

Bachelorarbeit

# Untersuchung von 3D-Audio Systemen und binauralen 3D-Kopfhörersimulationen

*Vergleich im Hinblick auf die Qualität einer authentischen Raumwiedergabe am Beispiel einer  
Livemusik-Audioaufzeichnung*

vorgelegt von Sascha Coronini

Matrikel-Nr.: 25743

Bearbeitungszeitraum: 06.11.2015 bis 05.02.2016

Hochschule der Medien

Nobelstraße 10

70569 Stuttgart

Studiengang Audiovisuelle Medien

Wintersemester 2015/2016

Erstprüfer: Prof. Oliver Curdt

Zweitprüfer: Dipl. Ing. Bernd Kugler

## Kurzfassung

Diese Bachelorarbeit befasst sich neben einer geschichtlichen Einführung in die Mehrkanaltechnik mit verschiedenen Möglichkeiten zur Wiedergabe und Produktion von 3D-Audio.

Es werden darüber hinaus eine Vielzahl verschiedener 3D-Audioformate abgehandelt und die Grundlagen von kanal- bzw. objektbasiertem 3D-Audio erklärt.

Außerdem werden eine Reihe von Softwareprogrammen vorgestellt, welche unter anderem als Panoramaregler<sup>1</sup> für 3D-Audio genutzt werden können. Das Hauptaugenmerk liegt allerdings in der Möglichkeit dieser Programme, 3D-Audiowiedergabesysteme unter der Verwendung von Kopfhörern zu simulieren.

Der praktische Teil der Arbeit stellt einen Hörversuch dar. Dieser befasst sich mit der Fragestellung, ob eine binaurale Simulation eines Musikkonzertes ähnlich gute räumliche und klangliche Eigenschaften besitzen kann, wie ein auf einer großen Anzahl von Lautsprechern basierendes 3D-Audiosystem.

## Abstract

This bachelor thesis considers, apart from a historical introduction into multichannel audio, several possibilities to play and produce 3D-Audio.

Furthermore it deals with various 3D-Audioformats and explains the principles of channel- and object-based audio.

Besides, a number of software programs will be introduced, which can be used as pan-pots in 3D-Audio. However, the focus lies on the simulation of 3D-Audio systems played on headphones.

The practical part of this work constitutes a listening experiment.

It deals with the question whether a binaural simulation of a musical concert can be competitive in spatial and sound qualities, with a 3D-Audio system, that is based on a large number of loudspeakers.

---

<sup>1</sup> Legt die Abbildungsrichtung eines Tonsignals zwischen den Lautsprechern fest. (vgl. Dickreiter, 2014b, S. 529)

## Vorwort

Die folgende Arbeit entstand im achten Fachsemester meines Studiums, in der Fachrichtung Audiovisuelle Medien an der Hochschule der Medien in Stuttgart.

An dieser Stelle möchte ich mich bei allen Personen bedanken, die mich bei der Erstellung dieser Arbeit unterstützt haben.

Ein besonderer Dank gilt meinem betreuenden Professor Oliver Curdt, welcher mir während der Arbeit stets mit gutem Rat zur Seite stand.

Einen besonderen Dank möchte ich auch den Mitarbeitern der Firma B&R Medientechnik aus Kürten und insbesondere dem Geschäftsführer Dipl. Ing. Bernd Kugler aussprechen, der es mir ermöglichte, einen Hörversuch unter Anleitung und mit professioneller Technik durchzuführen. Durch ihn erhielt ich auch die Chance die ICOSA 2015 in Graz, Österreich (International Conference of Spatial Audio) zu besuchen. Durch diese Veranstaltung konnte ich viel über das Thema 3D-Audio lernen.

An dieser Stelle möchte ich mich auch bei meiner Familie für die große Unterstützung und den Rückhalt während meines gesamten Studiums bedanken.

## Eidesstattliche Erklärung

Hiermit erkläre ich, meine Bachelorarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel zur Bearbeitung herangezogen zu haben. Ich versichere, dass ich bisher keine Prüfungsarbeit mit gleichem oder ähnlichem Thema bei einer Prüfungsbehörde oder einer anderen Hochschule vorgelegt habe.

Ort, Datum

Sascha Coronini

## Inhaltsverzeichnis

Kurzfassung.....	I
Vorwort.....	II
Eidesstattliche Erklärung.....	III
Inhaltsverzeichnis.....	IV
1. Einleitung.....	1
2. Von der Monofonie zur Stereophonie.....	1
3. Mehrkanal-Surround.....	3
4. 3D-Audio.....	6
4.1. Kanalbasierte Audioformate.....	8
4.2. Objektbasierte Audioformate.....	9
4.3. 3D-Audio Wiedergabeverfahren.....	10
4.3.1. Wellenfeldsynthese.....	10
4.3.1.1. WFS: Praktische Umsetzung.....	11
4.3.2. Iosono.....	13
4.3.2.1. Iosono: Produktion.....	15
4.3.2.2. Iosono: Anwendungen.....	16
4.3.3. Auro 3D.....	17
4.3.3.1. Auro 3D: AuroMax.....	18
4.3.4. Dolby Atmos.....	20
4.3.4.1. Dolby Atmos: Anwendung im Kino.....	20
4.3.4.2. Dolby Atmos: Anwendung im Heimkinobereich.....	21
4.3.4.3. Dolby Atmos: Mobile Anwendung.....	21
4.3.4.4. Dolby Atmos: Produktion.....	21
4.3.5. Ambisonics.....	22
4.3.6. DTS:X.....	23
4.3.7. Enveloping Surround.....	24
4.3.8. Sonic Emotion Absolute 3D.....	25
4.3.9. NHK 22.2.....	26
5. Binauraltechnik.....	27
5.1. Software mit Möglichkeit zur binauralen Simulation.....	29
5.1.1. New Audio Technology Spatial Audio Designer.....	30
5.1.2. StiffNeck.....	31

5.1.3. Logic Pro 9.....	32
5.1.4. Auratorium.....	33
6. Hörversuch.....	34
6.1. Bauhausbühne Dessau.....	35
6.2. Mikrofonierung.....	36
6.2.1. Mikrofonierung für 5.1 bzw. Stereo.....	36
6.2.2. 3D-Mikrofonierung.....	37
6.3. Postproduktion für den Hörversuch.....	39
6.3.1. Band.....	39
6.3.2. Atmo-Mikrofone.....	40
6.3.3. Effekte.....	40
6.4. Wiedergabetechnik im Tonstudio.....	40
6.4.1. Konfiguration 1: Abhören über die 3D-Audio-Lautsprecherkonfiguration.....	41
6.4.2. Konfiguration 2: Abhören über die 3D-Kopfhörersimulation.....	43
6.5. Panning.....	44
6.5.1. Lautsprecher: Iosono Spatial Audio Workstation.....	44
6.5.2. Binaural: New Audio Technology Spatial Audio Designer.....	45
6.6. Signalbearbeitung.....	46
6.7. Probanden.....	47
6.8. Versuchsablauf.....	47
6.9. Auswertung des Fragebogens.....	48
6.10. Diskussion der Ergebnisse.....	49
6.10.1. Detaillierte Ergebnisbetrachtungen.....	49
7. Fazit.....	51
8. Literaturverzeichnis.....	52
8.1. Bücher und andere Veröffentlichungen.....	52
8.2. Internetquellen.....	55
9. Anhang.....	i
9.1. Anhangsverzeichnis.....	i

## Abbildungsverzeichnis

Abbildung 1: Quadrofone Lautsprecheranordnungen.....	3
Abbildung 2: Lautsprecheranordnung für 5.1 nach ITU-R BS.775-1 .....	5
Abbildung 3: Empfehlung zur Lautsprecheranordnung bei 7.1.....	6
Abbildung 4: Workflow einer Kanalbasierten Audioproduktion.....	8
Abbildung 5: Workflow einer Objektbasierten Audioproduktion.....	9
Abbildung 6: Prinzip der WFS mit diskreten Lautsprechern.....	10
Abbildung 7: Das Huygens Frensel-Prinzip übertragen auf die Audio-Wiedergabe.....	10
Abbildung 8: Anwendungsszenarien für den Audioprozessor IPC 100.....	16
Abbildung 9: Auro 9.1.....	17
Abbildung 10: AuroMax 26.1.....	19
Abbildung 11: Lautsprecher-Anordnung für Dolby Atmos im Kino.....	20
Abbildung 12: Empfohlene 7.1.2 Setup mit „upward-firing-Lautsprechern“.....	21
Abbildung 13: Sonic Emotion WaveDesigner.....	25
Abbildung 14: Positionen der Ebenen bei NHK 22.2.....	26
Abbildung 15: Neumann KU100, Dummy Head.....	28
Abbildung 16: Spatial Audio Designer: Mix-Modul.....	30
Abbildung 17: Motorisierte Lautsprecherkonfiguration mit Messaufbau.....	31
Abbildung 18: Logic Pro 9 Binaural Panner.....	32
Abbildung 19: Auratorium Raumakustik Simulation.....	33
Abbildung 20: Konzertsaal im Bauhaus Dessau. Rot: Bühne, grün: Zuschauerbereich.....	35
Abbildung 21: Produktion und Reproduktion einer Raumakustik.....	37
Abbildung 22: Atmos 5.1 Controller & Brauner ASM 5 Mikrophon.....	37
Abbildung 23: 45° Anordnung des Höhenmikrofons.....	38
Abbildung 24: Atmo-Mikrofonierung. Das mittige 5.1 Array besteht aus zwei Ebenen.....	38
Abbildung 25: Spurenplan; Culcha Candela.....	39
Abbildung 26: Signalverlauf Iosono.....	41
Abbildung 27: 23.1 Lautsprecheranordnung.....	42
Abbildung 28: Signalverlauf Kopfhörer.....	43
Abbildung 29: Verteilung der Audio-Objekte in der Iosono Spatial Workstation.....	44
Abbildung 30: Gruppe 1 bis 9 äquivalent zur 9.1 Lautsprecheraufstellung.....	45
Abbildung 31: Fünf Testpersonen hören die binaurale Kopfhörersimulation.....	47
Abbildung 32: Durchschnittswerte der Testergebnisse.....	48

## Abkürzungsverzeichnis

AES	Audio Engineering Society
CAD	Computer-aided design
CEO	Chief executive officer
DAW	Digital Audio Workstation
DSP	Digital Signal Processing
EBU	European Broadcasting Union
HDTV	High Definition Television
IRT	Institut für Rundfunktechnik
ITU	International Telecommunications Union
LFE	Low Frequency Emitter (Tiefbasslautsprecher)
MADI	Multichannel Audio Digital Interface
PA	Public adress
PCM	Pulse-code modulation
RMU	Rendering & Mastering Unit
UHDTV	Ultra High Definition Television
VDT	Verband Deutscher Tonmeister e.V.

# 1. Einleitung

Aus der immer weiter fortschreitenden Entwicklung von Audio-Systemen, die von einer mono- über eine stereofone Wiedergabe bis hin zu einer den Hörer horizontal umgebenden Beschallung verlief, folgt die Verwirklichung einer dreidimensional einhüllenden Wiedergabetechnik als nächster logischer Schritt.

Einige Hersteller haben bis dato verschiedene Wiedergabeformate auf den Markt gebracht, welche mit teilweise unterschiedlichen Herangehensweisen versuchen, eine möglichst authentische Klangwiedergabe zu erreichen. Um diesen dreidimensionalen Höreindruck auch mobil bzw. ohne großen Aufwand zu realisieren, werden unter der Verwendung von Binauraltechnik verschiedene Werkzeuge angeboten, die es ermöglichen 3D-Audiosysteme zu simulieren und über Kopfhörer wiederzugeben.

Ein im Rahmen dieser Arbeit durchgeführter Hörversuch vergleicht eine 3D-Audiosimulation der Firma New Audio Technology mit einem mit 3D-Audiosystem der Firma Iosono, das 24 Lautsprecher verwendet.

## 2. Von der Monophonie zur Stereophonie

Erst 1877 wurde es mit der Erfindung des Phonographen durch Thomas Alva Edison möglich Schall- bzw. Tonsignale zu speichern.

Vom Ende des 19. bis Anfang des 20. Jahrhunderts wurden zur Schallspeicherung rein mechanisch-akustisch arbeitende Vorrichtungen verwendet, welche zunächst mit Wachswalzen und seit 1887 auch mit Hartgummiplatten funktionierten. Zehn Jahre später kam zur Schallspeicherung die Schellackplatte als zusätzliches Speichermedium dazu, welche zu diesem Zweck bis etwa 1955 genutzt wurde.

(vgl. zu diesem Abschnitt Dickreiter, 2014a, S. 422)

Noch bevor im Jahr 1936 die magnetische Tonspeicherung Einzug hielt, wurde die mechanische Aufzeichnung nach und nach verbessert. Eine entscheidende Verbesserung brachten die im Jahr 1948 eingeführte „Langspielplatte“ (LP) und die ein Jahr später eingeführte „Single“ mit sich.

Diesem vorangegangen war die Erfindung des Mikrofons und des Verstärkers ab etwa 1925. Mit der Tonsignalspeicherung auf magnetischem Tonband kam zu der mechanischen Speicherung ein neues Format hinzu. (vgl. zu diesem Abschnitt Dickreiter, 2014a, S. 422)

*„Mit diesem sog. Magnettonverfahren stand erstmals ein zunehmend qualitativ hochwertiges Verfahren zur Verfügung, das eine umfangreiche Nachbearbeitung der Aufnahmen auch für die anschließende mechanische Speicherung zuließ, damit wurden Tonproduktionen im Bereich der Musik und des Hörspiels mit umfangreichen Bearbeitungen des Tonmaterials überhaupt erst möglich.“*

*(Dickreiter, 2014a, S. 422)*

Schallaufzeichnungen waren sowohl in ihrer mechanischen, als auch in ihrer magnetischen Form bis etwa 1955 monofon, also einkanalig. (vgl. Dickreiter, 2014a, S. 422)

Das Problem an der lange genutzten monofonen Technik war, dass nur ein sehr eingeschränkter Raumeindruck hörbar gemacht werden konnte. So konnte die Raumgröße, welche über den Hallanteil und die Entfernung der Schallquelle zum Mikrofon vermittelt wird, nur grob abgeschätzt werden. Dies hatte mitunter zur Folge, dass sich der entstehende Raumeindruck durch Schallreflexionen im Abhörraum deutlich von der Raumakustik des Aufnahmeraums unterschied. Darunter litt unter anderem die Sprachverständlichkeit.

(vgl. zu diesem Abschnitt Schätzlein, 2005, S. 187)

So war der Übergang von der ein- zur zweikanaligen Schallaufzeichnung, also von der Monofonie zur Zweikanal-Stereofonie<sup>2</sup>, ein wichtiger Entwicklungsschritt. Dieser ereignete sich ab etwa 1955, während Übergang zur Mehrkanal-Stereofonie ab etwa 1995 stattfand.

(vgl. zu diesem Abschnitt Dickreiter, 2014a, S. 422)

Im Jahre 1954 wurden für den Heimgebrauch erste zweikanalige Aufnahmen der Firma RCA auf vorbespielten Tonbändern veröffentlicht. Vier Jahre später erschienen dann die ersten stereofonen Vinylplatten. Schließlich wurde ab 1961 in den USA und zwei Jahre später auch in

---

2 Das Wort Stereofonie, welches auch oft in seiner Abkürzung Stereo verwendet wird, leitet sich von den griechischen Worten *stereos* (fest oder starr) und *phone* (Ton, Stimme) ab und wurde in vielen Kontexten, die sich mit zwei Audio-Kanälen beschäftigen verwendet. (vgl. Ahrens, 2012, S. 4)

Deutschland ein Verfahren zur stereofonen Rundfunkübertragung (Pilotonverfahren) angewendet. Am 30. August 1963 wurde Deutschlands erstes Rundfunkkonzert vom Sender Freies Berlin in Stereo übertragen.

(vgl. zu diesem Abschnitt Slavik/Weinzierl, 2008, S. 610)

Experimente mit stereofonen Übertragungen fanden jedoch auch schon deutlich früher statt. Bereits 1881 erlangte der französische Ingenieur Clément Ader ein Patent für eine zweikanalige Übertragung von Audiosignalen. Dieses System, welches zu jener Zeit noch über nicht verstärkte Telefonleitungen funktionierte, wurde im gleichen Jahr verwendet um eine Aufführung aus der Pariser Oper stereofon an einen Ausstellungsraum auf der ersten *Internationalen Elektrotechnischen Ausstellung* in Paris zu übertragen.

Auch im Rundfunk wurden stereofone Übertragungen schon Mitte der 1920er Jahre, damals noch versuchsweise über zwei unabhängige Sendefrequenzen, durchgeführt. Im Jahre 1933 ermöglichte es der Ingenieur Alan Dower Blumlein eine zweikanalige Aufzeichnung in Form der Zweikomponentenschrift in einer Schallplattenrinne unterzubringen.

(vgl. zu diesem Abschnitt Slavik/Weinzierl, 2008, S. 610)

### 3. Mehrkanal-Surround

Aus den ersten Experimenten mit mehr als zwei Lautsprechern resultierte laut Slavik und Weinzierl die Quadrofonie, bei der anfangs noch unsynchronisierte, mono- oder stereofone Bänder und später vier-kanalige Magnetformate benutzt wurden. Das vierkanalige

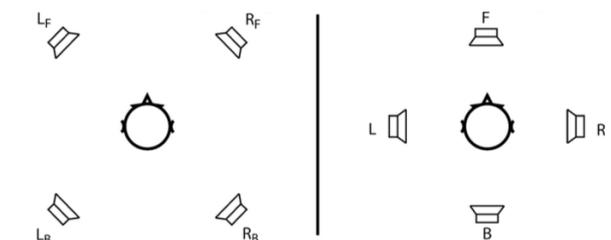


Abbildung 1: Quadrofone Lautsprecheranordnungen

Wiedergabeformat kann, wie in Abbildung 1

(Slavik/Weinzierl, 2008, S. 612)

veranschaulicht, in zwei verschiedenen Variationen aufgestellt werden; in Form des überwiegend verwendeten *Scheiber Array* (links im Bild) oder als sogenanntes *Dynaquad Array*. Der Filmtone und die elektroakustische Musik<sup>3</sup> waren die ersten Einsatzmöglichkeiten der Quadrofonie.

