dem Nutzsignal verrechnet [1, p. 572]), kann bei Kopfhörerwiedergabe die Position des Schallereignisses im Idealfall originalgetreu reproduziert werden. Werden Stereo-Inhalte auf Kopfhörern angehört, wandern auf horizontaler Ebene extrem auf eine Seite gepannte Signale von ±30° (Wiedergabe im gleichseitigen Dreieck) hin zu ±90° (Kopfhörer). Dies lässt sich im 3D-Studio, sowie binauralisiert in etwa mit der Position der Surround-Lautsprecher vergleichen, zumal das Signal in diesem Fall auch jeweils auf dem anderen Ohr mit entsprechender Laufzeitverzögerung und Pegelabnahme ankommt.

4. Hörversuch

4.1 Konzeption des Hörversuchs

In einem Hörversuch soll nun erforscht werden, ob und inwiefern 3D für den Musikhörer einen Mehrwert gegenüber Stereo in Bezug auf die Musikrichtung darstellt. Um einen möglichen Mehrwert bei Musik feststellen zu können, muss aber zunächst der eigentliche Wert von Musik definiert werden. Sinnvoll hierbei scheint die Orientierung an der menschlichen Motivation, Musik zu hören.

Nach Schramm und Kopiez (2011) gehören "zu den Nutzungsmotiven […] beispielsweise das Regulieren von Stimmungs- und Erregungszuständen, die kognitive, emotionale und assoziative Auseinandersetzung mit dem musikalischen Werk, das Ausdrücken eigener Identität sowie das Abgrenzen von anderen sozialen Gruppen" [7, p. 253].

Gurgen (2016) zeigt zusätzlich mit ihrer Forschung auf, dass die Motivation Musik zu hören zum größten Teil aus persönliche Faktoren wie Genuss und Stimmung bzw. Gefühlslage (Enjoyment, Mood / Emotional State) besteht. Seltener begründeten die Probanden ihrer Studie den Grund Musik zu hören durch umgebungsbedingte (Freundeskreis, Familie etc.) oder musikalische Faktoren. [8, pp. 235-237]

Dadurch scheint es plausibel, die Ermittlung des 3D-Mehrwertes eben über diese persönlichen Faktoren durchzuführen. Die Teilnehmer des Hörversuchs sollen zum einen angeben, wie sehr ihnen das Stück gefällt [a] und zum anderen die Intensität der bei ihnen ausgelösten emotionalen Erregung [b] bewerten. Diese eigenständige Bewertung anhand einer numerischen Skala wird als gängige Methode zur Feststellung von Gefühlsempfindungen angesehen [32, p. 40]. Die emotionale Erregung wird dabei ganzheitlich betrachtet, also nicht in verschiedene emotionale Zustände aufgeteilt. Dennoch werden den Probanden vor dem Hörversuch acht Oberbegriffe potenzieller emotionaler Eigenschaften von Musik vorgelegt, um ein Bewusstsein für ihre möglichen Wirkungsweisen zu schaffen (siehe Anhang 2). Diese entsprechen den Oberbegriffen der acht verschiedenen Cluster des Hevnerschen Adjektivzirkels, der der Gliederung und Beschreibung musikalischer Emotionen dient, also ernsthaft, traurig, sentimental, ruhig, humorvoll, fröhlich, aufregend und majestätisch [33, pp. 27-28]. Zudem ist in den freien Zeilen Platz für eigene Bemerkungen, Beobachtungen und Kommentare [c].

Über die Vorgehensweise der verbundenen Stichprobe wird jeder Teilnehmer unter jeder experimentellen Bedingung, also jedem Hörbeispiel, getestet. Somit können individuelle Differenzen wie unterschiedliche Musikpräferenzen oder Vorerfahrungen mit 3D-Audio als Fehlervarianz

vernachlässigt werden. Um systematische Sequenzeffekte, also die gegenseitige Beeinflussung der Titel und Formate in Abhängigkeit der Wiedergabereihenfolge zu vermeiden, wird die Reihenfolge der 14 Hörbeispiele rein zufällig bestimmt. Für jeden Durchgang besteht eine vorher festgelegte, zufällige Reihenfolge (siehe Anhang 4), die nachträglich für die Auswertung wieder zurückcodiert wird. Während des Versuchs achtet der Versuchsleiter strikt darauf, den Probanden nicht zu beeinflussen. Zudem versichert jeder Proband mit der Teilnahme am Hörversuch, über ein unbeeinträchtigtes Hörvermögen zu verfügen. [34, pp. 10-39]

Auszug aus dem Beiblatt des Hörversuchs (siehe Anhang 2):

Für den Hörversuch wurden 7 Stücke verschiedener Musikrichtungen ausgewählt, für die jeweils ein 1-minütiger Ausschnitt festgelegt wurde. Für diesen Ausschnitt bestehen jeweils eine geläufige Stereo- und eine 3D-Version. Der Test umfasst somit 14 kurze Hörbeispiele aus 7 verschiedenen Stücken. Die Wiedergabereihenfolge der Hörbeispiele wurde vor dem Test rein zufällig bestimmt, es ist also beispielsweise gut möglich, dass die zwei Versionen eines Stückes nicht nacheinander wiedergegeben werden. Ob es sich bei dem gespielten Beispiel um Stereo- oder 3D-Inhalt handelt, wird nicht angezeigt.

Letztgenannter Entschluss hat das Ziel, eine reine Bewertung des Gehörten zu erhalten und die Manipulation der Ergebnisse durch Suggestion zu vermeiden.

Zur Durchführung des Hörversuchs besteht ein Fragebogen (siehe Anhang 1), der anfangs das Alter und Geschlecht des Probanden, seine Beziehung zur Musik, seine Vorerfahrung mit 3D-Audio und persönliche Genrepräferenzen abfragt. Dies schafft die Möglichkeit, die Hörer bei der späteren Auswertung in verschiedene Probandengruppen aufzuteilen. Nachdem Seite 1 des Fragebogens ausgefüllt ist, wird das Beiblatt (siehe Anhang 2) ausgeteilt. Dies erklärt, wie bereits angeschnitten, den eigentlichen Hörversuch im Detail. Auf Seite 2-4 des Fragebogens macht der Proband Angaben zum Gefallen und der emotionale Erregung des jeweiligen Hörbeispiels auf einer Skala von 0 bis 10. Für Anmerkungen steht ein zusätzliches Kommentarfeld bereit.

Um den Überblick über Teilnehmer und Durchgänge zu wahren, wurden zusätzliche Notizblätter (siehe Anhang 3) verwendet.

Während des Hörversuchs ist im Studio Platz für maximal vier Personen (zwei Zweierreihen um den Sweetspot). Die Lautstärken wurden sowohl zwischen Stereo- und 3D-Versionen, als auch zwischen den verschiedenen Titeln über das Gehör angeglichen, sodass bei allen Hörbeispielen der gleiche Lautstärkeeindruck entsteht. Es wurde ein lauter, dennoch angenehmer einheitlicher Abhörpegel

gewählt. Zur Wiedergabe der Stereo-Beispiele wurde der LFE zur Erweiterung des Bassbereichs zugezogen.

4.2 Forschungshypothesen

Verschiedene Forschungen im Bereich 3D-Audio zeigen bereits eine tendenzielle Steigerung verschiedener Bewertungskriterien, u.a. auch des Gefallens und der emotionalen Erregung durch die Wiedergabe in 3D auf. [9, 12] Dies legt den Grundstein für die ersten beiden Hypothesen.

H₁: 3D-Audio steigert das Gefallen des Musikstücks im Vergleich zu Stereo.

H₂: 3D-Audio steigert die emotionale Erregung des Musikstücks im Vergleich zu Stereo.

Zudem zeigt *Hahn* (2017) in seinen Untersuchungen einen tendenziell größeren Anstieg der emotionalen Erregung auf [9], was zur dritten Hypothese führt.

H₃: 3D-Audio bewirkt eine größere Steigerung der emotionalen Erregung, als des Gefallens.

Elvers, Omigie et al. (2015) legen in ihrer Arbeit unterschiedliche Tendenzen im Musikgeschmack bei Aufteilung der Probanden in verschiedene Gruppen dar [35]. Auch Hahn (2017) weist auf Unterschiede zwischen verschiedenen Gruppen hin [9].

- H₄: Die Bewertungen der Items unterscheiden sich zwischen den verschiedenen Hörgruppen bei der gruppenweisen Auswertung.
- H₅: Die Bewertungsdifferenzen der Items zwischen 3D und Stereo unterscheiden sich zwischen den verschiedenen Hörgruppen bei der gruppenweisen Auswertung.

Die genannten Forschungshypothesen werden in Bezug auf die jeweils eigenständigen Hörbeispiele der verschiedenen Musikrichtungen ausgewertet.

Neben der Überprüfung der theoretisch hergeleiteten Hypothesen, wird zur weiteren Auswertung der Daten die Methode der explorativen Datenanalyse gewählt und sich dementsprechend andeutenden Tendenzen weiter nachgegangen. Diese Methode findet gerade bei wenig erforschten Sachverhalten Verwendung. Typisch ist auch die für diese Arbeit zutreffende hohe Anzahl an Variablen. [36, p. 621]

5. Auswertung

Die Hörversuche fanden vom 19. bis 28. Februar 2019 im 3D-Audio Filmtonstudio der *HdM* statt. Im Anschluss wurden die Daten mit Hilfe der Statistik- und Analysesoftware *SPSS* ausgewertet.

Insgesamt wurden in 33 Durchgängen 61 Probanden im Alter von 18 bis 59 Jahren befragt. Der Altersschnitt lag bei 27,7 Jahren (SD = 10,8). 52 Versuchspersonen (85,2%) waren *männlich*, 9 (14,8%) weiblich. 20 Probanden (32,8%) gaben an, reine Musikhörer zu sein. 18 (29,5%) kategorisierten sich als Hobbymusiker und 23 Probanden (37,7%) gaben eine professionelle Beschäftigung mit Musik/Tontechnik an. Des Weiteren gaben 34 Probanden (55,7%) an, schon einmal Musik auf einem 3D-Audio-System gehört zu haben. Für 27 Versuchspersonen (44,3%) war es dagegen das erste Mal.