(vgl. zu diesem Abschnitt Slavik/Weinzierl, 2008, S. 612f)

3 Verwendet überwiegend Lautsprecher als Klangquellen. (vgl. Slavik/Weinzierl, 2008, S. 612)

Auch das Kino vollzog einen Wandel von der monofonen Wiedergabe hin zu einer stereofonen und später mehrkanaligen Wiedergabetechnik.

Anfangs wurde das Publikum noch von einem Lautsprecher beschallt, welcher sich hinter der Leinwand befand. Da jedoch zu diesem Zeitpunkt bereits zweikanalige Tonbandmaschinen und Plattenspieler für den Heimgebrauch existierten, rüsteten schließlich auch die Kinos mit einer stereofonen Wiedergabe nach.

Aufgrund der vielen Sitze des Kinosaals konnten jedoch nur wenige Hörer eine richtige stereofone Wiedergabe vernehmen. Für viele Hörer, vor allem auf den seitlichen Plätzen, war keine richtige Lokalisation von Phantomschallquellen möglich. Für diese Hörer wurde der Dialog nicht mittig wahrgenommen, sondern fälschlicherweise von rechts bzw. links.

(vgl. zu diesem Abschnitt Friesecke, 2014, S. 785)

Dieses Problem wurde später mit einem Dialog-Lautsprecher gelöst, welcher aufgrund seiner mittigen Positionierung auch Center-Lautsprecher genannt wurde. So entstand eine Mischung aus einer mono- und stereofonen Wiedergabe, bei der der Dialog, vorausgesetzt er wurde auf den Dialog-Lautsprecher gemischt, nur aus jenem gehört und somit eindeutig lokalisiert werden konnte.

Für die hinteren Reihen war eine Lokalisation jedoch nur schwer möglich. Aufgrund der durch die Entfernung immer kleiner werdenden Stereo-Breite entsprach das Geschehen immer mehr einer monofonen Wiedergabe. Dieser Effekt intensivierte sich, je weiter der Platz von der Leinwand bzw. den Lautsprechern entfernt war. Die Zuschauer der hinteren Reihen mussten bei dieser Form der Wiedergabe zusätzlich den Nachteil in Kauf nehmen, dass der Raumeindruck des Kinosaals aufgrund der Entfernung zur Beschallung den künstlichen Raum des Films überlagerte und der Zuschauer sich folglich nicht in den Film hineingezogen fühlen konnte.

Um diesen Effekt zu minimieren und einen besseren künstlicheren Raumeindruck zu generieren, wurden zusätzliche Lautsprecher im Raum installiert.

Diese Lautsprecher, welche in Anbetracht der Positionierung um den Zuschauer herum Surround-Lautsprecher genannt werden, wurden ab dem ersten Drittel des Kinosaals seitlich an den Wänden verteilt. Um zu verhindern, dass die Lautsprecher lokalisierbar waren und so das Risiko bestand, dem Zuschauer die Illusion des Films zu zerstören, wurden in der Regel viele kleinere Lautsprecher verwendet um ein sehr diffuses Schallfeld zu erzeugen. Außerdem war die Verwendung von von Dipolstrahlern<sup>4</sup> und der Einsatz von Tiefpassfiltern vorteilhaft, da so die

---

4 Lautsprecher mit der Richtcharakteristik Acht. Phasenrichtiger und Phasengedrehter Schall wird gleichzeitig abgestrahlt. (vgl. Friesecke, 2014, S. 787)

Lokalisation nochmals deutlich verschlechtert wurde. Die Surround-Lautsprecher waren zu Anfang noch einkanalig, wurden jedoch aufgrund des Einsatzes von Effekten, welche beispielsweise hinter dem Hörer stattfanden, zu stereofonen Surround-Lautsprechern. Auch die Wiedergabe von Rauminformationen wurde dadurch verbessert.

(vgl. zu diesem Abschnitt Friesecke, 2014, S. 785ff)

Das daraus entstandene Format 5.1 wird bis heute verwendet, wenn eine Wiedergabe von diskretem Mehrkanalton gewünscht ist. Darüber hinaus stellt es die Grundlage für beinahe alle mehrkanaligen Wiedergabesysteme dar. Weiterentwickelt wurde das Format von Tomlinson Holman (THX), Gerhard Steinke (VDT) und Günther Theile (IRT). Es trug ursprünglich den Namen 3/2-Format, was die drei Frontkanäle sowie zwei Surround-Kanäle definiert. Bei der späteren 5.1 Empfehlung wurde noch ein weiterer Kanal für tieffrequente Effekte, genannt LFE, hinzugefügt. Dieses Verfahren wurde bereits im Jahr 1994 von der *International Telecommunications Union* standardisiert und trägt den Namen ITU-R BS. 775-1.

(vgl. zu diesem Abschnitt Slavik/Weinzierl, 2008, S. 643)

Wie in Abbildung 2 zu sehen, sind die Lautsprecher nach ITU-R BS. 775-1 mit gleichem Abstand zum Hörer angeordnet und befinden sich dabei auf einer Kreislinie.

Hier sind die vorderen zwei Lautsprecher um  $\pm 30^\circ$  und die hinteren Surround-Lautsprecher um  $\pm 110^\circ$  zum Center-Lautsprecher versetzt.

Eine Abweichung von  $5\text{-}10^\circ$  ist hierbei

zulässig. (vgl. zu diesem Abschnitt Weinzierl, 2008, S. 643)

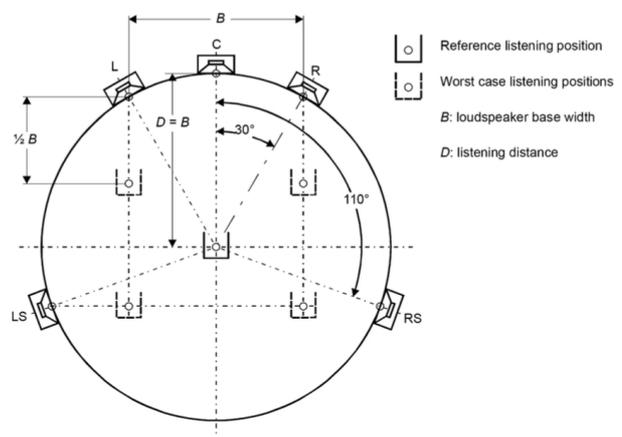


Abbildung 2: Lautsprecheranordnung für 5.1 nach ITU-R BS.775-1

(Slavik/Weinzierl, 2008, S. 643)

Des Weiteren gibt es die Formate 6.1 und 7.1. Dabei hat 6.1 einen zusätzlichen Lautsprecher für die rückseitige Wiedergabe, welcher den Namen *Back-Surround* (BS) trägt. Dieser ist genau zwischen den hinteren beiden Surround Lautsprechern angeordnet.

(vgl. zu diesem Abschnitt Weinzierl, 2008, S. 643)

Bei diesem Format können mit Hilfe von drei Surround-Kanälen, unabhängig vom Sitzplatz des Hörers, Bewegungen rund um den Rezipienten verwirklicht werden. (vgl. Friesecke, 2014, S. 788f)

Das Format 7.1 ist in Bezug auf das Format 5.1 um zwei zusätzliche Lautsprecher erweitert. Beim 7.1-Format wird der beim 6.1-Format verwendete zusätzliche dritte Surround-Kanal ergänzend auf zwei Kanäle aufgeteilt, um auch den hinteren Surround-Kanal in Stereo wiederzugeben. Diese werden BSL und BSR (Back-Surround links/rechts) genannt.

(vgl. zu diesem Abschnitt Friesecke, 2014, S. 788f)

Die empfohlene Aufstellung des 7.1 Formats ist in Abbildung 3 beschrieben.

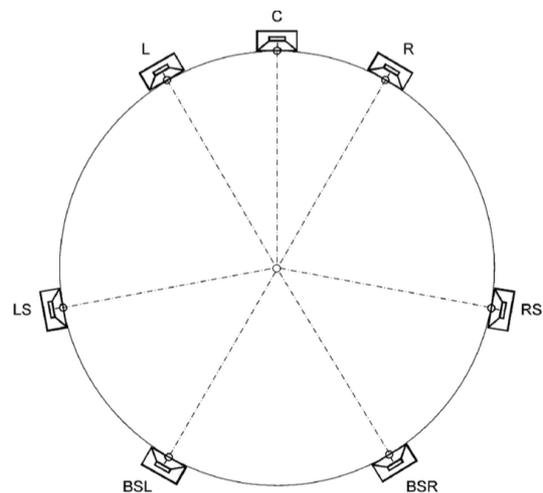


Abbildung 3: Empfehlung zur Lautsprecheranordnung bei 7.1

(Slavik/Weinzierl, Stefan, 2008, S. 643)

## 4. 3D-Audio

*„In einer alternativen Anwendung zusätzlicher Kanäle wird die Horizontalebene verlassen. Lautsprecheranordnungen in der Ebene oberhalb des Hörers vervollständigen den räumlichen Gestaltungsbereich und ermöglichen in bestimmten Grenzen den Aufbau eines dreidimensionalen Klangbilds“*

(Theile, 2014, S. 311f)

In der Videotechnik wird eine 3D-Wiedergabe durch den Eindruck von Tiefe erreicht. Im Gegensatz dazu wird in der Ton-Wiedergabe durch Einbeziehung der Raumdimension Höhe eine 3D-Wiedergabe erzeugt. (vgl. zu diesem Abschnitt Theile, 2014, S. 312)

Die Einführung einer neuen, vertikalen Achse führt laut Auro Technologies zu einer komplett neuartigen Hörerfahrung, welche im Gegensatz zu herkömmlichen zweidimensionalen Audio-Formaten deutliche Verbesserungen in Transparenz, Tiefe und Klangfarbe mit sich bringt. (vgl. Auro Technologies, 2015)

Die erste Musikaufnahme, welche die Höherebene mit einbezieht, wurde vor circa zehn Jahren

vom deutschen Klassik-Plattenlabel MDG (Musikproduktion Dabringhaus & Grimm) veröffentlicht.

Als Lautsprecherkonfiguration wurde das eigens entworfene 2+2+2-Format verwendet. Bei diesem Format handelte es sich um ein modifiziertes 5.1-Format, bei welchem statt dem Center- und Subwooferkanal zwei Lautsprecher oberhalb des rechten bzw. linken Front-Lautsprechers installiert wurden. Mit diesem Format wurde ein möglichst authentisches Klangbild angestrebt, das dem Klangbild des Konzertsaals entsprechen sollte. Als Datenträger wurde hierfür die Audio-DVD verwendet. (vgl. zu diesem Abschnitt Theile, 2014, S. 312)

Neben dem 2+2+2-Format gab es auch schon früher Audio-Formate, die es zuließen Klang in drei Dimensionen wiederzugeben. So sind seit 1990 die Verfahren Ambisonics und die Wellenfeldsynthese, zu welchen bereits 20 Jahre zuvor die Grundlagen formuliert wurden, Gegenstand ausgeprägter akustischer Forschung und Entwicklung.

Den Grundstein für alle zukünftigen Strategien stellt die immer zunehmende Rechnerleistung dar, die für die hohe Anzahl an Kanalsignalen benötigt wird. Denn diese Kanalsignale werden in der Regel nicht auf einem Medium gespeichert, sondern erfordern es meist in Echtzeit berechnet zu werden. (vgl. zu diesem Abschnitt Slavik/Weinzierl, 2008, S. 657)

Das Bedürfnis, die akustische Wirkung mit einer immer höher werdenden Anzahl von Wiedergabekanälen vollkommener zu machen, ist laut Stoll und Hartmann ein verständliches Anliegen. Dieses Anliegen gehe allerdings mit der Gefahr einher, den Basisstandard 5.1 aufzuweichen und somit Konsumenten und die produzierende Industrie zu verunsichern. Darüber hinaus wird die systemübergreifende Kompatibilität von mit hohem Aufwand produzierten Mehrkanal-Produktionen gefährdet. Ein möglicher Lösungsansatz für dieses Problem sei, zumindest in der zukünftigen Rundfunkübertragung, die Nutzung objektorientierter Audio-Übertragungsverfahren, welche es im Gegensatz zu den heute üblichen kanalbasierten Audiosystemen möglich machen die Audio-Information getrennt von den räumlichen Eigenschaften zu übermitteln. (vgl. zu diesem Abschnitt Stoll/Hartmann, 2014, S. 1209)

Das folgende Kapitel unterscheidet zwischen objekt- und kanalbasierten Systemen und gibt Einblick in bestehende Wiedergabesysteme, die auf unterschiedliche Weise 3D-Audio erzeugen können.

## 4.1. Kanalbasierte Audioformate

Kanalbasierte Audio-Formate verfolgen nicht das Ziel einer physikalischen Synthese von realen oder virtuellen Schallfeldern, sondern versuchen Phantomschallquellen basierend auf dem Phänomen der Summenlokalisierung zu produzieren. Im Fall einer dreidimensionalen Wiedergabe werden laut Slavik und Weinzierl Signalpegel für Lautsprecher-Tripel berechnet. Realisiert wird dies mit dem sogenannten *Vector Base Amplitude Panning* (VDAP), was für beliebige zwei-, bzw. dreidimensionale Lautsprecherkonfigurationen eingesetzt werden kann. VDAP bietet eine mathematische Formalisierung des stereofonen Panoramareglers und macht es möglich, Audiosignale zusätzlich in einer dritten Dimension, der Vertikalen, zu verteilen.

(vgl. zu diesem Abschnitt Slavik/Weinzierl, 2008, S. 657)

Das Audiomaterial wird mit den Positionsdaten für eine bestimmte Lautsprecherkonfiguration berechnet, woraus für jeden Lautsprecher eine diskrete Spur entsteht. Diese Spuren werden dann auf Speichermedien, wie z.B. Streaming-Servern oder anderen digitalen Medien gelagert und können bei Bedarf abgerufen werden.

(vgl. zu diesem Abschnitt Auro Technologies/Barco, 2015, S. 13)

Folgende Abbildung zeigt einen typischen Work-Flow einer kanalbasierenden Audioproduktion.

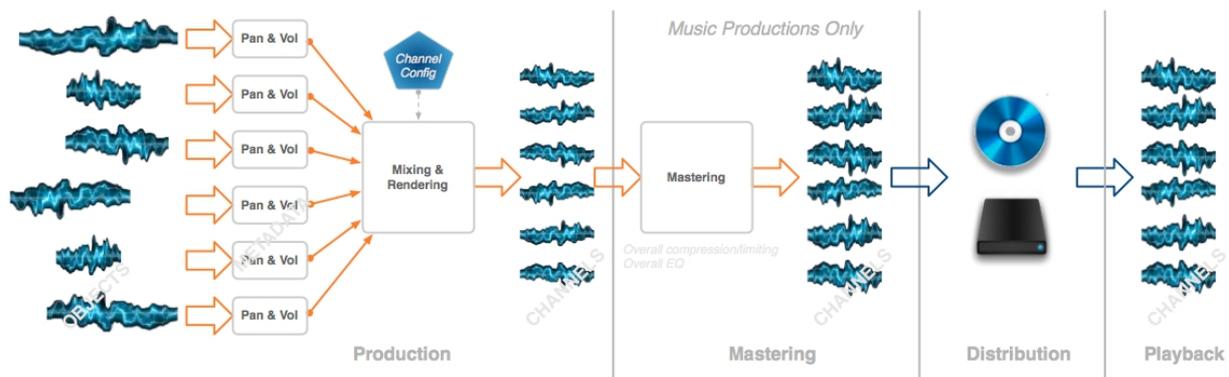


Abbildung 4: Workflow einer Kanalbasierten Audioproduktion

(Auro Technologies/Barco, 2015, S. 13)

## 4.2. Objektbasierte Audioformate

Bei objektbasierten Wiedergabeverfahren werden Audioelemente mit ihren dazugehörigen Metadaten nicht zu Spuren für einzelne Kanäle, sondern zu Audio-Objekten verarbeitet. Diese beinhalten die Audioinformation sowie Metadaten getrennt voneinander und werden in einem Container<sup>5</sup> gespeichert.

(vgl. zu diesem Abschnitt Auro Technologies/Barco, 2015, S. 14)

Als Containerformat kommt beispielsweise das *Advanced Authoring Format* (\*.aaf) in Frage, welches einen objektorientierten Austausch von synchronem Ton und Video zwischen unterschiedlichen Audio- und Video-Workstations realisiert. (vgl. Slavik, 2014, S. 944)

Die Metadaten beschreiben die momentanen Eigenschaften des Audiosignals und können Parameter wie Pegel, Frequenzgang, Koordinaten im Raum oder Phasenbezug zu anderen Kanälen sein. (vgl. Kompendium, 2016)

Bei der Wiedergabe ist ein Signalprozessor notwendig, der die Audio-Informationen und Metadaten ausliest und an die installierte Lautsprecherkonfiguration verteilt.

Ein objektorientiertes Verfahren hat den Vorteil, dass ein Mix theoretisch für jegliche Lautsprecher-Konfigurationen benutzt werden kann.

(vgl. zu diesem Abschnitt Auro Technologies/Barco, 2015, S. 14)

Die Positionierung der Objekte ist in der Praxis allerdings stark abhängig von der verwendeten Lautsprecherkonfiguration. (vgl. Friesecke, 2014, S. 785) Denn „[e]in Klangobjekt kann nur aus einer Richtung wieder gegeben werden, in der sich auch ein installierter Lautsprecher befindet“ (Friesecke, 2014, S. 785).