Bewertungen zum generellen Gefallen der Musikrichtungen und zum Gefallen und der emotionalen Erregung der Hörbeispiele wurden von den Probanden auf einer Likert-Skala von 0-10 angegeben (siehe Anhang 1). Hier wurde die explizite Formulierung der Extreme 0: "gar nicht" und 10: "sehr gut" bzw. "sehr stark" als möglichst unmissverständliche Minima und Maxima der menschlichen Empfindung gewählt. Zudem wurde gezielt eine ungerade Anzahl an Antwortmöglichkeiten gewählt, um mit der grau hinterlegten Antwortmöglichkeit mit dem Wert 5 einen mittleren, neutralen Orientierungspunkt zu bieten. Dies impliziert gleichermaßen die Äquidistanz der Skala, die in der Praxis zur Auswertung der eigentlich ordinalskalierte Likert-Skala mit intervallskalierten Methoden berechtigt [37].

Die folgenden Diagramme und Tabellen sind, sofern nicht anders angegeben, wie in Abbildung 21 beschrieben farblich kodiert.

- ☐ Beliebtheit der Musikrichtung
- Gefallen der Stereo-Version
- Gefallen der 3D-Version
- Emotionale Erregung der Stereo-Version
- Emotionale Erregung der 3D-Version
- \square p > 0,05: nicht signifikant
- \square p \leq 0,05: signifikant
- \square p \leq 0,01: sehr signifikant
- \square p \leq 0,001: hoch signifikant

Abbildung 21: Legende für Diagramme (links) und p-Wert-Tabellen (rechts)

5.1 Gesamtanalyse

Einen ersten Überblick über die Datenlage bietet die grafische Darstellung der Bewertungsmittelwerte aller Probanden.

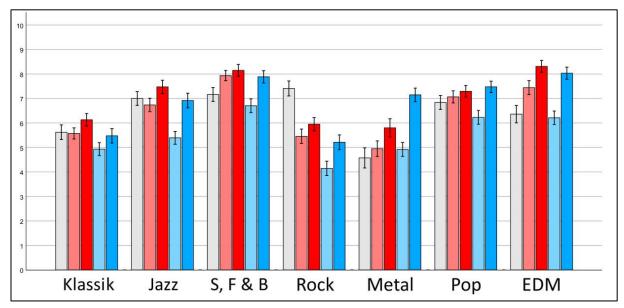


Abbildung 22: Mittelwerte der Beliebtheit der Musikrichtungen, des Gefallens und der emotionalen Erregung der Hörbeispiele in Stereo und 3D; Fehlerbalken: ± 1 Standardfehler

Kolmogorov-Smirnov-Tests zeigen, dass lediglich die Werte der Variablen Beliebtheit von Klassik und emotionale Erregung des klassischen Hörbeispiels in 3D normalverteilt sind. Dadurch werden für die paarweise Auswertung der verbundenen Stichproben (Vergleich von Stereo und 3D) Wilcoxon-Tests gewählt, die keine Normalverteilung der Daten voraussetzen.

Die Ergebnisse der Wilcoxon-Tests in Tabelle 4 zeigen, dass bei der Bewertung des Gefallens des Popund S, F & B-Hörbeispiels und der emotionalen Erregung des Klassik-Hörbeispiels keine signifikanten Unterschiede in Abhängigkeit des Wiedergabeformats auszumachen sind. Hierbei ist allerdings anzumerken, dass die beiden letztgenannten den Signifikanzbereich bei einem Signifikanzniveau von 5% nur leicht überschreiten, das Gefallen bei Pop aber eindeutig. Alle anderen Testergebnisse zeigen eine signifikante Steigerung des Gefallens und der emotionalen Erregung von Stereo zu 3D. Auffallend hierbei ist die durchgehend hoch signifikante Steigerung der emotionalen Erregung mit Ausnahme des klassischen Hörbeispiels. Jazz, Metal und EDM zeigen diese Eigenschaft für Gefallen und emotionale Erregung.

	Klassik	Jazz	S, F & B	Rock	Metal	Pop	EDM
Gefallen (S – 3D)	0,001	0,000	0,067	0,019	0,000	0,264	0,000
Erregung (S – 3D)	0,076	0,000	0,000	0,000	0,000	0,000	0,000

Tabelle 4: p-Werte der Wilcoxon-Tests: paarweiser Vergleich der Bewertungen in Stereo und 3D

In der Gesamtauswertung kann die Hypothese H_{1/Gesamt} für Klassik, Jazz, Rock, Metal und EDM angenommen werden, es steigert sich also das Gefallen dieser Musikrichtungen durch 3D. Keine signifikante Steigerung zeigt sich bei S, F & B und Pop.

Die Hypothese H_{2/Gesamt} kann mit Ausnahme der Klassik bestätigt werden. Die emotionale Erregung steigert sich bei allen Musikrichtungen außer Klassik signifikant durch die Wiedergabe in 3D.

Grafisch kann das Ausmaß der Steigerung über die Differenz der Mittelwerte (= Mittelwert 3D – Mittelwert Stereo) in einem Säulendiagramm (siehe Abbildung 23) dargestellt werden. So wird mehr die Relation der Bewertungen zwischen Stereo und 3D, als der absolute Wert betrachtet.

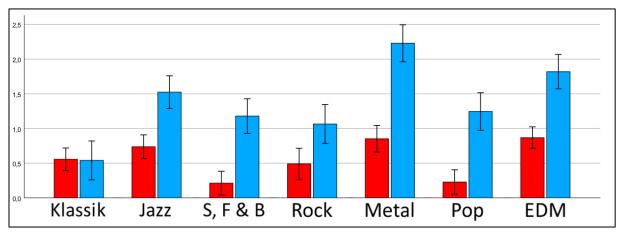


Abbildung 23: Mittelwertsdifferenz des Gefallens (rot) und der emotionalen Erregung (blau) zwischen 3D- und Stereo-Bewertung; Fehlerbalken: ± 1 Standardfehler

In Abbildung 23 deutet sich vermehrt eine höhere Steigerung der emotionalen Erregung als Steigerung des Gefallens bei Wiedergabe in 3D an. Weitere Wilcoxon-Tests bestätigen dies mit Ausnahme des klassischen Hörbeispiels mit durchweg hoher Signifikanz. Beim klassischen Hörbeispiel muss dagegen definitiv die Nullhypothese des Wilcoxon-Tests, dass also kein Unterschied besteht, angenommen werden. Den größten Mehrwert bewirkt das 3D-Format in der emotionalen Erregung bei Metal, EDM und Jazz.

	Klassik	Jazz	S, F & B	Rock	Metal	Pop	EDM
diff. Gefallen – diff. Erregung	0,995	0,006	0,000	0,009	0,000	0,000	0,000

Tabelle 5: p-Werte der Wilcoxon-Tests: paarweiser Vergleich der Mittelwertsdifferenz des Gefallens und der emotionalen Erregung

Es findet also eine signifikant höhere Steigerung der emotionalen Erregung, als des Gefallens durch 3D statt. Hypothese H_{3/Gesamt} kann für alle Musikrichtungen, mit Ausnahme der Klassik, bestätigt werden.

Zuletzt steht in der allgemeinen Betrachtung noch die Untersuchung aus, ob sich die zu Beginn abgefragte Beliebtheit der übergreifenden Musikrichtungen mit dem Gefallen der zugehörigen

Hörbeispiele in Stereo als gebräuchliches "Referenzformat" deckt. Dies ermöglicht zum einen Rückschlüsse auf die Übertragbarkeit und Aussagekraft des Hörbeispiels auf das Genre, zum anderen aber auch Rückschlüsse auf die Qualität und das Gefallen des Hörbeispiels bei den Probanden innerhalb eines Genres. Es dient somit der Evaluation der Hörbeispiele.

	Klassik	Jazz	S, F & B	Rock	Metal	Pop	EDM
Beliebtheit - Gefallen in Stereo	0,943	0,306	0,008	0,000	0,144	0,504	0,000

Tabelle 6: p-Werte der Wilcoxon-Tests: paarweiser Vergleich der Beliebtheit und des Gefallens in Stereo

Bei Klassik, Jazz, Metal und Pop muss die Nullhypothese der Wilcoxon-Tests in Tabelle 6 angenommen werden. Die zentralen Tendenzen von Beliebtheit und Gefallen unterscheiden sich nicht. Es kann also von einem auf die Musikrichtung zutreffenden Beispiel gesprochen werden. S, F & B und EDM übersteigen zudem die Beliebtheitswerte mit hoher Signifikanz. Einzig das Rock-Hörbeispiel fällt mit seiner Bewertung aus der Reihe. Wie in Abbildung 22 zu sehen, liegt die Beliebtheit der beliebtesten Musikrichtung weit über der Bewertung des Hörbeispiels.

5.2 Gruppenanalyse

5.2.1 Hörtyp

In Abbildung 24 sind die Bewertungsmittelwerte jedes Hörtyps aufgeführt. Jede Variable kann als Triplett gesehen werden. Der linke Balken beschreibt jeweils die Bewertung der reinen Musikhörer, der mittlere die der Hobbymusiker und der rechte Balken die Bewertung der Experten, also professionellen Tontechniker/Musiker.

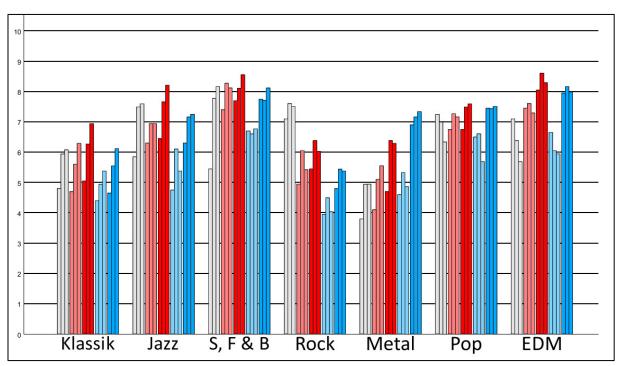


Abbildung 24: Bewertungsmittelwerte in Abhängigkeit des Hörtyps; Tripletts v.l.: reiner Musikhörer, Hobbymusiker, prof. Musiker/Tontechniker; im Sinne der Übersichtlichkeit keine Fehlerbalken

H_{4/Hörtyp}:

Die erste Frage bei Analyse der Daten in Abhängigkeit des Hörtyps ist die Prüfung auf signifikante Bewertungsunterschiede der gleichen Variablen zwischen den verschiedenen Hörtypen. Hierfür bietet sich der Kruskal-Wallis-Test an. Tabelle 7 beschränkt sich hierbei lediglich auf die Variablen mit signifikanten Unterschieden zwischen den Hörtypen.

Klassik	Jazz	S, F & B
Gefallen Stereo: 0,008	Beliebtheit: 0,021	Beliebtheit: 0,000
Gefallen 3D: 0,007	Gefallen 3D: 0,030	

Tabelle 7: signifikante p-Werte der Kruskal-Wallis-Tests: Prüfung auf unterschiedliche Tendenzen in der Bewertung der gleichen Variablen in Abhängigkeit des Hörtyps

Zwar deuten sich im Diagramm in Abbildung 24 weitere Tendenzen an, doch mit der gesunkenen Stichprobengröße (etwa 20 Probanden pro Hörtyp) steigt die Wahrscheinlichkeit, dass aufgekommene Unterschiede zufällig sind. Dies steht in direktem Zusammenhang mit der Signifikanz.

Dementsprechend vergrößern sich bei der Aufteilung in Hörtypen auch die aus Gründen der Übersichtlichkeit weggelassenen Fehlerbalken teilweise um ein Vielfaches.

H_{1,2/Hörtyp}:

Die schon in der Gesamtanalyse durchgeführten Wilcoxon-Tests zum Vergleich der Bewertungen von Stereo und 3D (siehe Tabelle 4) liefern bei der gruppenweisen Auswertung interessante Erkenntnisse (siehe Tabelle 8). Die Gruppe der reinen Musikhörer empfindet lediglich in vier Bewertungsitems eine signifikante Steigerung durch 3D. Bei den Hobbymusikern sind es bereits sieben und für die Expertengruppe erfahren bereits zehn der insgesamt 14 zu bewertenden Items einen signifikanten Mehrwert durch das 3D-Format. Zudem steigert sich durch die Wiedergabe in 3D das Gefallen sowie die emotionale Erregung des EDM-Beispiels bei allen drei Hörtypen signifikant. Dies gilt mit Ausnahme des Gefallens bei reinen Musikhörern ebenso für das Jazz- und Metal-Hörbeispiel.

	(S - 3D)	Klassik	Jazz	S, F & B	Rock	Metal	Pop	EDM
reine	Gefallen	0,083	0,587	0,127	0,311	0,123	1,000	0,033
Musikhörer	Erregung	0,730	0,002	0,074	0,188	0,002	0,071	0,015
Hobby-	Gefallen	0,070	0,033	0,782	0,318	0,006	0,572	0,017
musiker	Erregung	0,202	0,021	0,028	0,058	0,001	0,072	0,001
Experten	Gefallen	0,035	0,001	0,112	0,064	0,037	0,223	0,000
	Erregung	0,146	0,000	0,001	0,003	0,000	0,002	0,000

Tabelle 8: p-Werte der Wilcoxon-Tests: paarweiser Vergleich der Bewertungen in Stereo und 3D in Abhängigkeit des Hörtyps

H_{5/Hörtyp}:

Zuletzt zeigen weitere Kruskal-Wallis-Tests eine interessante Tatsache auf. Hierbei wurden die Bewertungsdifferenzen (3D – Stereo) zwischen den Hörtypen verglichen. Lediglich die Bewertungsdifferenz des Gefallens bei Jazz zeigt signifikante Unterschiede zwischen den Hörtypen auf (p = 0,018). Während sich das Gefallen am Jazz-Beispiel für den reinen Musikhörer durch 3D nicht signifikant ändert, steigt er für den Experten hoch signifikant an. Abbildung 24 veranschaulicht dies deutlich.

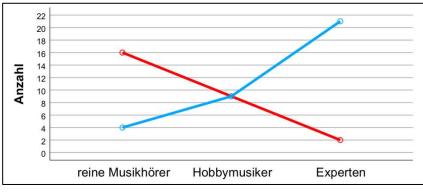


Abbildung 25: Vorerfahrung mit Musik in 3D in Abhängigkeit des Hörtyps; rot: keine Vorerfahrung, blau: Vorerfahrung

Durch die hohe Korrelation zwischen Hörtyp und Vorerfahrung mit Musik in 3D-Audio des jeweiligen Probanden, wurde letztere nicht separat ausgewertet. Während die meisten reinen

Musikhörer noch nie zuvor

Musik in 3D-Audio gehört hatten, war dies für die Expertenhörer nichts Neues.

5.2.2 Alter

Für die Analyse der Altersgruppen wurden die Probanden in die zwei etwa gleichgroßen Gruppen 18-23 Jahre (n = 32; 52,5%) und 24-59 Jahre (n = 29; 47,5%) aufgeteilt.

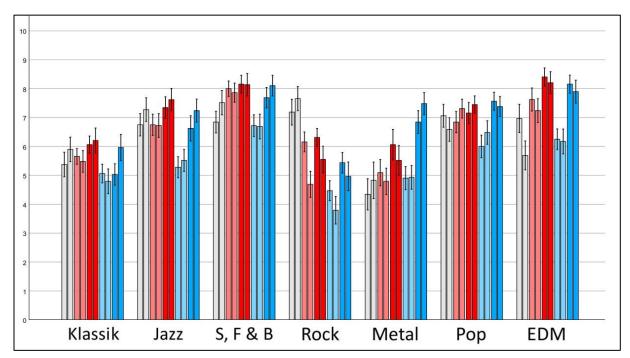


Abbildung 26: Bewertungsmittelwerte in Abhängigkeit des Alters; paarweise Gruppierung: links 18-23 J., rechts 24-59 J.; Fehlerbalken: ± 1 Standardfehler

H_{4/Alter}:

Da es sich in diesem Fall um lediglich zwei Stichproben handelt, bieten sich Mann-Whitney-U-Tests (nicht parametrische Tests zweier unabhängiger Stichproben) zur Untersuchung auf signifikante Bewertungsunterschiede an. Lediglich die Bewertung des Gefallens des Rock-Hörbeispiels in Stereo unterscheidet sich signifikant zwischen den Altersgruppen (p = 0,021). Sie fällt bei den älteren Probanden wesentlich niedriger aus. Nicht weniger erwähnenswert zeigt sich die Beliebtheit der Musikrichtung EDM (p = 0,058), die sich, wenn auch nicht signifikant, bei den jüngeren als beliebter andeutet.

H_{1,2/Alter}:

Gruppenweise durchgeführte Wilcoxon-Tests zum Vergleich der Stereo- und 3D-Bewertungen zeigen darüber hinaus, dass sich bei Klassik lediglich für die ältere Altersgruppe sowohl das Gefallen, als auch die emotionale Erregung signifikant steigern. Gleichermaßen verhält es sich mit dem Gefallen des Rock-Beispiels.

	(S – 3D)	Klassik	Jazz	S, F & B	Rock	Metal	Pop	EDM
18 - 23 J.	Gefallen	0,109	0,043	0,300	0,486	0,005	0,258	0,007
16 - 25 J.	Erregung	0,919	0,000	0,008	0,020	0,000	0,001	0,000
24 - 59 J.	Gefallen	0,002	0,001	0,108	0,009	0,014	0,676	0,000
	Erregung	0,011	0,000	0,001	0,005	0,000	0,030	0,000

Tabelle 9: p-Werte der Wilcoxon-Tests: paarweiser Vergleich der Bewertungen in Stereo und 3D in Abhängigkeit des Alters

H_{5/Alter}:

Um die Bewertungsdifferenzen (3D – Stereo) zwischen beiden Gruppen auf Unterschiede zu prüfen, wurden weitere Mann-Whitney-U-Tests angesetzt. Deren Nullhypothese, dass sich die zentralen Tendenzen der beiden unabhängigen Stichproben nicht unterscheiden, muss in allen Fällen angenommen werden. Als Besonderheit kann hier die beinahe signifikant größere Bewertungsdifferenz der emotionalen Erregung im klassischen Hörbeispiel durch die älteren Probanden (p = 0,059) genannt werden.

5.2.3 Geschlecht

Bei der Auswertung der Daten in Abhängigkeit des Geschlechts muss die unausgeglichene Gruppenverteilung mit 52 männlichen Probanden (85,2%) und 9 weiblichen Probanden (14,8%) im Hinterkopf behalten werden. Es gilt also, die Daten mit einer gewissen Vorsicht zu interpretieren. Dennoch ergeben sich bei der Analyse klare Tendenzen.

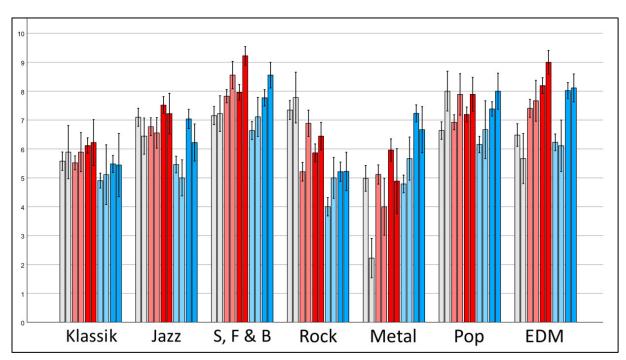


Abbildung 27: Bewertungsmittelwerte in Abhängigkeit des Geschlechts; paarweise Gruppierung: links männlich, rechts weiblich; Fehlerbalken: ± 1 Standardfehler

H_{4/Geschlecht}:

Zur Prüfung auf Unterschiede in der Bewertung der gleichen Items zwischen den beiden Gruppen wurden Mann-Whitney-U-Tests durchgeführt. Items mit signifikanten Bewertungsunterschieden sind in Tabelle 10 aufgeführt. Zudem ergaben sich beinahe signifikante Bewertungsunterschiede in der Beliebtheit (p = 0.061) der Popmusik und dem Gefallen des Stereobeispiels des gleichen Genres (p = 0.078).

S, F & B	Rock	Metal
Gefallen 3D: 0,037	Gefallen Stereo: 0,041	Beliebtheit: 0,019

Tabelle 10: signifikante p-Werte der Mann-Whitney-U-Tests: Prüfung auf unterschiedliche Tendenzen in der Bewertung der gleichen Variablen in Abhängigkeit des Geschlechts

H_{1,2/Geschlecht}:

Die bei den anderen Gruppen durchgeführten Wilcoxon-Tests, zum gruppenweisen Vergleich der Bewertungen von Stereo und 3D, ergeben in diesem Fall infolge unausgeglichener Gruppenverteilung keinen Sinn. Aufgrund der niedrigen Anzahl an Stichproben bei weiblichen Probanden liefern die Wilcoxon-Tests, trotz augenscheinlich großer Unterschiede (siehe Abbildung 27), nur wenige sich signifikant unterscheidende Bewertungen zwischen Stereo und 3D. Die Ergebnisse zwischen männlichen und weiblichen Teilnehmern können somit später nicht sinnvoll verglichen werden.