Um für jeden Lautsprecher ein individuelles Audiosignal zu generieren, benötigt der Prozessor zudem die akustischen Eigenschaften des Wiedergaberaumes, die durch Einmessen des Ortes erlangt werden können. (vgl.

Kompendium, 2016)

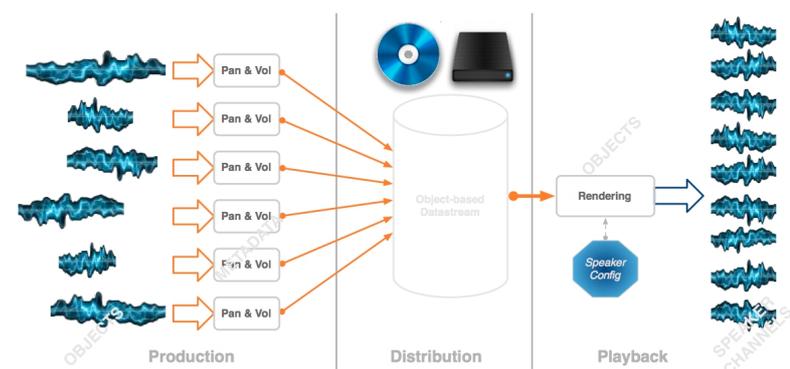


Abbildung 5: Workflow einer Objektbasierten Audioproduktion

(Auro Technologies/Barco, 2015, S. 14)

5 „[...] in der Computertechnik „logische Behälter“, in denen aus einem gemeinsamen Ursprung stammende und daher zusammengehörige Daten in einer definierten Struktur abgelegt und angeordnet sind.“ (Meyer-Schwarzenberger, 2014, S. 704)

## 4.3. 3D-Audio Wiedergabeverfahren

Für die Produktion und Wiedergabe von 3D-Audio stehen verschiedene Verfahren und Formate bereit. In diesem Kapitel wird eine Auswahl verschiedener Hersteller vorgestellt.

### 4.3.1. Wellenfeldsynthese

Die Wellenfeldsynthese (WFS) ist ein Wiedergabeverfahren, das es erlaubt durch Lautsprecherarrays virtuelle Schallfelder zu erzeugen (vgl. Wittek, 2004, S. 1), die in ihrer Schallausbreitung einer realen Schallquelle entsprechen. (vgl. Lehmann/Langhammer, 2013, S. 355). Dies wird erreicht indem die Schallwellen der einzelnen Lautsprecher überlagert werden und so, wie in Abbildung 6 zu sehen, ein synthetisiertes Schallfeld erzeugen. (vgl. Lehmann/Langhammer, 2013, S. 355) Dem zugrunde liegt das *Huygens-Fresnel-Prinzip*, welches beschreibt, dass jeder Punkt einer Wellenfront als Ausgangspunkt einer neuen, sich kugelförmig ausbreitenden Welle betrachtet werden kann. (vgl. Slavik/Weinzierl, 2008, S. 665)

Angewandt auf die Audiowiedergabe können Wellenfelder von einzelnen Schallquellen berechnet und mit Hilfe von Lautsprechern, wie in Abbildung 7 dargestellt, wiedergegeben werden. Diese Lautsprecher müssen dabei um den gesamten Wiedergaberaum verteilt sein. (vgl. zu diesem Abschnitt Lehmann/Langhammer, 2013, S. 355)

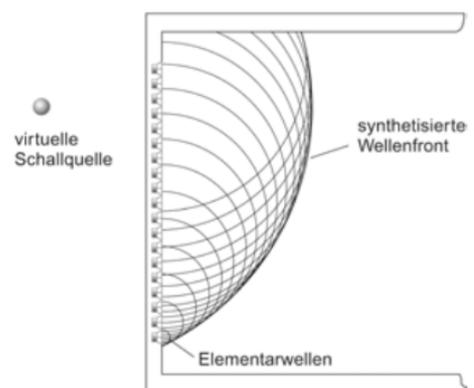


Abbildung 6: Prinzip der WFS mit diskreten Lautsprechern

(Schullan, Zuleeg, Hoeg, 2014, S. 583)

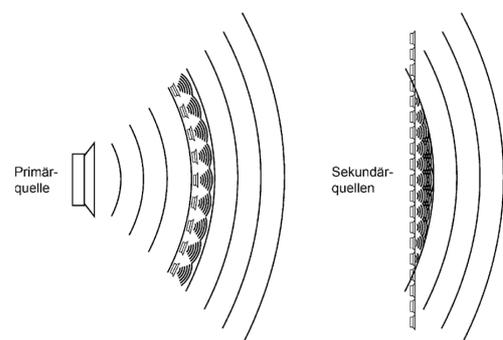


Abbildung 7: Das Huygens Fresnel-Prinzip übertragen auf die Audio-Wiedergabe

(Slavik/Weinzierl, 2008, S. 666)

Um im Gegensatz zu herkömmlichen Wiedergabesystemen keine Phantomschallquellen zu erzeugen, sondern eine physikalisch reale Resynthese eines räumlich ausgedehnten Schallfelds zu erreichen, ist eine hohe Anzahl von Lautsprechern nötig, welche dicht nebeneinander angeordnet sind. (vgl. Slavik/Weinzierl, 2008. S. 664f)

*„Dazu wird jeder Lautsprecher genau in dem Augenblick angesteuert, in dem die Wellenfront von der Originalquelle diesen Punkt passieren würde. Auf diese Art können identische Wellenfronten erzeugt werden, wie sie Schallquellen innerhalb dieser Lautsprecheranordnung abstrahlen.“*

*(Schulaaan, Zuleeg, Hoeg, 2014, S. 582f)*

Der Vorteil der sich daraus ergibt, liegt in der stabilen Lokalisation von virtuellen Schallquellen im Raum. Als *sweet spot*<sup>6</sup> kann beinahe der gesamte Zuschauerraum angesehen werden, da die einzelnen Lautsprecher nur noch als Teil der wiedergegeben Wellenfront erkannt werden. Somit wird die Beschallung nicht mehr als Entstehungsort des Schalls wahrgenommen.

(vgl. zu diesem Abschnitt Schullaaan, Zuleeg, Hoeg, 2014, S. 583)

#### 4.3.1.1. WFS: Praktische Umsetzung

Damit die Wellenfeldsynthese in der Realität ideal funktioniert, darf der Abstand der einzelnen Lautsprecher nicht größer sein als die Hälfte der Wellenlänge der höchsten abzustrahlenden Frequenz, was bei einer oberen Grenzfrequenz von 10 kHz maximal 1,7 cm beträgt und in praktischen Anwendungen nur schwer zu realisieren ist.

Damit jeder Lautsprecher sein individuelles Signal abstrahlen kann, ist es zudem notwendig, dass jeder Lautsprecher von einem eigenen Verstärkerkanal versorgt wird.

(vgl. zu diesem Abschnitt Schullaaan, Zuleeg, Hoeg, 2014, S. 582f)

Eine unmittelbare Umsetzung wäre laut Slavik und Weinzierl praktisch allerdings nicht möglich, denn eine Resynthese des Schalls mathematisch perfekt umzusetzen, „[...]würde eine unendlich große, auf einer geschlossenen Fläche um den Hörer unendlich dicht angeordnete Anzahl von

---

6 Punkt/Bereich im Raum an dem die besten Lokalisations- und Klangresultate zu hören sind. (vgl Barco Audio Technologies, 2015)

monopol<sup>7</sup>- und dipol-förmig abstrahlenden Lautsprechern erfordern“ (Slavik/Weinzierl, 2008, S. 667).

Ein weiterer Nachteil ergibt sich laut Schullaan, Zuleeg und Hoeg in der Anwendung in Räumen mit unzureichender akustischer Dämpfung. Störende Reflexionen beeinträchtigen die WFS, worauf sich der räumliche Eindruck verringert. Folglich ist die WFS nur für Räume mit ausreichender akustischer Dämpfung anwendbar.

(vgl. zu diesem Abschnitt Schullaan, Zuleeg, Hoeg, 2014, S. 583)

Das Konzept der nicht exakten Wellenfeldsynthese entstand Ende der 1980er Jahre, welches von Berkhout formuliert und schließlich von De Vries und Kollegen an der Technischen Universität in Delft weiterentwickelt wurde.

Diese Formulierung sieht neben einer Ebene von Lautsprecherarrays, bestehend aus näherungsweise monopol-förmigen Lautsprechern, einen diskreten Abstand der Lautsprecher vor. Dieser Abstand sollte nach praktischen Erfahrungen 10 cm bis maximal 30 cm betragen, damit die Schallsynthese ohne gravierende Beeinträchtigungen in der Lokalisation von Schallquellen funktioniert. Artefakte, welche aus einer nicht perfekten Umsetzung der WFS entstehen, können allerdings verbessert werden. So können auf Grundlage von Hörversuchen Strategien entwickelt werden, um beispielsweise die „fehlenden Schallquellen“ durch einen vorgegeben Abstand des Hörers zum Lautsprecherarray zu kompensieren.

(vgl. zu diesem Abschnitt Slavik/Weinzierl, 2008, S. 667f)

Da die real umsetzbare Wellenfeldsynthese trotz ihrer Abstriche zur Theorie mit einem enormen technischen Aufwand verbunden ist, ist die Anzahl der WFS-Installationen für größere Hörschichten noch überschaubar. Größere Installationen sind in den Ilmenauer *Lindenlichtspielen*, im Pariser *Cinémathèque française* sowie im Münchner Erlebniskino der *Bavaria Filmstadt* zu finden. Außerdem ist eine Art Hybrid mit konventioneller Frontbeschallung und einer Raumsimulation über ein Wellenfeldsynthese-System namens *Bregenz Open Acoustics* (BOA), welches für Opernaufführungen benutzt wird, auf der *Bregenzer Seebühne* zu finden. Weiter werden Systeme zur Wellenfeldsynthese in der Fachhochschule Düsseldorf sowie in der TU Berlin für Forschungszecke und öffentliche Veranstaltungen eingesetzt. Im Bereich der Psychoakustik wird die WFS zur Schallemissionsforschung und in der Fahrzeugtechnik für die Lautsprecherbestückung von Fahrzeugen eingesetzt. (vgl. zu diesem Abschnitt Slavik/Weinzierl, 2008, S. 670)

---

7 Z.B. konventionelle Lautsprecher mit geschlossenem Gehäuse. (vgl. Maier, 2008, S. 285)

### 4.3.2. Iosono

*Iosono* ist ein Verfahren der Firma Barco Audio Technologies, welches neben einem eigenen Prozessor eine eigene Software besitzt. Neben einer Installation im Opernhaus in Sydney sind weltweit über 80 Installationen von Iosono Systemen im Betrieb. Der Prozessor unterstützt objekt- sowie kanalbasierte Verfahren. Außerdem unterstützt der Prozessor bis zu 128 Ein- und Ausgänge, welche über MADI oder AES/EBU zugänglich sind.

(vgl. zu diesem Abschnitt Barco Audio Technologies, 2015, S. 1f)

*Iosono* ist ein 3D-Audioverfahren, welches auf dem Prinzip der Wellenfeldsynthese basiert und vorzugsweise im Event-Bereich und in Kinoinstallationen verwendet wird. In der Anzahl der Wiedergabekanäle kennt das System keine Obergrenze. Kinomischungen werden jedoch laut Slavik mit maximal 64 Quellkanälen gefertigt, welche auf einem *Cinema Master* codiert werden. Bei der Wiedergabe können diese Quellkanäle dann in Echtzeit für bis zu 128 Lautsprecherpositionen berechnet werden. (vgl. zu diesem Abschnitt Slavik, 2014, S. 972) Spezielle Kinoton-Mischungen im *Iosono* 3D-Audio-Format *icf*, die in ein *Digital Cinema Package* (DCP) eingebettet werden können, liegen bis dato für zwei Kinofilme vor: „Immortals – Krieg der Götter“ und „My Way“ (beide 2011).

(vgl. Lehmann/Langhammer, 2013, S. 356)

Durch die Verwendung des Prinzips der Wellenfeldsynthese ist es laut Barco Audio Technologies nicht mehr nötig sich als Hörer im *sweet spot* aufzuhalten.

(vgl. Barco Audio Technologies, 2015)

*„Aufgrund der verwendeten Wellenfeldsynthese ergeben sich in der Praxis große Sweet Spots und eine genaue, vom Sitzplatz nahezu unabhängige Ortung der Schallquellen.“*

*(Slavik, 2014, S. 972)*

Das Verfahren kann unter der Voraussetzung, dass die Filmtonegrie und der Wiedergabesaal mit entsprechenden Lautsprechern ausgerüstet sind, eine realistische 3D-Wiedergabe erreichen. So können auch Signale wie beispielsweise Regen, welche bei anderen Verfahren von oben oder

seitlich wiedergegeben werden, realitätstreu von unten wiedergegeben werden.  
(vgl. zu diesem Abschnitt Slavik, 2014, S. 971)

*„Klänge können mit IOSONO nicht nur um den Hörer herum platziert werden, sondern auch innerhalb und außerhalb der Hörfläche oder über den Köpfen der Hörer.“*

*(Lehmann/Langhammer, 2013, S. 353)*

Da der Prozessor objektbasiertes Arbeiten unterstützt, können Inhalte lautsprecherunabhängig produziert werden. Dies reduziert den Aufwand und die Kosten, viele verschiedene Lautsprechersysteme mit den entsprechenden Formaten zu versorgen.

(vgl. zu diesem Abschnitt Lehmann/Langhammer, 2013, S. 353f)

Laut Lehmann und Langhammer können somit unterschiedliche Lautsprecheranordnungen mit einem Format bespielt werden, ohne dass 3D-Effekte verloren gehen. Außerdem sei es möglich aus einer *Iosono*-Mischung eine Surround-Mischung zu generieren, ohne dafür eine neue Surround-Mischung anfertigen zu müssen.

(vgl. zu diesem Abschnitt Lehmann/Langhammer, 2013, S. 357)

Wird eine 5.1 oder 7.1 Mischung abgespielt, generiert der Prozessor einen *Upmix* auf das *Iosono*-System. Dieser *Upmix* führt zu einer Erweiterung des *sweet spots* und darüber hinaus zu einer stabilen Ortung der Schallquellen. (vgl. Slavik, 2014, S. 972)

Weitere Vorteile des Systems sind neue kreative Freiheiten, wie Räume akustisch zu vergrößern bzw. zu verkleinern oder Kinobetreibern die Möglichkeit zu bieten nicht mit jeder neuen Entwicklung eines Formats ein neues Wiedergabesystem installieren zu müssen.

(vgl. Lehmann/Langhammer, 2013, S. 353f)

*„[...] ein kleiner Raum kann beispielsweise plötzlich wie eine Kirche klingen.“*

*(Lehmann/Langhammer, 2013, S. 353)*

#### 4.3.2.1. Iosono: Produktion

Für die Produktion von Audio-Inhalten kommt die Software *Spatial Audio Workstation* (SAW) zum Einsatz, die ein objektorientiertes Pannen ermöglicht. (vgl. Slavik, 2014, S. 971)

Die *Spatial Audio Workstation* ist nur als Plug-in für die DAW *Nuendo* verfügbar. Für die Wiedergabe von 3D-Klängen wird der Prozessor *Spatial Audio Processor IPC 100* verwendet. Im Plug-in werden die einzelnen in *Nuendo* sichtbaren Spuren als Audio-Objekte dargestellt und können im Raum platziert werden. Ebenso können den Audio-Objekten verschiedene Eigenschaften wie z.B. Bewegungen zugewiesen werden. So ist es möglich, mit einem Joystick in Echtzeit Bewegungen, beispielsweise bei einer Live-Anwendung, zu verwirklichen. Diese Bewegungspfade können darüber hinaus als Automation abgespeichert werden.

(vgl. zu diesem Abschnitt Lehmann/Langhammer, 2013, S. 354)

Die Positionseigenschaften werden neben zusätzlichen Metadaten, welche zur „[...] Optimierung und Anpassung der Inhalte an alle Wiedergabe-Umgebungen“ dienen, mit den Audiodaten in das Iosono-Audioformat eingebettet“ (Lehmann/Langhammer, 2013, S. 354). Ebenso wird ein zukunftssicherer Master erstellt, der es dem Audioprozessor ermöglicht, den Inhalt an das verwendete Lautsprechersystem und den verwendeten Raum anzupassen.

(vgl. Lehmann/Langhammer, 2013, S. 354)

Die Audio-Objekte können darüber hinaus als drei Arten von Klangquellen definiert werden. Als *ebene Welle* (plane waves) definierte Audio-Objekte werden aus einer bestimmten Richtung wahrgenommen. Audio-Objekte, die als *Punktquellen* (point sources) definiert sind, werden dagegen vom Hörer an einer bestimmten Position registriert. Beide Arten von Quellen werden dabei außerhalb der Hörfläche positioniert.

Als dritte Art werden *fokussierte Quellen* definiert, welche sich im Gegensatz zu ebenen Wellen und Punktquellen innerhalb des Hörraumes befinden und dort auch bewegt werden können.