H_{5/Geschlecht}:

Dagegen bieten sich Mann-Whitney-U-Tests zur Gegenüberstellung der Bewertungsdifferenzen (3D – Stereo) zwischen den beiden Gruppen an. Signifikant zeigt sich der Unterschied der Bewertungsunterschiede des Gefallens des Rock-Beispiels (p = 0,048) und der emotionalen Erregung des Metal-Beispiels (p = 0,025). Abbildung 27 verdeutlicht, dass 3D für die männlichen Teilnehmer in den eben genannten Punkten einen signifikant höheren Mehrwert darstellt.

5.3 Qualitative Analyse und Interpretation

Da die Kommentarfelder wichtige Informationen zur Interpretation der Ergebnisse liefern, werden die qualitative Analyse der Kommentarfelder und die Ergebnisinterpretation parallel durchgeführt.

Zur Auswertung der ausgiebig genutzten Kommentarfelder wurden die Kommentare in einer Tabelle gesammelt und sinngemäß zusammengefasst.

Abbildung 28 bietet hierbei eine Visualisierung der häufigsten Kommentare in Bezug auf das Wiedergabeformat unabhängig von der Musikrichtung. Größe und Rotanteil veranschaulichen die Häufigkeit der Kommentare (groß und rot = häufig).



Abbildung 28: Wortwolken der häufigsten Kommentare in Bezug auf das Wiedergabeformat

Auch in den Kommentaren zeichnet sich ein Mehrwert durch Wiedergabe in 3D ab. In vielen Punkten ist der direkte Vergleich zwischen beiden Formaten nicht abzustreiten. Dies sollte auch bei Betrachtung der Kommentare stets im Kopf behalten werden und die auf den ersten Blick negativ wirkenden Stereo-Bewertungen etwas relativieren. Ein Großteil der häufigsten Kommentare bezieht sich auf rein klangliche Aspekte des Gehörten, einige Kommentare aber auch auf Gefühlsempfindungen beim Hören.

Unter den Stereo-Beispielen häufen sich Kommentare zur räumlichen Wirkung des Gehörten. Die Stereo-Beispiele werden als *flach*, *schmal*, *eng* und *frontal* beschrieben. Das Pendant bietet dagegen die für viele Probanden eindrücklich beschriebene *Umhüllung* in 3D. Diese wird oft mit dem Gefühl, *mittendrin* zu sitzen bzw. *live* dabei zu sein, umschrieben. Auch die Worte *räumlich*, *weit*, *breit* und *offen* fallen. Einige Male werden die 3D-Beispiele allerdings auch als *zu diffus* wahrgenommen.

Ebenso häufig wird in den Kommentaren auf den Klang der 3D-Beispiele unabhängig vom Raumeindruck eingegangen. Die Mischungen werden als *voll, angenehm, transparent, ausgewogen* und *klar* beschrieben. Ähnliche Worte finden sich mit *klar, natürlich* und *voll* auch unter den Stereo-

Beispielen wieder. Oft werden aber auch Klangbeschreibungen wie *spitz, hart, direkt, dicht, anstrengend, nervig, kantig* und *verwaschen* gewählt. Dies geht Hand in Hand mit den Kommentaren zur *gegenseitigen Überlagerung* der verschiedenen Elemente in Stereo. Dieser Eindruck kommt im 3D-Beispiel nicht auf. Die Versuchspersonen haben mehrheitlich das Gefühl, dass *genügend Platz für alle Elemente* ist.

Eine mögliche Erklärung für den anstrengenden Höreindruck in Stereo gibt das in Kapitel 3.2.2 beschriebene Phänomen. Durch die bessere Separation der Instrumente im 3D-Raum kann in 3D sowohl bei Frequenz-, als auch Dynamikbearbeitung natürlicher gearbeitet werden. Im Umkehrschluss bedeutet dies ein adynamischeres und unnatürlicheres Klangbild der Stereo-Version im direkten Vergleich. Bei Angleichung der Musiktitel hinsichtlich Lautstärkeeindruck und anschließender Festlegung einer einheitlichen Wiedergabelautstärke wird eine Sache deutlich: Wird eine für die Titel in 3D laute und dennoch angenehme Lautstärke gewählt, ist dies für die Stereo-Wiedergabe schon deutlich zu laut und unangenehm. Zum einen durch den hohen Pegel, der alleine von den vorderen beiden Lautsprechern aus einer einzigen Richtung wiedergegeben wird, zum anderen aber auch durch das dichtere Klangbild. Durch die geringere Dynamik tritt beim Hörer demnach eine schnellere Ermüdung ein [38]. Betrachtet man zudem beispielsweise die Stereo-Version des Jazz-Beispiels, wird das ohnehin höhenreiche Saxophon bei entsprechendem Abhörpegel schnell unangenehm. In der 3D-Version wurden die Soli wesentlich dezenter gefahren, da sie dort allein durch ihre Positionierung differenzierter wahrgenommen werden können. Die Ausgewogenheit des Klanges macht sich demnach auch bei lauteren Pegeln bezahlt. Für die Hörversuche wurde ein, auch bei der Stereo-Wiedergabe annehmbarer, einheitlicher Pegel gewählt.

Auffallend ist dagegen, dass auch einige der Probanden die Stereo-Versionen, gerade bei den energetischeren Beispielen, als *druckvoller* beschreiben.

Die bereits aufgezeigte höhere Emotionalität durch Wiedergabe in 3D wird auch in den Kommentaren ersichtlich. Starke Indikatoren sind hierbei die mehrmals aufgezählte Gänsehaut, der Bewegungsdrang, aber auch die mitreißenden Eigenschaften. Eine ebenso immer wieder aufgezählte Bemerkung verschiedener Probanden war, dass die Verteilung der Instrumente automatisch die Aufmerksamkeit beim Zuhören fördert.

5.3.1 Klassik

Der Mehrwert von 3D im Klassik-Beispiel fällt, wie in vorangegangenen Kapiteln beschrieben, vergleichsweise gering aus. Zudem handelt es sich hierbei um das einzige Hörbeispiel, bei dem die Bewertung der emotionalen Erregung nicht höher als die des Gefallens ist. Eine interessante Tatsache ist auch die allgemein höhere Bewertung des Gefallens (Stereo und 3D) durch Experten. Dies deckt sich mit den Forschungsergebnissen von *Elvers, Omigie et al.* (2015), die bei Expertenhörern eine deutlich höhere Akzeptanz für anspruchsvolle Musik wie Klassik festgestellt haben [35]. Die älteren Hörer zeigen zudem im Gegensatz zur jüngeren Gruppe eine signifikante Steigerung beider Bewertungsitems durch 3D.

Die Auswertung der Kommentare zeigt eine klare Assoziation des Titels mit *Filmmusik* und *Walt Disney*. Die 3D-Version wird neben vielen *Emotionsbeschreibungen* von einigen als *weit* und *groß*, die Stereo-Version dagegen als *flach* und *schmal* beschrieben. Allerdings bestätigen auch einige Kommentare einen eher *kleinen Unterschied zwischen beiden Versionen*.

Dies lässt sich damit begründen, dass bei der Aufnahme klassischer Musik in der Regel versucht wird, den Höreindruck im Raum abzubilden. Die Bühne wird demnach mit ihrem Direktschall in der Front abgebildet [24, p. 23]. Von hinten erreicht den Hörer lediglich Diffusschall - ein Höreindruck der auch bei der Stereo-Wiedergabe durch Wandreflexionen entstehen kann. Zudem trägt der relativ unauffällige, aber dennoch realistische Raumschall der 3D-Version zum Ergebnis bei. Natürlich kann auch hier beispielsweise ein anderer Klangkörper oder Raum zu anderen Testergebnissen führen.

5.3.2 Jazz

Ein allgemein starker Mehrwert bewirkt die Wiedergabe in 3D am Jazz-Beispiel. Mit Ausnahme des Gefallens bei den reinen Musikhörern steigern sich alle Items für jeden Hörtypen durch 3D. Zudem steigert sich das Gefallen durch 3D für den Experten um einiges mehr als für den reinen Musikhörer.

Möglichweise ziehen die Experten das Gefallen mehr aus dem Klangerlebnis als andere Hörtypen. Dies bleibt allerdings nur eine Vermutung.

Der häufigste Kommentar bezieht sich in der 3D-Version auf das Gefühl, *mittendrin* zu sitzen bzw. *live* dabei zu sein. Zudem hat das Stück eine entspannende Wirkung und funktioniert laut Probanden gut bei dem Genre. Einige nehmen an, *sehr lange zuhören* zu können, da die Verteilung im Raum nach eigenen Angaben die *Aufmerksamkeit fördert*. Vereinzelt werden auch Stimmen laut, die die *Verteilung der Instrumente* aufgrund bestehender Hörgewohnheiten als *ungewohnt* empfinden.

Einige wenige Kommentare unter der Stereo-Version beschreiben diese als *spitz, anstrengend, zu frontal* und *flach*. Ein einzelner merkt an, dass der *Stereo-Hall* nach der 3D-Version *künstlich* klingt.

Zudem bezieht sich ein Kommentar auf das *Solo-Saxophon*, das *nicht mehr mittig* zu sein scheint. Die vermeintliche Verschiebung des Instrumentes liegt allerdings im Gesetz der ersten Wellenfront begründet. Vereinfacht besagt dies, dass der Richtungseindruck durch den zuerst eintreffenden Schall bestimmt wird [17, p. 131]. Die Position des Saxophons wird in Stereo über eine Phantomschallquelle in der Mitte gebildet. Sitzt der Hörer nicht auf der Mittelachse, kommt der Schall des näheren Lautsprechers früher am Gehör an und wird auch näher an ihm lokalisiert. Durch den in 3D verwendeten Center-Lautsprecher wird das Saxophon unabhängig von der Sitzposition auch dort lokalisiert. Derartige Kommentare kamen auch in den weiteren Hörbeispielen auf. Dies unterstreicht die Theorie, dass Realschallquellen zur maximalen Ausdehnung des Sweetspots beitragen. [24, p. 23]

5.3.3 Soul, Funk & Blues

Neben der klaren Steigerung der emotionalen Erregung durch 3D, steigert sich das Gefallen in der Gesamtauswertung knapp nicht signifikant. Eine möglicher Grund dafür ist die im Vergleich zu anderen Musikbeispielen positive Klangbewertung der Stereo-Version (*toller Klang, schöne Räumlichkeit* und *groovig*), wobei auch die für Stereo üblichen negativen Bemerkungen fallen.