(vgl. zu diesem Abschnitt Lehmann/Langhammer, 2013, S. 355)

*„Ein Gewitter kann also als aus der Ferne nahendes Donnern wahrgenommen werden. Die einzelnen Regentropfen können aber auch so platziert werden, dass sie auf den Schultern der Zuhörer landen.“*

*(Lehmann/Langhammer, 2013, S. 355)*

Diese Effekte, welche bisher nur bei dicht angeordneten Lautsprecherarrays möglich waren, werden durch die Technologie von *Iosono* auch ohne einen solchen Aufwand ermöglicht. (vgl. Lehmann/Langhammer, 2013, S. 355f) Damit die Audioquellen jedoch nicht aus den einzelnen Lautsprechern wahrgenommen werden, sollte der Abstand der Lautsprecher zueinander und der Abstand zur Zuschauerfläche mindestens ein Verhältnis von 1:2 aufweisen. (vgl. Lehmann/Langhammer, 2013, S. 358)

#### 4.3.2.2. Iosono: Anwendungen

Der Prozessor *IPC 100* kann Audiosignale von Speichermedien wie Medienservern, DVD, Blu-Ray oder CD verarbeiten. Selbstverständlich können Audiosignale auch direkt aus der *Spatial Audio Workstation* an den Audioprozessor geleitet werden. (vgl. zu diesem Abschnitt Lehmann/Langhammer, 2013, S. 355) Dieser kann neben einer Touchscreen-Steuerung auch netzwerkseitig gesteuert werden. Letztere wird beispielsweise bei einer Anwendung im Kino genutzt, wenn mit einer zentralen Mediensteuerung mehrere Geräte, wie Projektoren, Player oder Audioprozessoren gesteuert werden. (vgl. Lehmann/Langhammer, 2013, S. 359)

Bei einer Anwendung im Kino ist der Prozessor mit einem *Digital Cinema Server* verbunden. Bei einer Live-Anwendung kann der Prozessor Audiosignale von Mischpulten oder Sequenzern zu erhalten, welche mit Programmen wie beispielsweise *Max/MSP* über das *Open Sound Control*-Protokoll gesteuert werden. Die Berechnung der Daten erfolgt im Audioprozessor basierend auf der eingestellten und somit verwendeten Lautsprecherkonfiguration. Die Bearbeitung und Erstellung der Ausgangssignale, welche über Wandler an die Lautsprecher verteilt werden, erfolgt in Echtzeit. (vgl. zu diesem Abschnitt Lehmann/Langhammer, 2013, S. 355ff)

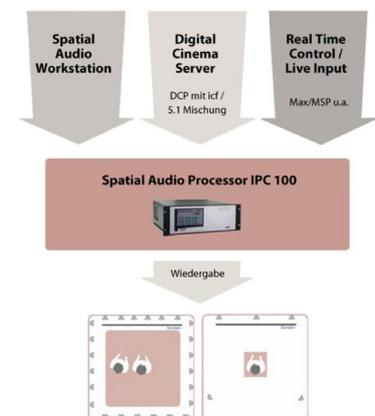


Abbildung 8:  
Anwendungsszenarien für den Audioprozessor IPC 100

„Ein größeres Kino mit einer IOSONO- Anlage verfügt im Durchschnitt über etwa 60 Lautsprecher.“

(Lehmann/Langhammer, 2013, S. 357)

### 4.3.3. Auro 3D

Ein weiteres 3D-Audiosystem bietet die Firma Auro Technologies mit *Auro 3D*, das im Jahr 2006 vorgestellt wurde. In der Basisversion von *Auro 3D* wird das 5.1-Format mit vier Höhenlautsprechern ergänzt, welche über den vier Ecklautsprechern des 5.1-Formats platziert werden. (vgl. zu diesem Abschnitt Theile, 2014, S. 312)

Entwickler dieses Systems war Wilfried Van Baelen; Gründer und CEO der Galaxy Studios. Dieser legte bereits 2004 den Grundstein zu *Auro 3D*, als er ein dreidimensionales Wiedergabesystem als Vorläufer entwarf, welches Teil seines *Auro 3D* Konzepts war.

(vgl. zu diesem Abschnitt Van Daele/Van Baelen, 2012, S. 1)

*Auro 3D* verfügt neben sechs eigenen Lautsprecherkonfigurationen über ein eigenes Codec-Verfahren. Die Lautsprecherkonfigurationen reichen von *Auro 8.0*<sup>8</sup> bis hin zu *AuroMax 26.1*, das neben einem Kanal für den LFE weitere Lautsprecher einbezieht. (vgl. zu diesem Abschnitt Auro Technologies, 2015, S. 9)

*AuroMax* ist dabei das einzige Format, welches in Kooperation mit der Firma Barco Audio Technologies objektbasiertes Audio verwendet. (vgl. Auro Technologies/Barco, 2015, S. 8)

Laut Van Daele und Van Baelen ermöglicht *Auro 3D* dem Rezipienten, über eine zusätzliche obere

Lautsprecherebene eine dritte Sound-Dimension zu

erfahren. Aufgrund der Benutzung der drei orthogonal angeordneten Achsen x, y und z wird hier echter dreidimensionaler Sound erzeugt, welcher zudem zu 100 Prozent mit den Formaten 5.1 und Stereo kompatibel ist.

(vgl. zu diesem Abschnitt Van Daele/Van Baelen, 2012, S. 1)

Als wesentliches Merkmal besitzt *Auro 3D* in der Variante 9.1 die dem Kubus ähnliche Anordnung von acht Lautsprechern, womit frühe Reflexionen über den gesamten oberen Halbraum wiedergegeben werden können. Somit lässt sich laut Theile eine subjektive räumliche Diffusität des Nachhalls in ausreichendem Maße reproduzieren. Praktisch nicht realisierbar seien allerdings Phantomschallquellen, welche zwischen den unteren und oberen Lautsprechern oder über dem Hörer lokalisiert werden sollen. (vgl. zu diesem Abschnitt Theile, 2014, S. 312)

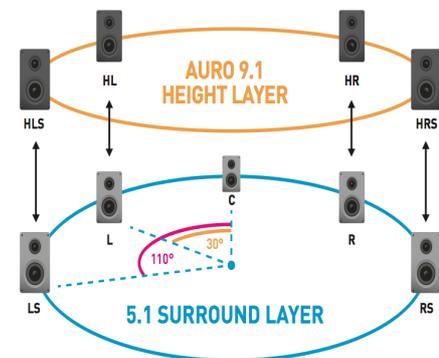


Abbildung 9: Auro 9.1

(Auro Technologies, 2015, S. 12)

<sup>8</sup> L, R, Ls, Rs, HL, HR, HLS, HRS (vgl. Auro Technologies, 2015, S. 9)

Bei der Aufnahme von Audio-Signalen erziele man laut Auro Technologies die besten Ergebnisse, wenn direkt im *Auro 3D*-Format mit einem geeigneten Auro Mikrofonarray aufgenommen werde. Mit kohärenten Signalen auf den Mikrofonkanälen gelinge somit eine gute Umhüllung des Hörers.

Für das Mikrofonarray ist als Grundlage ein 5.1 Aufnahmearray vorgesehen. Über diesem befinden sich vier nach oben gerichtete Mikrofone, welche als Quelle für die obere Lautsprecherebene dienen. (vgl. zu diesem Abschnitt Van Daele/Van Baelen, 2012, S. 1) Wichtig ist hierbei laut Theile eine möglichst genaue akustische Kanaltrennung der Mikrofone. Nur so könne eine räumliche Auflösung mit Direktschall, Atmo, Diffusschall und Reflexionen erfolgen. Diese Kanaltrennung sei bei einem Mikrofonarray mit mehreren Mikrofonen durchaus kompliziert. Soll Crosstalk, also unerwünschtes Übersprechen, vermieden werden, seien Methoden wie *OCT Surround* oder das Wählen von größeren Abständen zwischen den Mikrofonen notwendig. (vgl. zu diesem Abschnitt Theile, 2014, S. 316)

Beim *OCT-Surround*-Verfahren werden zwei um 90° zu den Seiten ausgerichtete Mikrofone mit Supernierencharakteristik und ein nach vorne ausgerichtetes Mikrofon mit Nierencharakteristik verwendet. (vgl. Theile, 2001, S. 13)

Der hauseigene Codec erlaubt es, mehrkanaliges Audio in einem PCM-Stream mit weniger Kanälen zu verarbeiten, bei dem beispielsweise 9.1 in einem 5.1 oder 2.0 Stream untergebracht werden kann. Um die originale Audioqualität zu erhalten, benötigt es bei der Wiedergabe allerdings einen Decoder, welcher die Kanäle wieder zu 100 Prozent separiert. Weiter bietet der Codec eine virtuelle, verlustlose Komprimierung und eine hohe Kompatibilität zu bereits existierenden Wiedergabe- und Distributionsverfahren.

(vgl. zu diesem Abschnitt Auro Technologies, 2011, S. 2)

#### 4.3.3.1. Auro 3D: AuroMax

Neben den kanalbasierten Formaten bieten die Firmen Auro Technologies und Barco zusammen ein objektbasiertes Format an. Dabei stellt Auro Technologies die Lautsprecherkonfiguration, während sich Barco um die Berechnung der objektbasierten Inhalte kümmert.

(vgl. zu diesem Abschnitt Auro Technologies/Barco, 2015, S. 5)

Als Audio-Prozessor kommt der *AuroMax Audio-Processor* zum Einsatz, welcher für die

Decodierung des Bitstreams und für die Berechnung der Audio-Objekte und *Beds*<sup>9</sup> zuständig ist. (vgl. Auro Technologies/Barco, 2015, S. 29)

Wie alle Auro 3D Lautsprecherkonfigurationen basieren die *AuroMax*-Formate ebenfalls auf dem 5.1 Surround-Format. (vgl. Auro Technologies/Barco, 2015, S. 11)

Dabei werden neue „Zonen“ geschaffen, welche die traditionellen Surround-Lautsprecher in kleinere Gruppen von Lautsprechern unterteilen und dem objektbasierten Ansatz eine größere Auflösung geben sollen. (vgl. Auro Technologies/Barco, 2015, S. 19)

Insgesamt existieren drei *AuroMax* Systeme. Die kleinste Konfiguration nennt sich *AuroMax 20.1* und basiert auf dem Standard *Auro 11.1*. Hier werden die Surround Lautsprecher sowie die Lautsprecher der oberen Ebene in zwei Zonen für jede Wand geteilt.

Für die nächstgrößere Lautsprecherkonfiguration *AuroMax 22.2* wird die obere Ebene zusätzlich in vier Zonen aufgeteilt.

Auro Technologies empfiehlt für die meisten Räume allerdings das größte Format *AuroMax 26.1* (siehe Abbildung 10), welches die höchste Kompatibilität mit Standard Surround-Formaten und *Auro 11.1* erreiche. (vgl. zu diesem Abschnitt Auro Technologies/Barco, 2015, S. 21-24)

Für die Erstellung von Audio-Inhalten steht die *Auro-3D Creative Tool Suite* zur Verfügung, welche die *Auro-3D Authoring*

*Tools* und *AuroMatic Pro*

*2D/3D* beinhaltet. Mit dem *Auro*

*Panner*, der Teil der *Auro-3D*

*Authoring Tools* ist und als Plug-in in Pro Tools verwendet

wird, können Audio-Objekte

und *Beds* in der Mischung

verteilt werden. Audio- und

Metadaten werden dann mit dem

*AuroMax Audio-Processor*

verarbeitet.

(vgl. zu diesem Abschnitt Auro

Technologies/Barco, 2015, S.

31f)

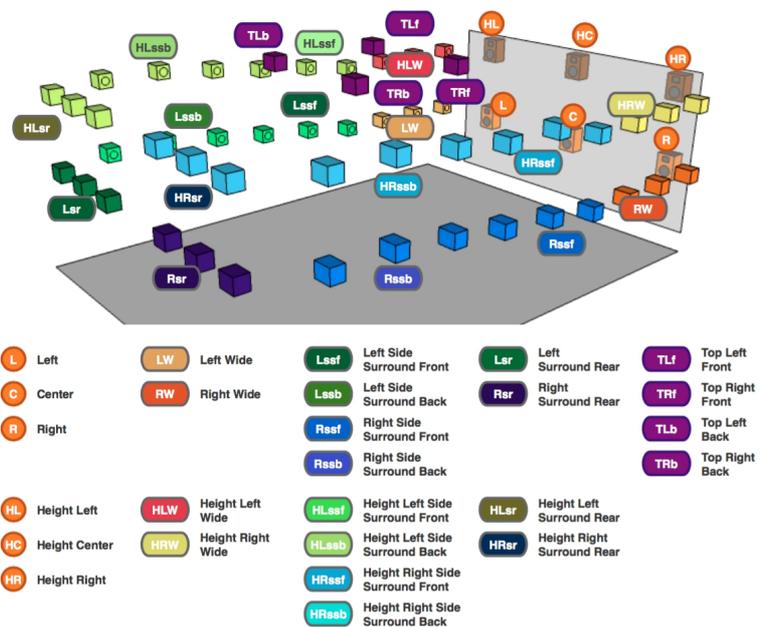


Abbildung 10: *AuroMax 26.1*

(Auro Technologies/Barco, 2015, S. 22)

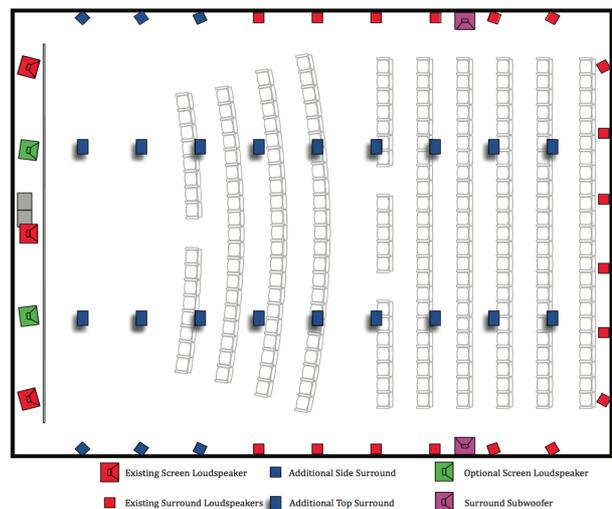
9 Statische Elemente, die hauptsächlich Musik und Atmosphären enthalten. (vgl. Slavik, 2014, S. 972)

#### 4.3.4. Dolby Atmos

Mit *Dolby Atmos* verfügt auch die Firma Dolby Laboratories über ein 3D-Audiosystem, das sowohl für die Anwendung im Kino sowie im Heimkinobereich, als auch für unterwegs mit Smartphone oder Tablet gedacht ist. (vgl. Dolby Laboratories, 2015a)

##### 4.3.4.1. Dolby Atmos: Anwendung im Kino

Für die Wiedergabe im Kino stehen bei *Dolby Atmos* bis zu 64 Wiedergabekanäle und bis zu 128 Quellkanäle zur Verfügung. Als Container dient Dolby dabei das *Digital Cinema Package (DCP)*. Dolby Atmos ist abwärtskompatibel zu 5.1 und 7.1 und unterstützt Upmixing sowie Downmixing. (vgl. Slavik, 2014, S. 972)



Für die Berechnung der großen Anzahl einzelner Lautsprecher, welche sich sowohl vorne, hinten, seitlich als auch in der Decke befinden, sorgt als RMU der *Dolby Atmos cinema processor CP850*. Dieser berechnet

Abbildung 11: Lautsprecher-Anordnung für Dolby Atmos im Kino

(Dolby Laboratories, 2014, S. 13)

und verteilt in Echtzeit die einzelnen Audio-Objekte und *Beds* auf die Lautsprecher. (vgl. zu diesem Abschnitt Dolby Laboratories, 2014, S. 7)

Dolby gibt zwar vor, wie die Lautsprecher im Saal zu verteilen sind (Siehe Abbildung 11), dennoch können, abhängig von Größe und Form eines Raumes, abweichende Lautsprecherpositionen gewählt werden. (vgl. Dolby Laboratories, 2014, S. 13)

#### 4.3.4.2. Dolby Atmos: Anwendung im Heimkinobereich

Auch für das Heimkino bietet *Dolby Atmos* eine Lösung, dreidimensionalen Ton wiederzugeben. Liegt ein *Dolby Atmos* Audioinhalt beispielsweise auf einer Blu-ray-Disc vor, kann dieser auf einem herkömmlichen Blu-ray-Player, welcher an einen *Dolby Atmos*-fähigen AV-Receiver angeschlossen ist, abgespielt werden. *Dolby Atmos*-fähige AV-Receiver unterstützen bis zu 24 am Boden und bis zu zehn an der Decke befindliche Lautsprecher.



Abbildung 12: Empfohlene 7.1.2 Setup mit „upward-firing-Lautsprechern“

(Dolby Laboratories, 2015b)

Für die Höhenwiedergabe stehen zwei

unterschiedliche Optionen zur Verfügung. Entweder können Lautsprecher in der Decke installiert werden oder es werden „upward-firing-Lautsprecher“ verwendet, die auf den äußeren Lautsprechern platziert werden um Reflexionen von der Decke zu erzeugen (Siehe Abbildung 12). Laut Dolby ist die Anzahl der Lautsprecher ausschlaggebend für die präzise Abbildung von Audioobjekten. (vgl. zu diesem Abschnitt Dolby Laboratories, 2015b)

#### 4.3.4.3. Dolby Atmos: Mobile Anwendung

Um auch unterwegs 3D-Audio Inhalte zu hören, ist ein Smartphone mit *Dolby Atmos* Unterstützung nötig. Damit können *Dolby Atmos* Inhalte unter Verwendung von Kopfhörern abgespielt werden. (vgl. zu diesem Abschnitt Dolby Laboratories, 2015c)

#### 4.3.4.4. Dolby Atmos: Produktion

Für die Erstellung von Audio-Inhalt stehen Toningenieuren mehrere Mixing-Tools zur Verfügung. Die Erstellung einer *Dolby Atmos*-Mischung ist in der *Digital Film Console* (DFC) der Firma Neve nativ möglich. Ebenfalls ist eine Integration in Pro Tools mittels des *Dolby Atmos Panner* Plug-ins realisierbar. Hierbei werden Metadaten vom Plug-in an die Dolby Atmos RMU, die *Dolby Atmos Rendering and Mastering Unit*, gesendet, welche es erlaubt Objekte in

einem Mix anzuordnen. Auch mit einer *MPC*- oder *Trion*-Konsole der Firma Harrison, die unter dem Betriebssystem *IKIS* laufen, ist eine direkte Verbindung zur *Dolby Atmos* RMU möglich. (vgl. zu diesem Abschnitt Dolby Laboratories, 2015d)

Diese Tools unterstützen bis zu 128 Spuren, welche für *Beds* oder Audio-Objekte verwendet werden können. (vgl. Dolby Laboratories, 2014, S. 7).

Die sogenannten *Beds* erschaffen als statische Elemente die Grundlage der Audio-Mischung und enthalten hauptsächlich Musik und Atmosphären, die mit herkömmlichen kanalbasierten Verfahren wie 5.1 oder 7.1 erstellt wurden. (vgl. Slavik, 2014, S. 972)

Laut Dolby Laboratories können *Beds* auch in Formaten wie 9.1, welche die Höhenlautsprecher mit einbeziehen, kreiert werden. Die *Bed*-Kanäle werden dabei direkt auf die vorgesehenen Lautsprecher verteilt, während die Audio-Objekte individuell im Raum verteilt werden können. Diese werden bei der Wiedergabe in Echtzeit, abhängig von der Lage der Lautsprecher, vom *Dolby Atmos Cinema Processor CP850* berechnet. Dafür werden für die einzelnen Audio-Objekte in ihren Parametern wie beispielsweise Delay oder Equalizer angepasst. (vgl. zu diesem Abschnitt Dolby Laboratories, 2014, S. 6f)

#### 4.3.5. Ambisonics

Eine weiteres 3D-Audioformat ist das von Michael Gerzon entwickelte Verfahren *Ambisonics*. Gerzon verfasste Anfang der 1970er Jahre die theoretischen Grundlagen am Mathematischen Institut in Oxford. Dieses Verfahren bietet die Möglichkeit, dreidimensionale Schallfelder mit theoretisch beliebiger Genauigkeit zu übertragen sowie virtuelle Klangquellen zu synthetisieren. Die Genauigkeit ist dabei von der Anzahl der verwendeten Übertragungskanäle bei der Aufnahme und der Anzahl der verwendeten Lautsprecher bei der Wiedergabe abhängig. Obwohl dieses Verfahren als Alternative zu Quadrofonie und Surround-Verfahren diskutiert wurde, konnte es sich auf dem Tonträgermarkt nicht durchsetzen.

Studien zu *Ambisonics* höherer Ordnung begannen 1999, wodurch *Ambisonics* zunehmend als Alternative zur Wellenfeldsynthese gesehen wurde. Durch *Ambisonics* höherer Ordnung wurde eine Synthese realer oder virtueller Schallfelder mit einer hohen Zahl von Lautsprechern möglich. (vgl. zu diesem Abschnitt Slavik/Weinzierl, 2008, S. 659f)

Ziel dieses Systems war es nach Richard Elen, ein Surround-System zu entwickeln, das vor allem musikalische Darbietungen in herkömmlichen Wohnzimmern wiedergeben kann. Dabei

sollten der originale Sound sowie die originalen akustischen Eigenschaften des Konzertumfelds wiedergegeben werden können.

Um Audio-Inhalte für die Wiedergabe zu erstellen, war es von Anfang an vorgesehen, diese mit speziellen Mikrofonen aufzunehmen. Zum Einsatz kommt hierfür beispielsweise das *Soundfield* Mikrofon von Soundfield Research, welches auch in Verbindung mit speziellen Panoramareglern benutzt wird. (vgl. zu diesem Abschnitt Elen, 2001, S. 2)

Dieses Mikrofon besteht aus vier Kapseln in Koinzidenz-Anordnung. Diese sind in der Form eines Tetraeders ausgerichtet und erzeugen vier Signale *W*, *X*, *Y* und *Z*.

*W* enthält dabei die gesamte Druckinformation eines Mikrofons mit Kugelcharakteristik. *X*, *Y* und *Z* stellen daneben die Audiosignale der drei Raumachsen dar. Der Vorteil dieser Anordnung ist neben der Raumabbildung die hohe Stereo- und Monokompatibilität.

(vgl. zu diesem Abschnitt Slavik, 2014, S. 934)

Aus der Aufnahme entsteht schließlich ein Signal aus vier Kanälen (Auch B-Format genannt), für dessen Wiedergabe mindestens vier Lautsprecherkanäle notwendig sind.

Um eine Lokalisation von Audioquellen zu verwirklichen und den *sweet spot* zu vergrößern, arbeitet *Ambisonics* mit der Rekonstruktion von Wellenfronten, welche mit Phasen- und Lautstärkeunterschieden überlagert werden.

Ein großer Vorteil dieses Systems ist die Tatsache, dass die Lautsprecher in nahezu jeglicher Aufstellung betrieben werden können. Die verwendete Aufstellung ist jedoch einem Decoder mitzuteilen, der dann die Signale der einzelnen Lautsprecher berechnet. Weiter ist *Ambisonics* kompatibel mit einer Stereo- oder 5.1-Wiedergabe, womit jedoch keine Höheninformation mehr wiedergegeben werden können. (vgl. zu diesem Abschnitt Elen, 2001, S. 2)

#### 4.3.6. DTS:X

Die Firma Digital Theater Systems (DTS) das 3D-Format *DTS:X* an.

*DTS:X* ist ein objektbasiertes Audio-Format, welches unabhängig von der Lautsprecherkonfiguration arbeitet.

Somit können verschiedene Konfigurationen bedient werden; eingebaute TV-Lautsprecher sowie Surroundsysteme bis hin zu Kinosystemen mit einer Vielzahl von Lautsprechern. Um *DTS:X* zu nutzen, werden spezielle AV-Receiver benötigt, die diesen Standard unterstützen. (vgl. zu

diesem Abschnitt Denon, 2016). Die Firmen Onkyo, Integra, Denon, Pioneer und Sony bieten dafür eine Reihe von AV-Receiver an. (vgl. DTS, 2016a)

Einhergehend stellt DTS das Format *DTS Headphone:X* zur Verfügung. Hier werden 3D-Audio Mischungen von Stereo bis zu 11.1 räumlich simuliert und für die Wiedergabe über ein mobiles Gerät bzw. Kopfhörer optimiert. Für die optimale Nutzung stehen eine Reihe von auswählbaren Kopfhörermodellen zur Verfügung. (vgl. zu diesem Abschnitt DTS, 2016b)

Weiter bietet DTS die Möglichkeit, 3D-Inhalte aus schon existierenden Audio-Inhalten zu generieren. Dies geschieht mit dem Algorithmus *DTS Neo:X*. Dabei können Quellen von 2.0 bis 7.1 zu 3D-Formaten<sup>10</sup> umgewandelt werden. (vgl. zu diesem Abschnitt DTS, 2013, S. 1)

#### 4.3.7. Enveloping Surround

Das Konzept zum 3D-Audiosystem *Enveloping Surround* liegt den Prinzipien der Firma *Saalakustik.de* zugrunde, die mit vier bis sechs Lautsprechern elektroakustisch eine Verbesserung der Akustik von Konzertsälen verwirklicht.

(vgl. von Schultendorff/Höpfner/Batke, 2015, S. 1)

Das 3D-Format hat den Anspruch den Direktschall von bisher existierenden Surround-Systemen zu nutzen und somit hochgradig kompatibel zu bestehenden Tonaufnahmen zu sein.

Zusätzlich wird ein reproduzierter oder echter Diffus-Schall mit einem umhüllenden Effekt auf weiteren Lautsprechern abgespielt. Dies wird verwirklicht, indem beispielsweise ein bereits vorhandenes 5.1 Surround-System mit zusätzlichen Lautsprechern ergänzt wird, welche das Ziel haben eine hohe Diffusität für den Umhüllungsklang zu erreichen. Diese Lautsprecher werden so positioniert, dass der Diffus-Klang aus Arealen kommt, welche vorteilhaft für eine Umhüllung des Hörers sind. Um dies zu erreichen wird ein Audioinhalt, der nur Diffus-Schall enthält, unter Verwendung von Raumimpulsantworten erstellt. Dieser kann darüber hinaus auch durch einen Upmix eines Mono- oder Stereosignals generiert werden. Dadurch wird erreicht, dass ein Hörer sich überall im durch das Surround-System abgegrenzten Raum aufhalten kann, ohne Einbußen in der Wiedergabequalität wahrzunehmen. Weiter wird der Einsatz eines Subwoofers überflüssig, da tiefe Frequenzen auf natürliche Weise wiedergegeben werden.

(vgl. zu diesem Abschnitt von Schultendorff/Höpfner/Batke, 2015, S. 4f)

---

<sup>10</sup> 9.1 oder 11.1 (vgl. DTS, 2013, S. 1)

### 4.3.8. Sonic Emotion Absolute 3D

Für die Unterhaltungsindustrie sowie auch für professionelle Anforderungen bietet die in der Schweiz gegründete Firma Sonic Emotion Lösungen für 3D-Audio an. Im Gegensatz zu anderen Herstellern setzt Sonic Emotion mit *Absolute 3D* in der Consumer-Sektion nicht auf eine Vielzahl von Lautsprechern, sondern nur auf ein einzelnes Lautsprechergehäuse, welches für 3D-Audio mit sechs Lautsprechern und einem optionalen, externen Subwoofer ausgestattet ist. Für den Vertrieb von *Absolute 3D* kooperiert Sonic Emotion mit einer Reihe von Herstellern, wie beispielsweise Teufel oder Samsung, welche diese Technik nutzen.

Ebenfalls steht eine App namens *Headquake Pro* für das iPhone zur Verfügung, welches eine Konvertierung von Stereo in 3D-Audio verspricht. (vgl. zu diesem Abschnitt Sonic Emotion, 2016)

Im professionellen Bereich bietet Sonic Emotion

den *Wave I 3D Sound-Processor*. Dieser arbeitet mit einem Algorithmus namens *Sound Field Control*, welcher nach dem Prinzip der Wellenfeldsynthese funktioniert. Dieser Algorithmus reduziert nach Sonic Emotion allerdings die für eine Wellenfeldsynthese benötigte Lautsprecheranzahl um den Faktor 10.

(vgl. zu diesem Abschnitt Sonic Emotion, 2013, S. 2)

Gesteuert wird der Prozessor über die Software *WaveDesigner* (Abbildung 13). Hier werden die Maße des Raums und die Positionen der verwendeten Lautsprecher eingetragen.

Für die Platzierung der Audioobjekte ist das Programm *WavePerformer* zuständig, welches als Plug-in in Form von z.B. AU, MAX MSP oder VST verfügbar ist.

(vgl. zu diesem Abschnitt Sonic Emotion, 2013, S. 6f)

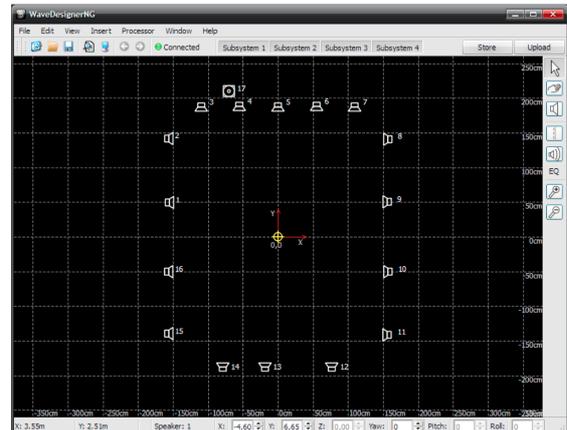


Abbildung 13: Sonic Emotion WaveDesigner

(Sonic Emotion, 2013, S. 6)

#### 4.3.9. NHK 22.2

Auf der Expo 2015 präsentierte die japanische Rundfunkgesellschaft NHK ein offen standardisiertes System mit 22 Haupt- und zwei Tieftonstrahlern.

(vgl. Kompendium, 2016a)

Entwickelt wurde dieses Format, welches momentan das Maximum für eine diskrete Mehrkanalstereofonie darstellt, für die sogenannte *Ultra High Definition (UHD)* Fernsehübertragung.

Mit diesem System wird eine Beschallung eines horizontalen Betrachtungswinkels von  $100^\circ$  angestrebt, bei dem beim Zuschauer das Gefühl entstehen soll, mitten im Geschehen zu sein. Im Gegensatz zu einem normalen HDTV-Bild ist das UHDTV-Bild in Höhe und Breite um den Faktor vier vergrößert. (vgl. zu diesem Abschnitt Stoll/Hartmann, S. 1208)

Da es sich um eine diskrete Mehrkanalstereofonie handelt, sind im Gegensatz zu beispielsweise *Dolby Atmos* keine Hardwareprozessoren notwendig.

Das System funktioniert in drei Ebenen. In der obersten Ebene sind neun Lautsprecher ringförmig oberhalb des Hörers angeordnet, wobei ein Lautsprecher von der Decke abstrahlt. Auch in der Mitte strahlen zehn Lautsprecher ringförmig auf Höhe der Ohren um den Hörer ab, während in der untersten Ebene drei Lautsprecher den Hörer von vorne unten beschallen. Die zwei Tieftöner befinden sich ebenfalls in der untersten Ebene.

Da die primären Audiokanäle von der mittleren Ebene wiedergegeben werden, ist das System kompatibel mit 5.1-, 6.1- bzw. 7.1-Formaten, da diese ebenfalls ihre Hauptkanäle in der mittleren Ebene haben. (vgl. zu diesem Abschnitt Kompendium, 2016a)

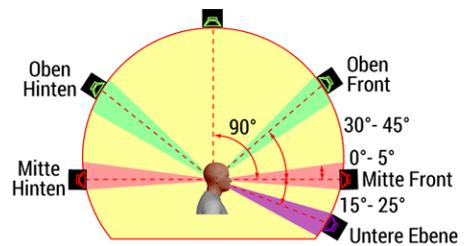


Abbildung 14: Positionen der Ebenen bei NHK 22.2

(Kompendium, 2016b)

## 5. Binauraltechnik

Bei einer normalen stereofonen Wiedergabe kann laut Spikofski keine Umhüllung des Hörers erreicht werden.

Mit einer binauralen Wiedergabe hingegen kann eine Umhüllung und die Abbildung verschiedener Entfernungen realisiert werden. Inhalte werden z.B. mit Hilfe von Kunstköpfen erzeugt. (vgl. zu diesem Abschnitt Spikofski, 2014, S. 345)

Die wohl einfachste Variante binaurales Audio-Material zu erstellen, ist die Aufnahme von akustischen Signalen direkt vor dem Trommelfell, wie es auch beim natürlichen Hören geschieht. Dies kann auf zwei Varianten geschehen: Entweder es werden Mikrofone direkt im Gehörgang eines Hörers platziert oder es kommt ein Kunstkopf zum Einsatz, welcher bereits Druckempfänger an den richtigen Stellen installiert hat.

(vgl. zu diesem Abschnitt Weinzierl, 2008, S. 586)

Die Form des Kunstkopfes ist sowie „[...]die akustischen Eigenschaften wie Richtcharakteristik, Pegel- und Laufzeitdifferenzen, dem menschlichen Kopf weitestgehend nachgebildet“ (Spikofski, 2014, S. 345).

Nur besitzt der Kunstkopf Mikrofone anstelle von Trommelfellen. (vgl. Spikofski, 2014, S. 345)  
Auf diese Weise kommt es ohne künstliches Eingreifen zu einer Filterung durch die Außenohrübertragungsfunktion<sup>11</sup> (AOÜF), „[...]welche für jede Schalleinfallrichtung die Wirkung unserer Kopfanatomie auf den einfallenden Schall beschreibt“ (Weinzierl, 2008, S. 586).

Wird dieses aufgenommene akustische Ereignis nun über Kopfhörer wiedergegeben, wird somit im Idealfall das aufgenommene Schallfeld originalgetreu reproduziert. Die Filterung durch die AOÜF wird durch akustische Abschattung, Beugung, Verzögerung, Resonanzen und Reflexionen durch Torso, Schulter, Kopf, Ohrmuscheln (pinnae), den Eingang in den Ohrkanal (cavum conchae) und den Ohrkanal selbst definiert. Die am stärksten ausgeprägte Filterung in der Außenohrübertragungsfunktion findet durch Kopf und Ohrmuscheln statt. Doch auch Torso und Schulter beeinflussen das Signal bei bestimmten Frequenzen um etwa  $\pm 3$  dB bzw. 5 dB, was allerdings nicht bei allen Kunstkopfsystemen berücksichtigt wird. Dabei ist diese Frequenzbeeinflussung im Gegensatz zum Einfluss des Ohrkanals abhängig von der Schalleinfallrichtung. (vgl. zu diesem Abschnitt Weinzierl, 2008, S. 586)

---

<sup>11</sup> Oft auch englisch HRTF (head-related transfer function) genannt. (vgl. Weinzierl, 2008, S. 586)

Ein erhebliches Problem bei der Aufnahme von binauralen Signalen ist der anatomische Unterschied zwischen verschiedenen Personen, welcher zu erheblichen Unterschieden in den Außenohrübertragungsfunktionen führt. So gelingt eine Aufnahme mit dem eigenen Kopf in der Regel besser, als Signale über einen der eigenen Anatomie fremden Kunstkopf, aufzunehmen. Diese Kunstköpfe stellen in der Regel den menschlichen „Durchschnittskopf“ dar, zu welchem eine Person mehr oder weniger kompatibel sein kann.