Die Gruppenauswertung zeigt einzig eine Steigerung der emotionalen Erregung bei Hobbymusikern und Experten. Zudem gaben die weiblichen Versuchsteilnehmer ein wesentlich höheres Gefallen bei Wiedergabe in 3D an.

Die in der Produktionsbeschreibung (siehe Kapitel 3.2.3) erwähnte *Panning-Fahrt der Stimmen* wurde in den Kommentaren am häufigsten positiv angemerkt, zwei Personen empfanden dies dagegen als *zu verspielt*. Auch in der 3D-Version wird der *Klang gelobt*, dagegen aber auch *gute Umhüllung* und *Gänsehaut* angemerkt. Eine Bemerkung, die es bei diesem Stück auf den Punkt bringt, beschreibt die 3D-Version als *szenischer*. Zudem wird die 3D-Version vereinzelt als *weniger tight*, also rhythmisch ungenauer empfunden. Dies liegt in der klaren Separation der Instrumente begründet. In Stereo überlagern sich diese und vor allem die Bläsersätze verschwimmen zu einem Klangkörper.

5.3.4 Rock

Zur Einordnung des Hörbeispiels ist die Beliebtheit der Musikrichtung zu erwähnen, die sich weit über dem Niveau der Gefallen-Bewertungen befindet (beliebteste Musikrichtung). Das Stück deckt sich also weniger mit dem Rockmusik-Ideal der Probanden. Das Rock-Beispiel erfährt dennoch in

Gefallen und emotionaler Erregung einen klaren Mehrwert durch 3D. In Bezug auf die Hörtypen dagegen ausschließlich für die emotionale Erregung bei den Experten. Für die ältere Hörgruppe steigert sich zum Gefallen auch die emotionale Erregung.

Spitzenreiter in den Kommentaren zur Stereo-Version ist neben der Bemerkung anstrengend, nervig und kantig auch der Konsens, dass das Stück zu überladen sei und ein Chaos durch viele Instrumente herrsche. Der 3D-Version wird dagegen mehr Platz für einzelne Elemente zugeschrieben. Für die einen besteht kein großer Mehrwert, für die anderen funktioniert es in 3D viel besser.

Eindrücklich zeigt sich auch in den Kommentaren die durch Positionierung eines Elements auf dem Lautsprecher entstehende Wirkung einer Realschallquelle. Die somit maximale Lokalisationsschärfe [24, p. 23] beschreiben die Hörer über das Gefühl, dass der *Sänger vor ihnen steht*.

5.3.5 Metal

Ein eindrücklicher Mehrwert durch Wiedergabe in 3D zeigt das Metal-Beispiel auf. Sowohl Gefallen, als auch emotionale Erregung steigern sich signifikant. Mit einem Anstieg von durchschnittlich mehr als zwei Punkten, führt Metal die Spitzengruppe der Mittelwertsdifferenz in der emotionalen Erregung an. Auch der Zuwachs des Gefallens liegt mit EDM im Spitzenbereich. Interessant ist auch, dass Metal in Bezug auf Alter und Hörtyp, mit Ausnahme des Gefallens beim reinen Musikhörer, in jedem Punkt einen Mehrwert erfährt.

Die Steigerung der emotionalen Erregung ist zudem bei männlichen Versuchsteilnehmern signifikant höher. Eine mögliche Erklärung liefert Abbildung 27. Sie zeigt, dass Metal bei weiblichen Teilnehmern deutlich unbeliebter ist. Findet eine Person also musikalisch keinen Zugang zu einer Musikrichtung, ist auch der relative Mehrwert in der emotionalen Erregung geringer. Im Umkehrschluss kann dies bedeuten: Möchte man beispielsweise einen Kunden vom emotionalen Mehrwert von 3D überzeugen, ist es ratsam, dies mit einer ihn ansprechenden Musikrichtung zu tun.

In den Kommentaren wird die Stereo-Version als anstrengend, nervig, kantig, hart, direkt und dicht beschrieben. Dagegen fällt bei dem Stereo-Beispiel aber auch öfter die Beschreibung druckvoll als beim 3D Beispiel. Dies besticht wiederum durch das Gefühl mittendrin zu sitzen. Die Versuchsteilnehmer beschreiben den Höreindruck als beeindruckend und merken einen tollen, transparenten und angenehmeren Klang an. Zudem findet die in Kapitel 3.2.5 angemerkte, kurzzeitige Benutzung der VOG bei einigen Probanden Anklang und führt sogar zu Gänsehaut. Eindrücklich beschreibt ein Proband den Höreindruck in 3D mit mächtig, ohne zu erdrücken.

5.3.6 Pop

Die S, F & B- und Pop-Beispiele weisen starke Ähnlichkeiten in ihren Bewertungen auf. Auch im Pop-Beispiel steigert sich selbst in der Gesamtbewertung lediglich die emotionale Erregung durch 3D signifikant. Gerade hier unterstützen die Bemerkungen zur Stereo-Version die bereits im S, F & B-Beispiel aufgeführte Theorie (siehe Kapitel 5.3.3). Die meisten Kommentare bemerken einen klaren, natürlichen, vollen Klang. Das Gehörte ist angenehm, macht Spaß und klingt nach Radio. Funktioniert also der Stereo-Mix für die Versuchspersonen entsprechend gut, steigert sich das Gefallen durch 3D nicht signifikant. Dagegen erfährt die emotionale Erregung eine klare Steigerung. Neben den, mit ähnlichen Attributen beschriebenen Klangeigenschaften der 3D-Version, gehen die Versuchsteilnehmer vermehrt auf die hohe emotionale Wirkung des Beispiels ein. Drei Versuchspersonen bemerken eine Gänsehaut. Die emotionale Wirkung wird zum einen mit der gefühlt starken Nähe zur Sängerin und der Umhüllung begründet.

Die Gruppenauswertung zeigt einzig bei Experten einen signifikanten, emotionalen Mehrwert. Zudem zeichnet sich eine tendenziell höhere Akzeptanz von Popmusik bei weiblichen Probanden ab, wenn auch nicht signifikant.

5.3.7 EDM

Der Mehrwert von 3D ist im EDM-Beispiel sowohl am Gefallen, als auch an der emotionalen Erregung nicht abzustreiten. Zudem zeigt das Bespiel als einziges Stück eine signifikante Steigerung durch 3D in beiden Items für alle Hörtypen. Erwähnenswert ist auch die nahezu signifikant höhere Beliebtheit der Musikrichtung bei jüngeren Versuchsteilnehmern. In Hinblick auf die Bewertung des Gefallens des eigentlichen Hörbeispiels, wird die Mittelwertsdifferenz zwischen den Altersgruppen dagegen gering.

In den Kommentaren zum 3D-Beispiel wird am häufigsten schönes, kreatives Panning, Umhüllung und der durch die Musik geförderte Bewegungsdrang aufgezählt. Ebenso bestätigen einige Versuchspersonen die Funktionalität des Stücks in 3D, einen detailreichen, transparenten Klang und eine schöne Verteilung des Glockenspiels im Raum. Im Vergleich zur Stereo-Version wird der Titel als interessanter, spannender und mitreißender beschrieben. Ein Teilnehmer schreibt, dass das Stück nicht nur zum Tanzen, sondern auch zum Zuhören anregt, zwei andere sprechen von einer Traumreise in eine neue Welt. Auch die Kommentare der Stereo-Version sind neben den wenigen für Stereo üblichen Anmerkungen weitestgehend positiv.

6. Fazit und Zukunft

Abschließend kann zusammengefasst werden, dass 3D-Audio in allen Musikrichtungen seine Berechtigung hat. Sowohl Gefallen als auch emotionale Erregung erfahren vielfach einen signifikanten Mehrwert durch das Format gegenüber Stereo. Besteht ein ansprechender Stereo-Mix, hält sich die Steigerung des Gefallens jedoch in Grenzen.

Doch gerade emotional kann die Wiedergabe in 3D einen enormen Mehrwert bedeuten. Dies zeigen zusätzlich zur quantitativen Auswertung auch die viel beschriebenen Gefühlsempfindungen der Probanden in Bezug auf die 3D-Versionen. Die drei bestbewerteten Stücke zeigen zudem, dass dies unabhängig der übermittelten Emotionen geschieht. Während das Metal-Beispiel von einigen als aufregend beschrieben wird, fällt bei EDM das Stichwort *Traumreise*. Das Jazz-Stück beschreiben viele als *entspannt* und *locker*.

Die weitere qualitative Analyse der Kommentare zeigt die von den Probanden geschätzte *Umhüllung* als Gefühl *mittendrin* zu sitzen und *live* dabei zu sein. Ebenso geschätzt wird der durch 3D mögliche *volle, transparente Klang*.

Die Auswertung der Hörtypen zeigt zudem, dass sich die Experten wesentlich einiger über den Mehrwert von 3D sind. Zwar deutet sich der Mehrwert auch bei den anderen Hörtypen in den meisten Fällen an (siehe Abbildung 24), doch kann dieser aufgrund der niedrigen Stichprobenzahl nach Aufteilung in Hörtypen oft nicht signifikant nachgewiesen werden. Umso aussagekräftiger sind wiederum signifikante Ergebnisse bei den Nicht-Experten, die den Großteil der späteren Kunden darstellen, also das Consumer Panel bilden [34, p. 13].

Einen deutlichen Mehrwert durch 3D erfährt das EDM-Beispiel, das einzig auch bei allen Hörtypen signifikante Anstiege durch 3D aufzeigt. Mögliches Anwendungsgebiet in der, gerade bei jungen Personen beliebten, Musikrichtung, ist die Implementierung in Diskotheken. Dies stellt eine Möglichkeit dar, ohne die hohen Anschaffungskosten für ein eigenes 3D-Audio-System in den Genuss des Formates zu kommen. In einigen Fällen wurde dies schon in die Praxis umgesetzt [39, 40].