Kunstköpfe für binaurale Aufnahmen gibt es in den verschiedensten Ausführungen für verschiedene Anwendungen. So wird bei Musik- und Sprachaufnahmen überwiegend auf den Kunstkopf *KU100* der Firma Neumann zugegriffen, wie er in Abbildung 15 zu sehen ist. In der akustischen Messtechnik, wo zum Beispiel Messungen zu binauralen akustischen Kriterien im Bereich *Sound Quality* oder beim akustischen Produktdesign erfolgen, werden meist andere Kunstkopfsysteme, wie der Brüel & Kjær *HATS 4100*, eingesetzt. Einige Modelle, wie die der Firma Kemar, werden für verschiedene Außenohrtypen angeboten, welche typisch für weibliche, männliche, amerikanische, europäische oder asiatische Hörer sind. (vgl. zu diesem Abschnitt Weinzierl, 2008, S. 587)



Abbildung 15:  
Neumann *KU100*,  
*Dummy Head*

(Neumann, 2016)

Historisch wurden bereits nach dem Ende der 1960 Jahre die ersten Kunstkopfsysteme für Untersuchungen zur Beurteilung der Hörsamkeit in Konzertsälen entwickelt und erste binaurale Hörspiele von Rundfunksendern veröffentlicht, wie das im Jahr 1973 erschienene *Demolition*, das von Rezipienten euphorisch aufgenommen wurde. Durch Klangverfärbungen bei der Wiedergabe auf Lautsprechern haben sich binaurale Kunstkopfaufnahmen im Musikbereich dagegen als wenig sinnvoll erwiesen. Auch ist es bei einer binauralen Musikaufnahme nicht möglich, die Klangbalance des Stückes ohne eine zusätzliche nahe Mikrofonierung zu kontrollieren, was die Möglichkeit zur guten Lokalisation von Schallquellen im Raum wiederum aushebelt. In den Bereichen Psychoakustik, Lärmschutz und Lärmwirkungsforschung stellt die Kunstkopfaufnahme hingegen ein wichtiges Werkzeug dar.

Ein weiterer Nachteil der Kunstkopfaufnahme ist die Schwierigkeit zwischen frontalen und rückwertigen Schalleinfallrichtungen mit gleicher interauraler Laufzeitdifferenz zu unterscheiden. Dieser Nachteil kann jedoch mit einem Positionssensor, welcher bei der Wiedergabe die Peilbewegungen des Kopfes abtastet, ausgeglichen werden.

(vgl. zu diesem Abschnitt Weinzierl, 2008, S. 589)

## 5.1. Software mit Möglichkeit zur binauralen Simulation

Neben der direkten Erstellung von Audio-Inhalten für die binaurale Wiedergabe, wie in Kapitel 5. erläutert, können mittels speziellen Downmix-Methoden Mehrkanalproduktionen räumlich auf Kopfhörern wiedergegeben werden, womit eine Lokalisation außerhalb des Kopfes stattfindet und somit Raum und Tiefe wahrgenommen werden können. (vgl. Stoll/Hartmann, 2014, S. 1244) Diese Technik wird BRS (Binaural Room Synthesis) genannt und erlaubt reale, wie auch modellierte Abhörsituationen als Abhörräume mit verschiedenen Lautsprecherkonfigurationen zu auralisieren.<sup>12</sup> Um dies zu realisieren, werden mittels Kunstkopfverfahren Außenohrübertragungsfunktionen sowie binaurale Raumimpulsantworten des Abhörraums gemessen, was zu einer virtuellen Darstellung des gemessenen Raums führt. Durch Faltung der Eingangssignale mit den gemessenen Daten wird dann schließlich die virtuelle räumliche Wahrnehmung erzeugt. (vgl. zu diesem Abschnitt Hoeg, 2014, S. 1328)

*„Bei optimaler virtueller Reproduktion entsprechen die Ausgangssignale des BRS-Prozessors den Kunstkopfsignalen an der Hörposition im Abhörraum, die mit Hilfe des Kopfhörers an den Ohren des Hörers authentisch reproduziert werden.“*

*(Hoeg, 2014, S. 1328)*

Um Vorn/Hinten-Vertauschungen zu vermeiden sowie eine dynamische Außenohrübertragungsfunktion zu generieren, können mittels Head-Tracking spontane Kopfbewegungen des Hörers ausgewertet werden.

Eine wichtige Anwendung ist nach Hoeg der virtuelle Surround-Sound Abhörraum. Dieser kann mittels dieses Systems ohne großen finanziellen Aufwand oder entsprechende Räumlichkeiten simuliert werden. (vgl. zu diesem Abschnitt Hoeg, 2014, S. 1328f)

Im folgenden Kapitel werden einige Programme vorgestellt, welche, vor allem in Bezug auf die 3D-Wiedergabe, Abhörräume mit verschiedenen Lautsprecherkonfigurationen simulieren können. Darunter sind sowohl Panoramaregler mit Möglichkeit zur binauralen Auralisation, als auch Anwendungen für spezielle Bereiche.

---

<sup>12</sup> Hörbarmachung von Raumimpulsantworten. (Ahnert, Tennhardt, 2008, S. 247)

### 5.1.1. New Audio Technology Spatial Audio Designer:

Das von Tom Ammermann entworfene Plug-in *Spatial Audio Designer* ist nicht ausschließlich zur binauralen Auralisation gedacht. Es kann als Panoramaregler zur Herstellung von Surround- sowie 3D-Mischungen verwendet werden und ist dabei plattformübergreifend und unabhängig vom gewünschten Format nutzbar. Besonders ist darüber hinaus die Fähigkeit, ausgewählte Lautsprecherkombinationen im Plug-in virtualisieren zu lassen. So können laut Ammermann mit *Headphone Surround 3D™* Lautsprecher-Virtualisationen für handelsübliche Kopfhörer erzeugt werden.

Neben allen derzeit auf dem Markt verfügbaren Audio-Formaten, wie z.B. *Auro 3D*, stehen auch ungewöhnliche Formate wie 12.1 oder 13.1 zur Verfügung. Abhängig von der Wahl des Audio-Formats können darüber hinaus verschiedene Räume mit unterschiedlicher Akustik abgerufen werden.

Da keine komplette Implementierung in Pro Tools oder anderen DAWs vorhanden ist, ist es notwendig verschiedene Plug-ins zu verwenden.

Hierzu wird das *Mix-Modul* (Abbildung 16) als Plug-in in eine Stereo Aux-Spur eingefügt, welche als Ausgang den Weg zum Kopfhörer definiert hat. Im *Mix-Modul* kann die Lautsprecherkonfiguration gewählt werden, welche in Presets abgespeichert ist. Die Auswahl des Presets ist bei der Bearbeitung jederzeit änderbar und wirkt sich nicht auf die Positionen der im *Mix-Modul* befindlichen Objekte aus. Jede Spur, die als Objekt im *Mix-Modul* angezeigt und positioniert werden soll, benötigt hierzu ein eingefügtes *Send-Modul*, das dem gewünschten *Mix-Modul* zugeordnet werden muss, damit es in der Liste des *Mix-Moduls* angezeigt wird. Außerdem ist es möglich mit mehreren *Mix-Modulen* gleichzeitig zu arbeiten, um parallel verschiedene Räume oder Lautsprecherkonfigurationen abzuhören. Damit das Zusammenspiel der einzelnen Plug-ins reibungslos funktioniert, ist es laut Ammermann wichtig, dass das Processing der Plug-ins in der richtigen Reihenfolge abläuft: Erst die *Send-Module* und anschließend die *Mix-Module*. Dies wird in Pro Tools erreicht, indem die

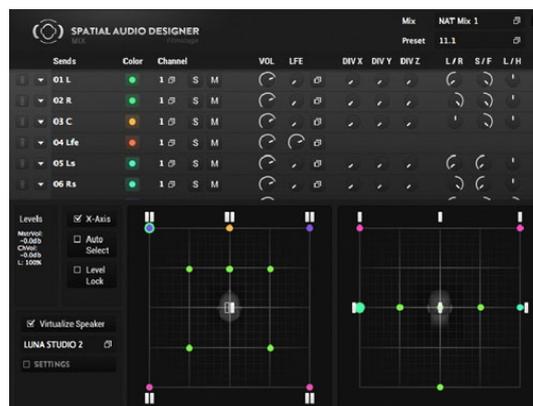


Abbildung 16: *Spatial Audio Designer: Mix-Modul*

(New Audio Technology, 2016a)

mit *Send-Modulen* ausgestatteten Spuren einen Bus als Ausgang zugewiesen bekommen, welcher den *Mix-Modulen* als Eingang dient.

Im *Mix-Modul* können nun die einzelnen Objekte auf der horizontalen sowie auf der vertikalen Ebene angeordnet werden. Auch eine stufenlose Pegelzuweisung einzelner Objekte auf den LFE ist möglich, falls die ausgewählte Lautsprecherkonfiguration diesen beinhaltet. Ebenso kann das Lautstärkeverhältnis der einzelnen Quellen im *Mix-Modul* angepasst werden. Eine Divergenz der einzelnen Quellen in X-, Y- und Z-Richtung kann ebenfalls realisiert werden, sodass die Quellen ihre Position in regelbarem Maße vergrößern. (vgl. zum gesamten Kapitel New Audio Technology, 2016b)

### 5.1.2. StiffNeck

Die Software *StiffNeck* wurde an der *University of Music and Performing Arts* in Graz entwickelt und stellt ein Werkzeug zur Auralisation von Lautsprecherkonfigurationen dar, die anhand von gemessenen Raumimpulsantworten

berechnet werden. *StiffNeck* wurde vor allem dafür entwickelt, ungewöhnliche Lautsprecher-Konfigurationen zu messen, welche für *in situ* Kompositionen eingerichtet wurden. Typisch für diese Art von Komposition ist, dass der Prozess derselben dort stattfindet, wo das Endergebnis präsentiert wird. Die Kernidee ist also, Künstlern die Möglichkeit zu verschaffen, auch an ihren Kompositionen zu arbeiten, wenn sie keinen Zugang zur



Abbildung 17: Motorisierte Lautsprecherkonfiguration mit Messaufbau

(Eckel, Rumori, 2014, S. 3)

konkreten Halle haben, für die sie eine Komposition entworfen haben. Die ungewöhnlichen Lautsprecherkonfigurationen wurden in der *György Ligeti Hall* in Graz gemessen. Diese bietet mehr als 100 fest installierte Lautsprecher, von denen 33 motorisch steuerbar sind, womit die Möglichkeit besteht in weniger als zwei Minuten die Lautsprecherkonfiguration zu ändern. Darüber hinaus bietet *StiffNeck* laut Eckel und Rumori eine große Anzahl von herkömmlichen

Lautsprecherkonfigurationen, deren Raumimpulsantworten ebenfalls in der *György Ligeti Hall* in Graz von Mikrofonarrays gemessen wurden. Diese können nach Belieben in ihren Eigenschaften wie Hörposition, Art des Mikrofonarrays oder Lautsprecher-Konfiguration, verglichen werden.

Außerdem ist *StiffNeck* in der Lage, Messpositionen und Orte der gemessenen Lautsprecher geometrisch zu visualisieren.

Praktisch zu verwenden ist *StiffNeck* als unabhängige Software oder im Rahmen der Sound-Programmierungsumgebung *SuperCollider Quark*. Ein- und Ausgänge werden über den *Jack audio server*<sup>13</sup> realisiert. (vgl. zum gesamten Kapitel Eckel/Rumori, 2014, S. 1-4)

### 5.1.3. Logic Pro 9

*Logic Pro 9* ist ein komplettes Audio- und Midi-Bearbeitungsprogramm, das als vollausgestattete DAW bezeichnet werden kann.

Obwohl *Logic Pro 9* kein Panoramaregler für 3D-Audio zur Verfügung stellt, bietet es die Möglichkeit, Signale in einem virtuellen 3D-Panorama zu verteilen. Hierzu wird kein externes Plug-in benötigt, da der binaurale Panoramaregler bereits in *Logic Pro 9* integriert ist. Hierzu muss der Output des Channel-Strips auf die Option „Binaural“ eingestellt werden. Ist dies ausgewählt, erscheint in der Panning-Ebene das Fenster *Binaural Panner*. (vgl. zu diesem Abschnitt Apple Inc., 2011, S. 848)



Abbildung 18: Logic Pro 9 Binaural Panner

Der *Binaural Panner* ermöglicht es, Signalquellen mit Hilfe von „Richtungs-Pucks“ in der Panning-Ebene zu verteilen. Hier ist es möglich aus zwei verschiedenen Modi zu wählen. Ist der Modus *Planar* ausgewählt, lässt sich das Signal auf einer Scheibe um den Hörer positionieren. Wählt man die Einstellung *Spherical* besteht die Möglichkeit, Quellen im dreidimensionalen Raum um die Hörer herum zu bewegen. Hier wird die Panning-Ebene als virtuelle Kugel veranschaulicht. Die Stereo-Breite eines Audiosignals kann ebenfalls mit Hilfe von Richtungs-Pucks verändert werden. Ebenso ist es möglich, den sogenannten Doppler-Effekt an- bzw. abzuwählen, welcher zum Tragen kommt, falls eine Bewegung der Quelle auf den Hörer zu oder von ihm weg stattfindet. (vgl. zu diesem Abschnitt Apple Inc., 2011, S. 850ff)

(Apple Inc., 2011, S. 852)

<sup>13</sup> Software-Schnittstelle für Audio-Software. (vgl. Jack Audio, 2016)

## 5.1.4. Auratorium

*Auratorium* basiert auf den theoretischen Grundlagen der Echtzeit-Simulationssoftware RAVEN, welche von Sönke Pelzer und Dr. Dirk Schröder am Institut für Technische Akustik der RWTH Aachen entwickelt wurde. Bei *Auratorium* handelt es sich um eine Neu-Implementierung mit einer intensiven Weiterentwicklung hinsichtlich der 3D-Audiowiedergabe und bietet so eine vollwertige Schallfeldrekonstruktion und Möglichkeit zur binauralen Auralisation. Darüber hinaus wurde die Simulationsgeschwindigkeit sowie die Performance und das User-Interface deutlich verbessert. Unterstützt werden Standard-Kopfhörer sowie beliebige Lautsprecher-Layouts, inklusive aller üblichen Surround-Formaten.

Der Simulationskern von *Auratorium* steht als eigenständige Software bereit, die über Plug-ins angesteuert wird. Die Plug-ins erfüllen dabei den Zweck *Auratorium* an Drittanbietersoftware zu koppeln. (vgl. zu diesem Abschnitt Schröder, 2016)

So ist es möglich mit herkömmlichen CAD-Programmen virtuelle Räume zu erstellen, die dann u.a. mittels Raytracing auralisiert werden. Die Fernsteuerung kann z.B. über die Schnittstelle *OpenSoundControl* (OSC) erfolgen und ist daher auch mit einem iPad oder Android-Gerät verwendbar. (vgl. zu diesem Abschnitt Audioborn, 2015a)

Außerdem ist es möglich die Software mit Sequencern zu steuern und Quellbewegungen zu automatisieren.

Die Anwendungen von *Auratorium* sind neben dem Bereich 3D-Audio Produktion vielfältig. So sind Anwendungen wie eine Optimierung der Raumakustik im Bereich der Architektur oder der Einsatz im Sachgebiet *Virtual/Augmented Reality* möglich. Weiter können Techniken wie Head-Tracking oder HMDs (Head-Mounted Display) z.B. von der Firma Oculus eingesetzt werden. (vgl. zu diesem Abschnitt Schröder, 2016)

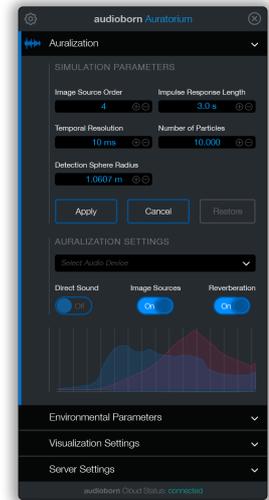


Abbildung 19:  
*Auratorium Raumakustik  
Simulation*

(Audioborn, 2015b)

## 6. Hörversuch

Der folgende Hörversuch vergleicht ein *Iosono* 3D-Audiosystem in der Lautsprecherkonfiguration 23.1 mit einer binauralen Simulation des *Spatial Audio Designers* von New Audio Technology.

Beide Formate sind in den Kapiteln 4.3.2. bzw. 5.1.1. beschrieben.

Im Hörversuch soll untersucht werden in welchem Maße sich die Qualität einer authentischen Raumwiedergabe bei beiden Wiedergabeformen unterscheidet.

Als Testsignal wurde eine Live-Aufnahme der Band *Culcha Candela* gewählt. Das Konzert fand im Rahmen der ZDF Konzertreihe *zdf@bauhaus* am 23. September 2015 im Bauhaus Dessau statt. Der Zuschauerraum wurde neben einer aufwendigen Stereo bzw. 5.1 Atmo-Mikrofonierung in 3D aufgezeichnet.

Aus der Aufnahme wird das Lied namens *Nie Genug* ausgewählt. Dieses Stück besitzt, im Gegensatz zu den anderen Stücken des Konzerts, mehr Publikumsgeräusche, welche für ein räumliches Hörerlebnis klar von Vorteil sind.

In den folgenden Kapiteln werden der Konzertsaal und die Mikrofonierung detailliert beschrieben. Später folgt eine Beschreibung der verwendeten Wiedergabetechnik, mit der der Hörversuch realisiert wird.

## 6.1. Bauhausbühne Dessau

Der Konzertsaal des Bauhaus Dessau kann auf eine lange Geschichte zurückblicken. So fanden hier schon in den 1920er Jahren legendäre Tanz- und Theaterexperimente statt.

(vgl. zu diesem Abschnitt Bauhaus Dessau, 2016)

Dass dieser Konzertsaal trotz seines Alters jedoch auch in die moderne Zeit passt, beweist das von *TVT.media* im Januar 2011 ins Leben gerufene

Fernsehformat *zdf@bauhaus*, welches für die Fernsehsender *zdf.kultur* und *3sat* produziert und als „[...]Laudatio auf die Bühnengeschichte des Bauhaus Dessau“ angesehen wird (Bauhaus Dessau, 2016).