Aktuelle Entwicklungen im Beschallungsbereich zeigen zudem die Möglichkeiten großräumiger 3D-Beschallungen auf. Schon jetzt werden Live-Konzerte in 3D durchgeführt [41]. Denkt man an eine Band wie Rammstein, die in Sachen Pyrotechnik keine Kosten und Mühen scheut, könnte die 3D-Beschallung in Anbetracht des emotionalen Mehrwerts bei Metal das Erlebnis noch um einiges intensiver machen.

Besteht die Bereitschaft, in ein 3D-System für den Heimgebrauch zu investieren, kann dagegen die gerade bei Hi-Fi-Enthusiasten beliebte Musikrichtung Jazz [42] einen großen Mehrwert durch 3D erfahren. Aber auch Klassik kann von der realistischen Abbildung des Formates profitieren. Hier ist auch mit der Einführung der *Sennheiser AMBEO Soundbar* im Mai 2019 ein weiterer großer Schritt des Formates in Richtung Konsument zu erwarten.

Damit sich die Anschaffung eines 3D-Systems für den Kunden allerdings lohnt, müssen zuerst genügend Inhalte bestehen. Möchte der Kunde native, kommerziell erhältliche 3D-Mischungen hören, stehen bislang nur wenige Titel zur Auswahl. Zwar bestehen Upmixing-Tools zur Aufbereitung gewöhnlichen Stereo-Materials hin zu 3D-Sound, doch fällt dann die so wichtige, kreative Komponente der eigentlichen Mischung in 3D weg. Auch fällt es schwer zu glauben, mit stereophonem Ausgangsmaterial die klanglichen und dynamischen Möglichkeiten von 3D auszureizen.

Weiter zeigt diese Arbeit, dass Stücke nicht zwingend für 3D produziert werden müssen, um für das Format zu funktionieren. Die Beispiele für Jazz, Rock, Metal und Pop wurden ursprünglich ohne Hintergedanken an eine 3D-Mischung produziert. Es können also auch bereits bestehende Aufnahmen mehrwertbringend in 3D umgesetzt werden. Zudem lässt sich mit den bereits editierten, fertigen Multitracks einer Produktion erstaunlich schnell eine funktionierende 3D-Mischung erstellen [43, p. 24].

Auch in der Filmindustrie, in der das Format mit Implementierung der Systeme in Kinos bereits recht weit vorangeschritten ist [4], ist großes Potential zu vermuten. Die Verwendung von 3D-Audio für Musikmischungen kann gerade in Kinofilmen aufgrund der hohen emotionalen Wirkung einen Mehrwert bringen.

3D-Audio ist und bleibt demnach ein Feld, dessen Entwicklung in Anbetracht technischer Neuerungen und der Akzeptanz bei Konsumenten in nächster Zeit spannend bleiben wird. In der Musikproduktion zeigt sich unumstritten ein Mehrwert durch das Format. Ob sich dies auch beim Endverbraucher durchsetzt, wird die Zukunft zeigen.

Anhang

Daten-CD

Die an der Rückseite der vorliegenden Arbeit angebrachte Daten-CD beinhaltet die jeweils nicht entzerrten Mono- und Stereo-Lautsprecherfeeds des Hörversuchs für die Stereo- und 3D-Version.

Das erste Hörbeispiel startet nach 20 Sekunden Vorlauf.

	Start	Hörbeispiel
1	0:20	Klassik: PC-24, ein musikalischer Alpenflug
2	2:20	Jazz: Pennsylvania 6-5000
3	4:20	Soul, Funk & Blues: Love On Top
4	6:20	Rock: Rallye
5	8:20	Metal: Creeping Decay
6	10:20	Pop: Change The World
7	12:20	EDM: Rain On Your Skin

Zudem sind auf der Daten-CD die zur Entzerrung verwendeten Impulsantworten zu finden.

Unterlagen zum Hörversuch

Anhang 1: Fragebogen zum Hörversuch	68
Anhang 2: Beiblatt zum Hörversuch	72
Anhang 3: Notizen des Versuchsleiters zum Hörversuch	73
Anhang 4: Wiedergabereihenfolge zum Hörversuch	74

Fragebogen zum Hörversuch der Bachelorarbeit:

Mehrwert von 3D-Audio in der Musikproduktion hinsichtlich Musikrichtung von Yannik Weber

Mit der Teilnahme an diesem Hörversuch stimmen Sie der anonymen Auswertung der hier angegebenen Daten zu Forschungszwecken zu. Darüber hinaus bestätigen Sie, über ein uneingeschränktes Hörvermögen zu verfügen.

1.	Aiter:	

- 2. Geschlecht: _____
- 3. Welche Beschreibung trifft auf Sie zu? (Einfachauswahl)
 - o reiner Musikhörer
 - o Hobbymusiker (regelmäßiges Musizieren)
 - o professionelle Beschäftigung mit Musik oder Tontechnik
- 4. Hören Sie heute zum ersten Mal Musik auf einem 3D-Audio System?

0	Ja	0	Nein
•	3 4	_	110111

5. Wie gern hören Sie die folgenden Musikrichtungen?

Musikrichtung	ungern	0	1	2	3	4	5	6	7	8	9	10	sehr gern
	Klassik	0	0	0	0	0	0	0	0	0	0	0	
	Jazz	0	0	0	0	0	0	0	0	0	0	0	
Soul, Fu	ınk & Blues	0	0	0	0	0	0	0	0	0	0	0	
	Rock	0	0	0	0	0	0	0	0	0	0	0	
	Metal	0	0	0	0	0	0	0	0	0	0	0	
	Pop	0	0	0	0	0	0	0	0	0	0	0	
EDM (Electronic Da	nce Music)	0	0	0	0	0	0	0	0	0	0	0	

Anhang 1: Fragebogen zum Hörversuch

2

6. Hörversuch

Hörbeispiel 1		0	1	2	3	4	5	6	7	8	9	10	
a)	gar nicht	0	0	0	0	0	0	0	0	0	0	0	sehr gut
b) c)	gar nicht	0	0	0	0	0	0	0	0	0	0	0	sehr stark

Hörbeispiel 2		0	1	2	3	4	5	6	7	8	9	10	
a)	gar nicht	0	0	0	0	0	0	0	0	0	0	0	sehr gut
b)	gar nicht	0	0	0	0	0	0	0	0	0	0	0	sehr stark
c)													

Hörbeispiel 3		0	1	2	3	4	5	6	7	8	9	10	
a)	gar nicht	0	0	0	0	0	0	0	0	0	0	0	sehr gut
b)	gar nicht	0	0	0	0	0	0	0	0	0	0	0	sehr stark
c)													

Hörbeispiel 4		0	1	2	3	4	5	6	7	8	9	10	
a)	gar nicht	0	0	0	0	0	0	0	0	0	0	0	sehr gut
b)	gar nicht	0	0	0	0	0	0	0	0	0	0	0	sehr stark
c)													

Hörbeispiel 5		0	1	2	3	4	5	6	7	8	9	10	
a) ga	ar nicht	0	0	0	0	0	0	0	0	0	0	0	sehr gut
b) ga	ar nicht	0	0	0	0	0	0	0	0	0	0	0	sehr stark

Fragen zum Hörversuch:

- a) Wie sehr hat Ihnen das Stück gefallen?
- b) Wie sehr hat Sie das Stück emotional bewegt?
- c) eigene Bemerkungen, Beobachtungen und Kommentare

3 Hörbeispiel 6 1 gar nicht sehr gut a) 0 0 0 0 0 0 0 0 b) gar nicht sehr stark c) Hörbeispiel 7 a) gar nicht 0 0 0 0 0 sehr gut 0 0 0 0 b) gar nicht o sehr stark c) Hörbeispiel 8 10 gar nicht sehr gut b) gar nicht o sehr stark c) Hörbeispiel 9 0 1 2 3 9 gar nicht 0 0 0 0 sehr gut a) 0 0 0 0 0 0 0 b) c) Hörbeispiel 10 0 a) gar nicht 0 sehr gut 0 0 0 0 0 0 0 0 0 0 b) gar nicht 0 sehr stark 0 0 0 0 0 0 0 0 c) Fragen zum Hörversuch: a) Wie sehr hat Ihnen das Stück gefallen? b) Wie sehr hat Sie das Stück emotional bewegt?

c) eigene Bemerkungen, Beobachtungen und Kommentare

4 Hörbeispiel 11 0 1 2 gar nicht 0 sehr gut 0 0 0 0 0 0 b) gar nicht sehr stark c) Hörbeispiel 12 gar nicht o 0 0 0 0 0 0 0 0 sehr gut 0 b) gar nicht o sehr stark c) Hörbeispiel 13 10 gar nicht o sehr gut b) gar nicht o sehr stark c) 0 1 Hörbeispiel 14 2 3 9 10 gar nicht o 0 0 0 0 0 0 0 0 sehr gut 0 0 b) sehr stark c)

Fragen zum Hörversuch:

- a) Wie sehr hat Ihnen das Stück gefallen?
- b) Wie sehr hat Sie das Stück emotional bewegt?
- c) eigene Bemerkungen, Beobachtungen und Kommentare

Beiblatt zum Hörversuch der Bachelorarbeit:

Mehrwert von 3D-Audio in der Musikproduktion hinsichtlich Musikrichtung von Yannik Weber

Nun startet der eigentliche Hörversuch.

Für den Hörversuch wurden 7 Stücke verschiedener Musikrichtungen ausgewählt, für die jeweils ein 1-minütiger Ausschnitt festgelegt wurde. Für diesen Ausschnitt bestehen jeweils eine geläufige Stereo- und eine 3D-Version. Der Test umfasst somit 14 kurze Hörbeispiele aus 7 verschiedenen Stücken. Die Wiedergabereihenfolge der Hörbeispiele wurde vor dem Test rein zufällig bestimmt, es ist also beispielsweise gut möglich, dass die zwei Versionen eines Stückes nicht nacheinander wiedergegeben werden. Ob es sich bei dem gespielten Beispiel um Stereo- oder 3D-Inhalt handelt wird nicht angezeigt.

Lassen Sie sich bitte während des Hörens voll und ganz auf die Musik ein - wenn förderlich auch gerne mit geschlossenen Augen.



Achten Sie beim Hören auf den Gefallen, den Sie am Beispiel finden

Wie sehr hat Ihnen das Stück gefallen?



und die Stärke der in Ihnen ausgelösten Emotionen.

Wie sehr hat Sie das Stück emotional bewegt?

Adjektive zur Beschreibung musikalischer Emotionen lassen sich in die Kategorien ernsthaft, traurig, sentimental, ruhig, humorvoll, fröhlich, aufregend und majestätisch einteilen. Versuchen Sie diese also bitte möglichst ganzheitlich zu betrachten.

Zur Unterscheidung von a) und b):

Ein Stück kann zum Beispiel eine sehr traurige Stimmung übermitteln und damit stark emotional bewegen. Dies muss dem Hörer aber nicht zwingend gefallen.



In den freien Zeilen ist bei Bedarf Platz für eigene Bemerkungen. Dort ist von 'klingt total blechern' bis 'Gänsehaut' alles erlaubt.

eigene Bemerkungen, Beobachtungen und Kommentare

Zwischen den Hörbeispielen gibt es kurze Pausen, in denen der jeweilige Abschnitt im Fragebogen ausgefüllt werden kann.

Die exakten Fragestellungen sind zusätzlich auf jeder Seite in der Fußzeile ausgeführt.

Mehrwei	rt von 3D-Au	dio in der Musi	kproduktion	hinsichtlich Musikrichtung von Yannik Webe
Datum	Uhrzeit	Durchg. Nr.	Teiln. Nr.	Name
×-				
7				
8				
32 3				

Anhang 3: Notizen des Versuchsleiters zum Hörversuch

Mehrwert													
Durchg.N 1	. 1	2 9	3 9	4 14	5	6	7 4	8 4	9 7	10 13	11 5	12 3	
2	4	10	12	1	3	4	9	9	3	12	3	7	
3	1	5	11	3	7	6	6	13	10	2	12	9	
4	7	6	2	4	1	12	2	11	4	1	13	2	
5	5	4	13	8	14	8	5	12	12	3	11	12	
6	13	14	8	2	6	11	14	8	1	6	10	10	
7	8	1	4	7	4	14	10	3	6	11	9	4	
8	10	7	1	9	5	13	13	2	13	9	2	5	
9	14	8	3	12	9	7	8	6	5	7	4	8	
10	6	2	14	6	12	10	12	7	11	4	8	1	
11		3	10	13	2	3	3	5	14	8	6 7	6	
12 13		12 13	5 7	11 5	8 10	9	1 7	14 1	9 8	5 10	14	13 14	
14	200000	11	6	10	11	5	11	10	2	14	1	11	
Durchg.N	. 13	14	15	16	17	18	19	20	21	22	23	24	
1	5	10	9	14	10	8	4	9	3	5	4	10	
2	1	2	2	13	7	11	10	8	14	8	11	11	
3	12	8	7	12	8	14	13	14	13	4	9	5	
4	11	1	4	10	9	12	5	12	12	13	3	9	
5	2	14	13	1	13 5	9	7 11	11	9 6	14	10 7	8	
7	4	3 9	1 5	6	5 14	13 10	9	2 5	5	2 10	12	14 12	
8	14	5	6	3	4	2	1	6	4	1	2	7	
9	8	11	8	8	6	4	14	1	7	12	13	2	
10	3	7	11	11	2	7	2	4	1	6	1	4	
11	13	12	14	7	1	6	8	7	11	3	5	1	
12	7	4	3	5	12	3	12	10	10	11	14	6	
13	6	6	10	4	3	5	3	3	8	9	6	13	
14	10	13	12	9	11	1	6	13	2	7	8	3	
Durchg.N	-	26	27	28	29	30	31	32	33	34	35	36	
1	13	11	3	8	1	12	10	13	6	1	11	5	
2	14 7	2 5	1 2	10 6	10 11	11 5	1 7	8	13 14	10 14	13 12	12 9	
4	2	4	2 11	5	2	5 14	11	9	5	3	4	2	
5	10	10	4	12	8	1	14	2	7	6	1	11	
6	11	12	8	4	7	10	5	5	12	4	3	6	
7	1	13	9	13	5	3	6	6	2	12	10	8	
8	5	8	5	2	6	8	4	11	1	8	14	14	
9	12	14	10	1	4	6	2	10	8	11	6	4	
10	4	3	6	3	3	13	13	4	4	7	2	3	
11	8	6	14	14	14	2	12	14	10	13	8	10	
12	6	9	13	9	13	4	3	1	11	2	5	1	
13	3	7	12	11	9	9	8	7	3	9	9	7	
14	9	1	7	7	12	7	9	12	9	5	7	13	

Anhang 4: Wiedergabereihenfolge zum Hörversuch

Abbildungsverzeichnis

Abbildung 1: Dimensionen der verschiedenen Tonformate	13
Abbildung 2: 3D-Audio-Filmtonstudio U48 der HdM Stuttgart (Foto: S. Hermentin)	15
Abbildung 3: Auro-13.1 (Grafik: J. Frey & Y. Weber)	16
Abbildung 4: Atmos 9.1.4 (Grafik: J. Frey & Y. Weber)	16
Abbildung 5: Ebenen zur Beschreibung der Lautsprecherpositionen (Grafik: Sengpiel)	16
Abbildung 6: Grundfläche des Filmtonstudios (Grafik: J. Frey)	17
Abbildung 7: Frequenzgangmessung bei 0° und 30° Mikrofonausrichtung	21
Abbildung 8: Messaufbau und -punkte	21
Abbildung 9: Einzelmessungen des rechten Lautsprechers an verschiedenen Messpunkten	23
Abbildung 10: Einzelmessung am Sweetspot und gemittelte Kurve aus allen Messungen	23
Abbildung 11: von REW erstellte Korrekturfilter	24
Abbildung 12: Filterkurve des rechten Lautsprechers in rePhase	25
Abbildung 13: stark vergrößerte, bereits beschnittene IR-Waveforms der Lautsprecher 1-16	26
Abbildung 14: Dolby Atmos 5.1.4 (Grafik: Dolby)	27
Abbildung 15: Auro-9.1 (oben) & 10.1 (unten) (Grafik: Auro)	27
Abbildung 16: die verschiedenen Layer bei Auro-3D (Grafik: Auro)	28
Abbildung 17: Spatial Audio Designer von New Audio Technology (Screenshot: NAT)	29
Abbildung 18: Max for Lives Surround Panner im Rock-Beispiel	30
Abbildung 19: Verteilung der Instrumente im Jazz-Beispiel (Grafik: J. Frey & Y. Weber)	35
Abbildung 20: Plakat von 3Disco (Grafik: J. Vitt & Y. Weber)	43
Abbildung 21: Legende für Diagramme (links) und p-Wert-Tabellen (rechts)	50
Abbildung 22: Mittelwerte der Beliebtheit der Musikrichtungen, des Gefallens und der emotion	nalen
Erregung der Hörbeispiele in Stereo und 3D; Fehlerbalken: ± 1 Standardfehler	51
Abbildung 23: Mittelwertsdifferenz des Gefallens (rot) und der emotionalen Erregung (blau) zwis	chen
3D- und Stereo-Bewertung; Fehlerbalken: ± 1 Standardfehler	52
Abbildung 24: Bewertungsmittelwerte in Abhängigkeit des Hörtyps; Tripletts v.l.: reiner Musikh	örer,
Hobbymusiker, prof. Musiker/Tontechniker; im Sinne der Übersichtlichkeit keine Fehlerbalken	54
Abbildung 25: Vorerfahrung mit Musik in 3D in Abhängigkeit des Hörtyps; rot: keine Vorerfah	rung,
blau: Vorerfahrung	55
Abbildung 26: Bewertungsmittelwerte in Abhängigkeit des Alters; paarweise Gruppierung: links 1	18-23
J., rechts 24-59 J.; Fehlerbalken: ± 1 Standardfehler	56

Abbildung 27: Bewertungsmittelwerte in Abhängigkeit des Geschlechts; paarweise Gruppierung: lin	nks
männlich, rechts weiblich; Fehlerbalken: ± 1 Standardfehler	57
Abbildung 28: Wortwolken der häufigsten Kommentare in Bezug auf das Wiedergabeformat	59

Abbildungsquellen:

Abbildung Abbildung 2	Quelle Foto: Simon Hermentin
Abbildung 3	Grafik: Johannes Frey (3D-Modell) und Yannik Weber (Beschriftung)
Abbildung 4	Grafik: Johannes Frey (3D-Modell) und Yannik Weber (Beschriftung)
Abbildung 5	http://www.sengpielaudio.com/RichtungshoerenInDreiEbenen.pdf
Abbildung 6	
Abbildung 14	https://www.dolby.com/us/en/technologies/dolby-atmos/dolby-atmos-home-
	theater-installation-guidelines.pdf?pdf=Installation-Guidelines
Abbildung 15	Wilfried van Baelen, "Der Auro-3D® Upmixer "AuroMatic™"," <i>VDT-Magazin</i> , no. 2, 2016
Abbildung 16	http://www.auro3d-heimkino.de/auro3d-technik-special/auro3d-funktionsprinzip.html
Abbildung 17 Abbildung 19 Abbildung 20	https://newaudiotechnology.com/products/spatial-audio-designer/ Grafik: Johannes Frey (3D-Modell) und Yannik Weber (Beschriftung) Grafik: Julian Vitt (3D-Modell) und Yannik Weber (Beschriftung)

Tabellenverzeichnis

Tabelle 1: Lautsprecherbestückung des 3D-Audio-Filmtonstudios der HdM
Tabelle 2: für den Hörversuch verwendete Musiktitel; []: Format des Hörversuchsausschnitts 32
Tabelle 3: Mikrofonierung von PC-24
Tabelle 4: p-Werte der Wilcoxon-Tests: paarweiser Vergleich der Bewertungen in Stereo und 3D 51
Tabelle 5: p-Werte der Wilcoxon-Tests: paarweiser Vergleich der Mittelwertsdifferenz des Gefallens
und der emotionalen Erregung
Tabelle 6: p-Werte der Wilcoxon-Tests: paarweiser Vergleich der Beliebtheit und des Gefallens in
Stereo
Tabelle 7: signifikante p-Werte der Kruskal-Wallis-Tests: Prüfung auf unterschiedliche Tendenzen in
der Bewertung der gleichen Variablen in Abhängigkeit des Hörtyps
Tabelle 8: p-Werte der Wilcoxon-Tests: paarweiser Vergleich der Bewertungen in Stereo und 3D in
Abhängigkeit des Hörtyps
Tabelle 9: p-Werte der Wilcoxon-Tests: paarweiser Vergleich der Bewertungen in Stereo und 3D in
Abhängigkeit des Alters
Tabelle 10: signifikante p-Werte der Mann-Whitney-U-Tests: Prüfung auf unterschiedliche Tendenzen
in der Bewertung der gleichen Variablen in Abhängigkeit des Geschlechts58