Hier werden auf einer historischen Bühne Live-Konzerte von internationalen Künstlern unmittelbar vor einem kleinen Publikum von circa 150 Zuschauern dargeboten. Diese Konzerte werden im Anschluss auf *zdf.kultur* und *3sat* gezeigt.

(vgl. zu diesem Abschnitt Bauhaus Dessau, 2016)

Der in Abbildung 20 grün umrandete Publikumsbereich diente ursprünglich als Backstage-Bereich, während der ehemalige bestuhlte Zuschauerbereich hinter der Bühne bei den Konzerten ausschließlich für den Auftritt bzw. Abgang der Band genutzt wird.

Akustisch gesehen ist der Zuschauerraum aufgrund der parallel strukturierten Wände nicht optimal. Auch die an einer Seite angebrachte Fensterwand bietet keine guten akustischen Eigenschaften. Um die

Akustik zu verbessern, wurden die Fenster mit schallabsorbierendem Molton verkleidet und an der Decke Akustikelemente zur Verbesserung des Schallverlaufs angebracht.

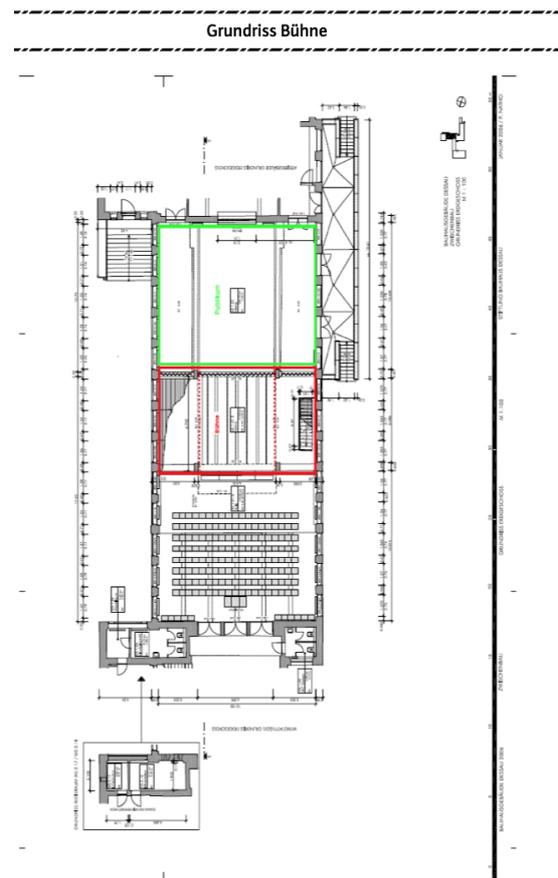


Abbildung 20: Konzertsaal im Bauhaus Dessau.

Rot: Bühne, grün: Zuschauerbereich.

(Bauhaus Dessau, 2006)

## 6.2. Mikrofonierung

Da die Konzertproduktion nicht primär als 3D-Audio-Vorhaben geplant war, sondern im Stereo- bzw. 5.1-Format vermarktet wurde, kam neben der eigentlichen 3D-Mikrofonierung eine klassische Atmo-Mikrofonierung für 5.1, welche zu Stereo abwärtskompatibel ist, zum Einsatz. Das 5.1 Mikrofonarray bot dabei allerdings die Basis für die 3D-Mikrofonierung.

### 6.2.1. Mikrofonierung für 5.1 bzw. Stereo

Um Publikumsreaktionen wie Applaus oder sonstige Geräusche aufzunehmen, wurden an jeder Seite der Bühne drei Mikrofone installiert. Die Platzierung der Mikrofone links und rechts am Bühnenrand können in der späteren Mischung als Erzeugung einer großen Stereobreite genutzt werden. Für diese Mikrofone wurden zwei verschiedene Richtcharakteristika eingesetzt: Um den unmittelbaren Zuschauerbereich abzudecken, wurden zwei Mikrofone mit Nierencharakteristik verwendet (Sennheiser MKH 40). Jeweils eines befand sich vorne direkt neben der Bühne, während an den Seitenwänden rechts und links ebenfalls ein solches Mikrofon befestigt war. Durch die Verwendung von zwei Mikrofonen für diesen Zweck wurde sichergestellt möglichst die gesamte Breite der jeweiligen Publikumsseite aufzunehmen. Um jedoch auch die Geräusche des mittig bis hinten sitzenden Publikums aufzuzeichnen, kam jeweils links und rechts ein Mikrofon mit Keulen- bzw. Supernierencharakteristik (Sennheiser MKH 60) zum Einsatz. Da diese aufgrund ihrer engen Richtcharakteristik jedoch nur den seitlichen Bereich abdeckten, wurden zusätzlich mittig über der Bühne zwei weitere Sennheiser MKH 60 Mikrofone auf einer Traverse installiert. Diese zeichneten zusätzlich den mittigen hinteren Zuschauerbereich auf. Die Richtcharakteristika Niere und Keule, welche nach hinten weitestgehend den Schall abdämpfen, wurden gewählt, da es sehr wichtig ist nur die Geräusche des Publikums aufzunehmen und den Lautstärkepegel der PA bzw. der Band möglichst auszublenden.

Die Mikrofonierung der Band wurde mit einer nahen Abnahme verwirklicht. So wurden neben gängigen Mikrofonen für den Gesang Ansteckmikrofone für die Bläsersektion genutzt. Das Schlagzeug wurde ebenfalls nah und mit zusätzlichen Overhead-Mikrofonen abgenommen.

## 6.2.2. 3D-Mikrofonierung

Die Hauptaufgabe der Mikrofonierung war in erster Linie den Publikumsbereich so aufzunehmen, dass er später möglichst realistisch reproduziert werden kann. Dafür war es notwendig die einzelnen Kanäle so diskret wie möglich aufzuzeichnen, da in der späteren Reproduktion kein Direktschall von oben ausgehen sollte. Von oben sollte später nur die Rauminformation, die durch Reflexions- und Hallanteile definiert ist, zu hören sein.

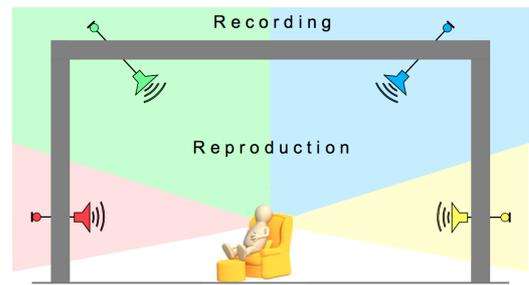


Abbildung 21: Produktion und Reproduktion einer Raumakustik

(Theile/Wittek, 2012, S. 11)

Um eine möglichst natürliche Aufnahme zu erhalten, ist es laut Theile und Wittek notwendig, direkten Schall möglichst diskret von Hall und Reflexionen zu trennen, um so ein Übersprechen zu verhindern. Mit einem Blick auf die spätere Reproduktion bei der Wiedergabe, wie auf Abbildung 21 zu sehen, wird klar, wie die Mikrofone platziert werden. (vgl. Theile und Wittek, 2012, S. 11f)

So bestand die eigentliche 3D-Mikrofonierung aus zwei Ebenen, welche später jeweils den zwei Ebenen des *Iosono* 3D-Systems<sup>14</sup> zugewiesen werden sollten.

Um die Mikrofonanordnung wie in Abbildung 21 zu realisieren, diente als Grundlage der ersten Ebene ein 5.1-Mikrofonarray, welches den Direktschall aus vier verschiedenen Richtungen aufnehmen sollte. Der Center-Kanal wurde später bewusst beim Mischvorgang vernachlässigt, da er aufgrund seiner Ausrichtung zur Bühne hauptsächlich den Pegel des PA Systems aufgezeichnete.

Zum Einsatz kam hier das *Adjustable Surround Microphone 5* (ASM 5) der Firma Brauner, welches mit Brauner *VM-1* Kapseln ausgestattet ist.

Diese sind direkt mit einem Multicore-Anschlusskabel<sup>15</sup> an die Steuereinheit *Atmos 5.1* der Firma SPL electronics GmbH angeschlossen.

So sind die Mikrofone unmittelbar den festgelegten Surround-Kanälen zugeordnet. Die Steuereinheit bietet die

dazugehörigen Vorverstärker. (vgl. zu diesem Abschnitt SPL electronics GmbH, 2003, S. 3f)



Abbildung 22: Atmos 5.1 Controller & Brauner ASM 5 Mikrofon

(SPL, 2016)

<sup>14</sup> Diese Mikrofonanordnung ist auch auf andere kanal- und objektbasierte Systeme anwendbar.

<sup>15</sup> „Kabel mit mehreren Signalleitungen“ (Friesecke, 2014, S. 899)

Die zweite Ebene bildeten vier jeweils auf den vier äußeren Mikrofonen des 5.1-Mikrofonarray befestigte Mikrofone mit Nierencharakteristik, wofür die Kapseln *MK 41* der Firma Schoeps verwendet wurden. Diese wurden, wie in Abbildung 23 zu sehen, im  $45^\circ$  Winkel direkt auf den einzelnen Mikrofonen des *ASM 5* platziert.

Eine Platzierung der Mikrofone beider Ebenen ohne Abstand zueinander erfolgte aus folgendem Grund:

Bei einer Aufnahme mit einer horizontalen Stereo-Mikrofonierung ist es bekannt, dass ein größerer Abstand beider Mikrofone zueinander einen, aufgrund der Wechselwirkung beider Kanäle, besseren räumlichen Eindruck erzeugt. Im Gegensatz dazu ist diese Wechselwirkung in der vertikalen Ebene weniger effektiv als in der horizontalen Ebene. Folglich bietet ein Abstand der Mikrofone in der Vertikalen weder in der räumlichen Abbildung noch in der Tonqualität Vorteile.

(vgl. zu diesem Abschnitt Lee, 2014, S. 6)

Um die 3D-Mikrofonierung abzurunden, wurden an der Publikumsrückseite (Bei Abbildung 24 rechts im Bild) zwei weitere Mikrofone mit Kugelcharakteristik installiert, um den räumlichen Schall im hinteren Bereich noch genauer abbilden zu können.

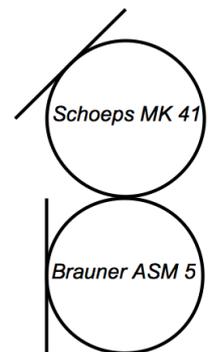


Abbildung 23:  $45^\circ$  Anordnung des Höhenmikrofons

(eigene Darstellung)

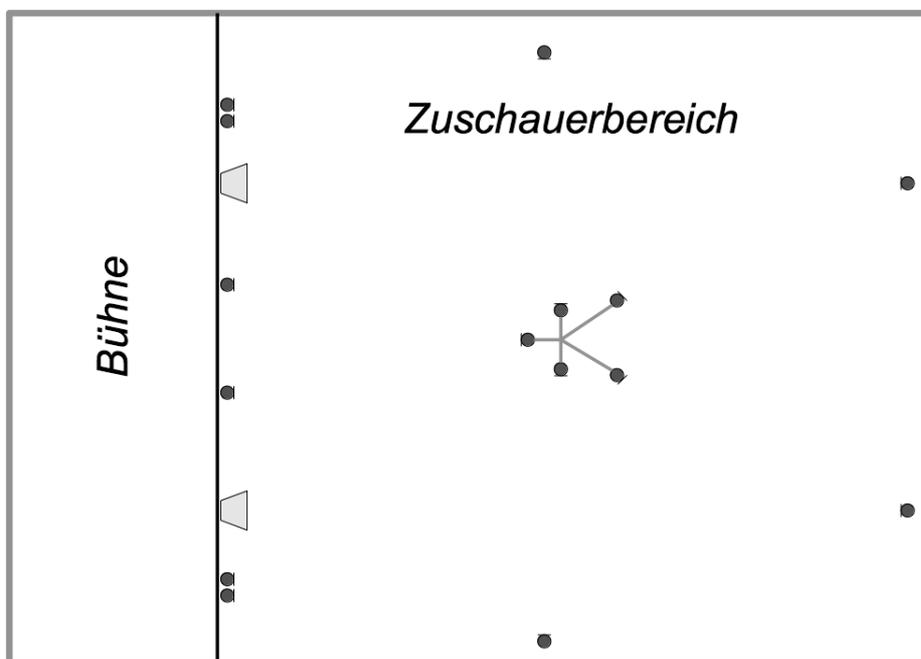


Abbildung 24: Atmo-Mikrofonierung. Das mittige 5.1 Array besteht aus zwei Ebenen.

(eigene Darstellung)

## 6.3. Postproduktion für den Hörversuch

Durch die Mikrofonierung analoger Schallquellen, Aufnahme von elektronischen Komponenten, Bearbeitung der Signale mit Delay- sowie Halleffekten und schließlich der Atmo-Mikrofonierung ergeben sich, wie in Abbildung 25 zu sehen, insgesamt 49 Spuren.

Um die Übersichtlichkeit besser zu gestalten, werden die Spuren in drei Kategorien untergliedert: Band, Atmo-Mikrofone und Effekte.

<b>Band</b>	<b>Atmo</b>	<b>Effekte</b>
Omar	K ML	D2 L
John	N ML	D2 R
Itchy	N MR	H3000 L
Chino	K MR	H3000 R
Curly	K L	EMT L
Trompete	N R	EMT R
Sax	Kugel L	Lexicon 1 L
Kick	Kugel R	Lexicon 1 R
Snare	N L	Lexicon 2 L
HH	K R	Lexicon 2 R
Tom	Atmos L	
Timbales	Atmos C	
Conga	Atmos R	
OH L	Atmos LS	
OH R	Atmos RS	
DJ L	Schoeps L	
DJ R	Schoeps R	
BV 1	Schoeps LS	
BV 2	Schoeps RS	
Spare		

Abbildung 25: Spurenplan; *Culcha Candela*

(eigene Darstellung)

### 6.3.1. Band

Die Band *Culcha Candela* setzt sich aus folgenden Tonquellen zusammen.

*Omar, John, Itchy, Chino* und *Curly* sind die Hauptsänger und werden von zwei Backing-Vocals (*BV 1* und *BV 2*) unterstützt.

In der Bläsersektion treten eine *Trompete* und ein Saxophon (*Sax*) auf.

Das Schlagzeug setzt sich aus den Instrumenten *Kick, Snare, High Hat (HH), Timbales, Tom, Conga* und den Becken zusammen, welche aufgrund der Mikrofonierung mit den Spuren Overhead links und rechts (*OH L/R*) repräsentiert werden.

Als elektronische Quelle fungiert ein DJ, welcher ein Stereo-Signal (*DJ L/R*) abgibt.

### 6.3.2. Atmo-Mikrofone

Wie bereits im Kapitel 6.2. beschrieben, kam neben einer normalen Stereo- bzw. 5.1 Mikrofonierung eine 3D-Mikrofonierung zum Einsatz. Die Kürzel *K* und *N* stehen im Spurenplan für die Richtcharakteristik Keule bzw. Niere. Als *Kugel L/R* sind zwei Schoeps *MK 2* aufgeführt, welche an der Rückwand des Konzertsaals installiert waren. Mit *Atmos* sind die fünf Mikrofone des 5.1 Arrays gemeint. *Schoeps* beschreibt die Mikrofontypen, die als zweite Ebene auf dem 5.1 Array installiert waren, als Schoeps *MK 41*.

### 6.3.3. Effekte

Um die Aufnahme zusätzlich noch räumlicher zu gestalten, kam während des beschriebenen Konzerts eine Reihe von Effektgeräten zum Einsatz, welche als Delay- oder Halleffekt für verschiedene Instrumente bzw. Gesänge verwendet wurden.

- *D2 L/R*: tc electronic D-Two
- *H3000 L/R*: Eventide H3000
- *EMT L/R*: EMT 251 Digital Reverberation System
- *Lexicon 1 L/R*: Lexicon 960
- *Lexicon 2 L/R*: Lexicon 960

## 6.4. Wiedergabetechnik im Tonstudio

Der mobile Ton-Übertragungswagen *Ü10* der Firma B&R Medientechnik in Kürten bietet alles, was zur Mischung und Bearbeitung von 3D-Audio-Inhalten notwendig ist. Er ist durch neueste 3D-Audiotechnik und durch eine komplette Schalldämmung mit schallabsorbierenden Wänden sowohl technisch als auch akustisch für beste Ergebnisse optimiert.

Für den Hörversuch wurden zwei Mischungen erstellt. Auf einem *Iosono 3D*-System der Firma Barco und im *Spatial Audio Designer* von New Audio Technology.

Im Folgenden wird der Signalfloss der verwendeten Technik veranschaulicht.

## 6.4.1. Konfiguration 1: Abhören über die 3D-Audio-Lautsprecherkonfiguration

Das Pro Tools HD10 System, das in der Version 10.3.10 verwendet wird, ist auf einem Macintosh Rechner bereitgestellt.

Dieses System, welches auch bei der Aufnahme der Signale eingesetzt wurde, dient bei der Mischung des Testsignals und bei der Vorführung des Hörversuchs als reines Abspielsystem der einzelnen Tonspuren.

Als Session dient die Aufnahme-Session, bei welcher der gesamte Mehrspurmitschnitt zur Verfügung steht.

Um die Verbindung zwischen Pro Tools und dem Mischpult Lawo MC80 zu gewährleisten, bedarf es eines MADI Interface. Zum Einsatz kommt hier ein *Xlogic Delta-Link MADI HD* der Firma SSL, womit 64 Kanäle über MADI übertragen werden können.

Um die Tonsignale weiter auf den Prozessor, der dann die Verstärker der Lautsprecher ansteuert, zu verteilen, muss im Falle des *Iosono*-Systems ein Umweg über die DAW Nuendo gemacht werden.