Literaturverzeichnis

- [1] A. Friesecke, *Die audio-Enzyklopädie: Ein Nachschlagewerk für Tontechniker*, 2nd ed. Berlin, Germany, Boston, Massachusetts: De Gruyter Saur, 2014.
- [2] WDR, Stichtag: 22. März 2009 Vor 45 Jahren: WDR-Hörfunk startet Stereo-Ausstrahlungen. Accessed on: Apr. 01 2019.
- [3] Wilfried van Baelen, "Der Auro-3D® Upmixer "AuroMatic™"," VDT-Magazin, no. 2, 2016.
- [4] Teufel Blog Redaktion, ENTERTAINMENT, FILM UND MUSIK: Übersicht: Dolby Atmos Kinos in Deutschland. [Online] Available: https://blog.teufel.de/uebersicht-dolby-atmos-kinos-in-deutschland/. Accessed on: Apr. 02 2019.
- [5] Sennheiser electronic GmbH & Co. KG, *AMBEO Soundbar: Unerhört guter Klang.* [Online] Available: https://de-de.sennheiser.com/ambeo-soundbar. Accessed on: Apr. 01 2019.
- [6] BR WISSEN, Musikforschung: Musik und Emotion. [Online] Available: https://www.br.de/themen/wissen/musik-forschung-psychologie-100.html. Accessed on: Apr. 02 2019.
- [7] H. Schramm and R. Kopiez, "3.4 Die alltägliche Nutzung von Musik," in *Rowohlts Enzyklopädie*, vol. 55661, *Musikpsychologie: Das neue Handbuch*, H. Bruhn, R. Kopiez, and A. C. Lehmann, Eds., 3rd ed., Reinbek bei Hamburg: Rowohlt-Taschenbuch-Verl., 2011, pp. 253–265.
- [8] Elif Tekin Gurgen, "Social and Emotional Function of Musical Listening: Reasons for Listening to Music," *Eurasian Journal of Educational Research*, no. 66, pp. 229–242, http://dx.doi.org/10.14689/ejer.2016.66.13, 2016.
- [9] Ephraim Hahn, "Evaluation des musikalischen Mehrwerts von 3D Audio in der klassischen Musikproduktion," Bachelorarbeit, Erich-Thienhaus-Institut, Hochschule für Musik, Detmold, 2017.
- [10] C. Guastavino and B. F. G. Katz, "Perceptual evaluation of multi-dimensional spatial audio reproduction," *The Journal of the Acoustical Society of America*, vol. 116, no. 2, pp. 1105–1115, 2004.
- [11] Marvin Pfeifer, "Evaluierung des produktionstechnischen Potenzials und konsumentenorientierten Mehrwerts von Binaural-3D-Audio im Kontext von Popmusik," Masterarbeit, Hochschule der Medien, Stuttgart, 2018.

- [12] A. Silzle, S. George, E. A. P. Habets, and T. Bachmann, "Fraunhofer Institute for Integrated Circuits IIS, Erlangen, Germany: Investigation on the Quality of 3D Sound Reproduction," *Proceedings of ICSA 2011*, pp. 334–341, 2011.
- [13] AURO TECHNOLOGIES NV, *Auro-3D® Home Theater Setup: Installation Guidelines.* [Online] Available: https://www.auro-3d.com/wp-content/uploads/documents/Auro-3D-Home-Theater-Setup-Guidelines_lores.pdf. Accessed on: Jan. 22 2019.
- [14] Dolby Laboratories Inc., *Dolby Atmos® Home Theater Installation Guidelines*. [Online] Available: https://www.dolby.com/us/en/technologies/dolby-atmos/dolby-atmos-home-theater-installation-guidelines.pdf?pdf=Installation-Guidelines. Accessed on: Jan. 22 2019.
- [15] Dolby Laboratories Inc., What is the LFE channel? [Online] Available: https://www.dolby.com/uploadedFiles/Assets/US/Doc/Professional/38_LFE.pdf. Accessed on: Jan. 11 2019.
- [16] Room EQ Wizard, Help Signal Generator: Pink Noise. [Online] Available: https://www.roomeqwizard.com/help/help_en-GB/html/siggen.html. Accessed on: Jan. 11 2019.
- [17] M. Dickreiter, Handbuch der Tonstudiotechnik, 8th ed. Berlin: De Gruyter, 2014.
- [18] Genelec, *Monitor Setup Guide*. [Online] Available: https://www.genelec.com/sites/default/files/media/Studio%20monitors/Catalogues/monitor_s etup_guide_2018_bbage125e.pdf. Accessed on: 15.01.19.
- [19] R. A. Katz, *Mastering audio: The art and the science*. Burlington, MA: Focal Press, Taylor & Francis Group, 2015.
- [20] Sonarworks, *About: Sonarworks story.* [Online] Available: https://www.sonarworks.com/about. Accessed on: Jan. 15 2019.
- [21] A. Friesecke, Studio-Akustik: Konzepte für besseren Klang; mit Checklisten, FAQs und Low Budget-Tipps; [inkl. CD mit Berechnungshilfen], 5th ed. Bergkirchen: PPVMedien, 2015.
- [22] Room EQ Wizard, *Help Analysis Preferences: Impulse Response Window Defaults.* [Online] Available: https://www.roomeqwizard.com/help/help_en-GB/html/analysis.html. Accessed on: Jan. 16 2019.

- [23] Surround Sound Info, Übersicht: Titel mit 3D-Sound: Musik. [Online] Available: https://surround-sound.info/3d-sound/titel/?15&musik&63&3&7&15&3&1&alle&0&3&0&1. Accessed on: Jan. 23 2019.
- [24] L. Nipkow, "Die Bedeutung von 3D bei Immersive Audio," VDT-Magazin, no. 1, pp. 22–25, 2019.
- [25] P. J. Rentfrow and S. D. Gosling, "The do re mi's of everyday life: The structure and personality correlates of music preferences," *Journal of Personality and Social Psychology*, vol. 84, no. 6, pp. 1236–1256, 2003.
- [26] Metal Hammer, *Heavy Metal*. [Online] Available: https://www.metal-hammer.de/genres/heavy-metal/. Accessed on: Jan. 23 2019.
- [27] Luzerner Zeitung, *Die Festival Strings heben ab mit dem PC-24*. [Online] Available: https://www.luzernerzeitung.ch/kultur/die-festival-strings-heben-ab-mit-dem-pc-24-ld.1029491. Accessed on: Jan. 24 2019.
- [28] Silent Work, *3D-Audio*. [Online] Available: http://silentwork.com/index.php/3d-audio. Accessed on: Jan. 24 2019.
- [29] Jochen Neuffer, *JOCHEN NEUFFER: JOCHEN NEUFFER TOBIAS BECKER BIGBAND «AUGMENTED REALITIY»*. [Online] Available: https://www.jochenneuffer.com/augmented-reality. Accessed on: Mar. 05 2019.
- [30] Dennis DeSantis et al., Ableton-Referenzhandbuch Version 9: für Windows und Mac OS, 2013.
- [31] Dipl.-Ing. Eberhard Sengpiel, *Kopfbezogene Übertragungsfunktion HRTF*. [Online] Available: http://www.sengpielaudio.com/KopfbezogeneUebertragungsfunktionHRTF.pdf. Accessed on: Jan. 28 2019.
- [32] L. Schmidt-Atzert, "Emotionspsychologie und Musik," in *Musikpädagogische Forschung*, K.-E. Behne, Ed., Laaber: Laaber-Verlag, 1980-1989, pp. 26–46.
- [33] J. Flecker, "Bedeutung von Musik für Emotionen und Verhaltensweisen von Menschen," in *Die Bedeutung von Musik für die Gestaltung von Markenpersönlichkeit*, J. Flecker, Ed., Wiesbaden: Springer Fachmedien Wiesbaden, 2014, pp. 9–72.
- [34] Wolfgang Ellermeier, Jürgen Hellbrück, Armin Kohlrausch, and Alfred Zeitler, *Kompendium zur Durchführung von Hörversuchen in Wissenschaft und industrieller Praxis.* [Online] Available:

- https://www.dega-akustik.de/fileadmin/dega-akustik.de/publikationen/Kompendium_Hoerversuche_2008.pdf. Accessed on: Jan. 25 2019.
- [35] P. Elvers, D. Omigie, W. Fuhrmann, and T. Fischinger, "Exploring the musical taste of expert listeners: musicology students reveal tendency toward omnivorous taste," (eng), *Frontiers in psychology*, vol. 6, p. 1252, 2015.
- [36] N. Döring and J. Bortz, Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016.
- [37] Universität Zürich, *UZH Methodenberatung Skalenniveau: 2. Hierarchie der Skalenniveaus.*[Online] Available: https://www.methodenberatung.uzh.ch/de/skalenniveau.html. Accessed on: Mar. 09 2019.
- [38] Nicolay Ketterer, Wie laut geht es noch?: Loudness War Interview mit Lautheitsforscher Rudi Ortner. [Online] Available: https://www.soundandrecording.de/tutorials/loudness-war-interview-mit-lautheitsforscher-rudi-ortner/. Accessed on: Apr. 04 2019.
- [39] Ableton, *4DSOUND:* Eine neue Dimension. [Online] Available: https://www.ableton.com/de/blog/4dsound-video-performance-stimming/. Accessed on: Apr. 06 2019.
- [40] Surround Sound Info, Weitere Disko-Einsätze für Dolby Atmos. [Online] Available: https://surround-sound.info/news/weitere-disko-einsaetze-fuer-dolby-atmos-2394/. Accessed on: Apr. 06 2019.
- [41] C. Bangert, *3D-Audio-Trends: Immersive Sound Systeme im Überblick.* [Online] Available: https://www.production-partner.de/story/3d-audio-trends-immersive-sound/#Dolby%20Atmos. Accessed on: Apr. 06 2019.
- [42] STEREO | Magazin für HiFi, High End & Musik, *Musik: Rezensionen und Musik-Playlisten*. [Online] Available: https://www.stereo.de/musik/. Accessed on: 06.04.19.
- [43] E. Wisse, "Michael Jackson hätte in 3D produziert," *VDT-Magazin*, vol. 33, no. 3, pp. 23–27, 2017.