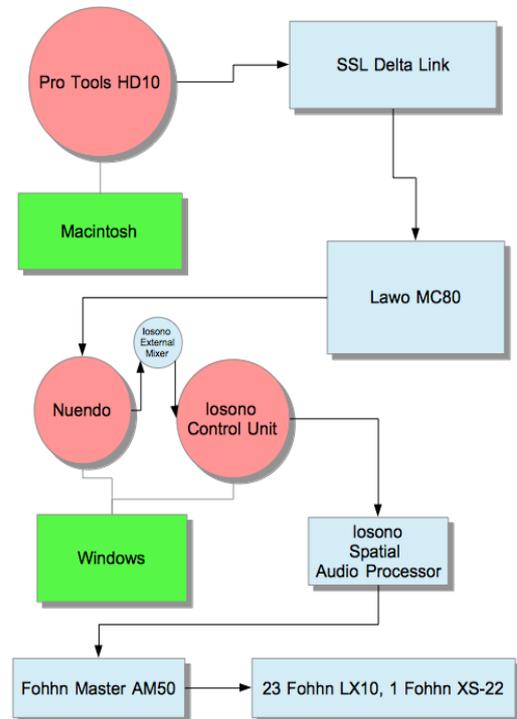


Abbildung 26: Signalverlauf Iosono

(eigene Darstellung)

Nuendo, welches auf einem Windows-System arbeitet, erhält die Audio-Signale und gibt diese über ein Plug-in namens *Iosono External Mixer* inklusive der Positionsdaten der hier erstellten Audio-Objekte an den *Iosono Spatial Audio Processor* weiter. Der *Iosono Spatial Audio Processor* wird aus Komfortgründen von der *Iosono Control Unit*, welche ebenfalls auf dem Windows Rechner ausgeführt wird, gesteuert.

Der *Iosono Spatial Audio Processor* berechnet die Verteilung der Audio-Objekte an die Lautsprecher, welche über zwei Schnittstellen (AM-40 AIREA Master Modul) mit der Systemendstufe *Fohhn Master AM50* über AES/EBU verbunden sind.

Für die Matrix-Verwaltung wird ein RME *Lock* verwendet.

Als Lautsprecherkonfiguration ist, wie in Abbildung 27 zu sehen, ein 23.1-Format installiert, welches in drei Ebenen funktioniert. Die erste Ebene bildet einen Kreis auf Ohrhöhe und ist mit

14 Lautsprechern bestückt. Die zweite Ebene bildet über Kopfhöhe ebenfalls einen Kreis und ist mit sieben Lautsprechern versehen, welche leicht nach unten geneigt sind. In der Decke sind als dritte Ebene schließlich noch zwei weitere *Voice of God*-Lautsprecher installiert, die auf einer horizontalen Linie angeordnet sind. Zum Einsatz kommen hierbei 2-Wege-Lautsprecher mit der Bezeichnung *LX-10 ASX* der Fohhn *AIREA* Serie, die eigene DSP besitzen und netzwerkseitig über AES/EBU angesteuert werden. Für den tieffrequenten Bereich sorgt ein *XS-22 Active Subbass*, ebenfalls von Fohhn.

Folgende Abbildung veranschaulicht die detaillierte Anordnung der Lautsprecher.

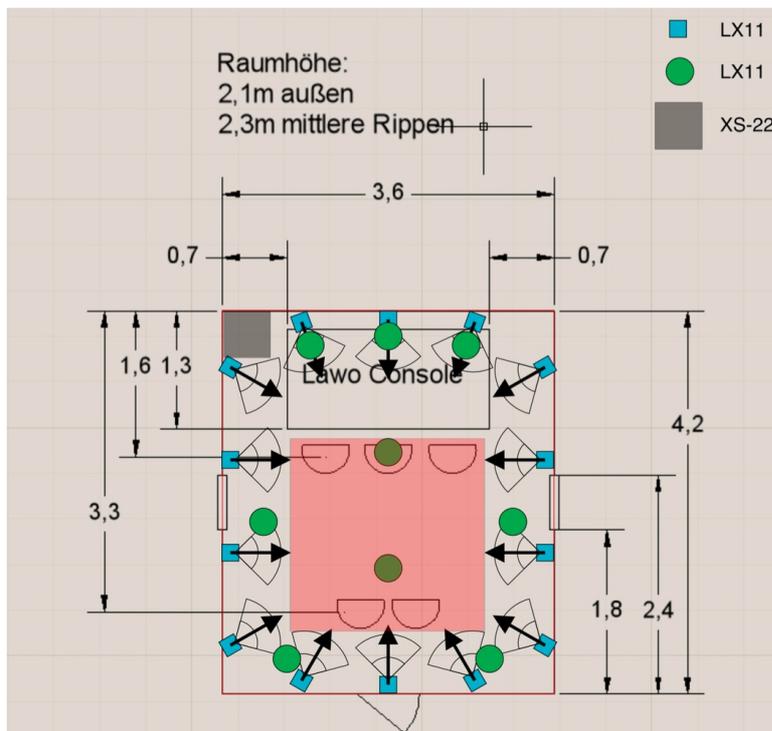


Abbildung 27: 23.1 Lautsprecheranordnung

blau: Ebene 1, grün: Ebene 2, moosgrün: Deckenlautsprecher

(Kugler, 2015, S. 27)

## 6.4.2. Konfiguration 2: Abhören über die 3D-Kopfhörersimulation

Für die Simulation der 3D Mischung wird das Plug-in *Spatial Audio Workstation* von New Audio Technology verwendet.

Theoretisch hätte es auch funktioniert dieses Plug-in im Pro Tools HD10 System zu benutzen. Es hat sich allerdings gezeigt, dass das Plug-in bei über 50 Spuren zu viel DSP belegt, sodass das Pro Tools HD10 System es nicht schafft in gleicher Zeit alle Spuren abzuspielen sowie die Berechnungen aller Plug-ins zu gewährleisten.

Da dies die Stabilität des bewährten Systems potentiell gefährden kann, werden die Berechnungen auf ein zweites Pro Tools 10 *Native* System ausgelagert.

Der Signalverlauf ist bis zum Mischpult mit dem der Lautsprecherwiedergabe identisch. So dient das

Pro Tools HD10 System lediglich als Abspielsystem des Mehrspurmitschnittes. Das Pro Tools 10 *Native* System ist über ein RME *MADIface* und eine *HDSP Breakoutbox* mit dem *Lawo MC80* verbunden und bekommt aufgrund der Beschränkung<sup>16</sup> von 32 Kanälen nicht alle einzelnen Spuren, sondern einige Spuren als *Stems*<sup>17</sup> gebündelt. Der Rückweg zum Mischpult dient dem Zweck die Ausspielwege des Mischpultes für die Wiedergabe zu verwenden. Diese speisen den Behringer *Powerplay Pro 8* Kopfhörerverstärker, welcher seine Signale wiederum an Sennheiser Kopfhörer weitergibt. Für den Hörversuch werden ausschließlich das Ohr umschließende, offene und hochwertige Studiokopfhörer verwendet, da sich in der Praxis gezeigt hat, dass ein 3D-Audio-Erlebnis so deutlich besser funktioniert als mit geschlossenen Kopfhörer. Da der gesamte Signalfluss mit MADI verwirklicht wird, kommt hierfür eine Matrix-Verwaltung mit einem RME Lock zum Einsatz.

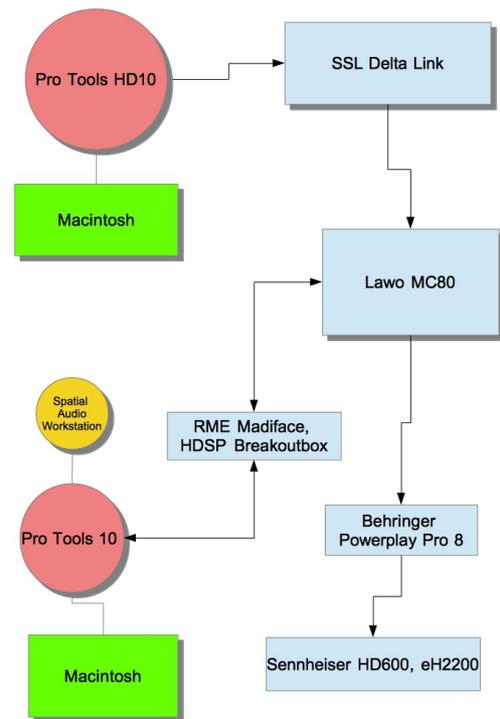


Abbildung 28: Signalverlauf Kopfhörer

(eigene Darstellung)

<sup>16</sup> Pro Tools 10 ist in seiner Native Version auf 32 Eingangskanäle beschränkt. (Avid Technology, 2016)

<sup>17</sup> „Zusammenfassung mehrerer Signale“ (Friesecke, 2014, S. 901)

## 6.5. Panning

Die Verteilung der Audio-Signale im 3D-Panorama erfolgt für das *Iosono* 3D-System in der *Iosono Spatial Audio Workstation*. Für die binaurale Mischung wird das Plug-in *Spatial Audio Designer* verwendet. Es soll eine möglichst identische Verteilung in beiden Programmen realisiert werden.

### 6.5.1. Lautsprecher: Iosono Spatial Audio Workstation

Die Bandaufstellung wird möglichst breit verteilt, damit die vordere Stereo-Breite vollständig ausgenutzt werden kann. Eine flächigere Wiedergabe wird erreicht, in dem Overhead- sowie Hi-Hat- und Backing-Vocal-Objekte in die zweite Ebene gepannt werden. Dies entspricht zwar nicht der realen Aufstellung, erzeugt aber eine gute Räumlichkeit. Obwohl es üblich ist die Gesangsmikrofone aus der Mitte zu hören, stellt sich heraus, dass es bei einer höheren Anzahl von Sängern, in diesem Fall fünf Hauptsänger, deutlich mehr Transparenz erzeugt, diese fast über die gesamte Stereo-Breite zu verteilen.

Bei den Atmo-Mikrofonen ist es wichtig eine möglichst originalgetreue Abbildung des Konzertraums zu schaffen. Daher sind diese an ihren jeweiligen Positionen um den Abhörraum im Ton-Übertragungswagen verteilt und sorgen für eine lückenlose Umhüllung mit Publikums- und Raumgeräuschen. Äquivalent zu ihrer Position im Publikumsbereich werden die Mikrofone *Kugel L/R* sowie die im 45° nach oben zeigenden 3D-Mikrofone *Schoeps L/R/LS/Rs* nach oben gepannt. Alle anderen Mikrofone erweisen auf der Ohrebene eine optimale, realistische Wiedergabe. Die Halleffekte werden zur Unterstützung der Räumlichkeit ebenfalls um den Hörer verteilt, während der Delay-Effekt, wie bei einem Live-Konzert üblich, von vorne zu hören ist. Im Allgemeinen ist es beim Panning laut Lee wichtig,

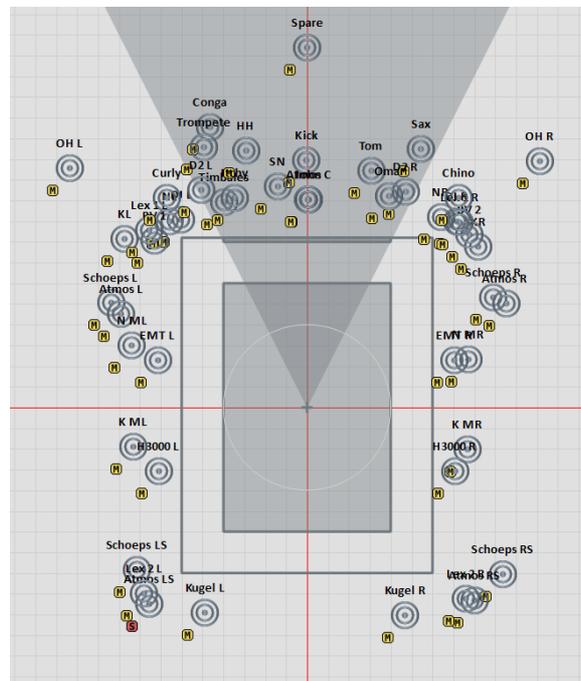


Abbildung 29: Verteilung der Audio-Objekte in der *Iosono Spatial Workstation*

keine tiefen Frequenzen in der oberen Ebene zu verteilen, da diese unabhängig von der physikalischen Lautsprecherposition am oder unter dem Ohr des Rezipienten geortet werden. (vgl. Lee, 2015, S. 4) Die Praxis zeigt zudem, dass das Pannen von tieffrequenten Signalen in eine höhere Ebene zu einem deutlich unschärferen Endergebnis führt.

Eine detaillierte Darstellung der einzelnen Audio-Objekte im Panorama ist im Anhang zu finden.

## 6.5.2. Binaural: New Audio Technology Spatial Audio Designer

Aufgrund der im Kapitel 6.4.2. beschriebenen Beschränkung von *Pro Tools 10 Native* auf 32 Kanäle wird im Rahmen des Versuchs die binaurale Simulation eines 9.1-Formates gewählt. Dies bietet den Vorteil, dass Audio-Objekte, die nicht zwischen zwei Lautsprecher gepannt werden sollen, im Mischpult auf neun Gruppen, welche den 9 Lautsprechern des 9.1-Formats entsprechen, verteilt werden.

Folglich stehen noch 23 Kanäle für frei verteilbare Objekte zur Verfügung.

Die Verteilung der Spuren erfolgt wie folgt:

Gruppe 1: *DJ L, Atmos L, K L, N L, Lexicon 1 L, BV 1*

Gruppe 2: *Spare, Atmos C* (beide nicht verwendet)

Gruppe 3: *DJ R, Atmos R, N R, K R, Lexicon 1 R, BV 2*

Gruppe 4: *Atmos Ls, Lexicon 2 L*

Gruppe 5: *Atmos Rs, Lexicon 2 R*

Gruppe 6: *OH L, Schoeps L*

Gruppe 7: *OH R, Schoeps R*

Gruppe 8: *Schoeps Ls, Kugel L*

Gruppe 9: *Schoeps Rs, Kugel R*

Die restlichen 23 Spuren sind als frei bewegliche Objekte nutzbar. Die Verteilung der Objekte erfolgt nach Anleitung des in der *Iosono Spatial Audio Workstation* angewandten Pannings und ist im Anhang näher veranschaulicht.

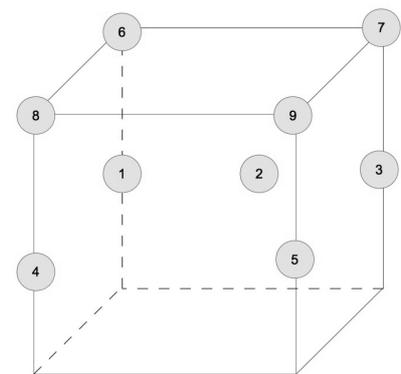


Abbildung 30: Gruppe 1 bis 9  
äquivalent zur 9.1  
Lautsprecheraufstellung.

(eigene Darstellung)

## 6.6. Signalbearbeitung

Die klangliche Bearbeitung der Quellen für den Hörversuch erfolgt ausschließlich im Mischpult Lawo *MC80*.

Bei der Mehrspuraufnahme ist im ersten Schritt eine Anpassung der Lautstärke-Verhältnisse notwendig.

Im nächsten Schritt erfolgt die Dynamikbearbeitung mit den Pult-eigenen Kompressoren. Dies ist vor allem bei den Stimmen wichtig, da die Sänger aufgrund der Party-Charakteristik der Musik sehr in ihrer Lautstärke schwanken. Auch bei Bläsern, Schlagzeug und Percussion wird in die Dynamik eingegriffen, um die lauten Stellen zu kontrollieren und sie teils besser hörbar zu machen.

Bei fast allen Atmo-Mikrofonen kommt ein Hochpass-Filter zur Anwendung, da der untere Frequenzbereich die Mischung negativ beeinflusst. Somit ist es ausreichend, die Rauminformationen, welche bei der Wiedergabe maßgeblich für die reale Konzertatmosphäre verantwortlich sind, erst ab circa 150 Hz wiederzugeben.

Mit dem Einsatz eines Equalizers, welcher vor allem für Gesang und Schlagwerk verwendet wird um störende Frequenzen zu eliminieren und als angenehm empfundene Frequenzen zu verdeutlichen, ist die Signalbearbeitung annähernd abgeschlossen.

Im Rahmen dieser Arbeit hat sich gezeigt, dass der Prozess der klanglichen Bearbeitung auf einem 3D-Audio-System deutlich weniger Zeit in Anspruch nimmt, als das Mischen von Stereo-Inhalten. Dies könnte an der natürlicheren Hörumgebung liegen, welche mit dem „geschaffenen Raum“ erzeugt wird.

Für die 3D-Kopfhörersimulation ist eine erneute Anpassung der Lautstärke-Verhältnisse notwendig. So muss die Lautstärke des Gesangs und des DJ deutlich angehoben werden, damit die Kopfhörermischung im Lautstärke-Verhältnis äquivalent zur Mischung des echten 3D-Audio-Systems ist. Somit ist ein besserer Vergleich gegeben.

Eine klangliche Anpassung findet, im Rahmen der Mischung für die Kopfhörersimulation, nicht statt.

## 6.7. Probanden

Um den Hörversuch so aussagekräftig wie möglich zu gestalten, nahmen sowohl Personen mit sowie Personen ohne berufliche Tonaffinität teil.

Die Aufteilung der beiden Personengruppen erfolgte paritätisch.

Die Altersspanne der Versuchspersonen reichte von 23 bis 55 Jahre.

## 6.8. Versuchsablauf

Nach Verteilung der Fragebögen an die Probanden und Klärung von Fragen hierzu, wurde jedem Proband ein Kopfhörer zugewiesen. Anschließend wurden sie über die Funktionen des Kopfhörerverstärkers instruiert. Diese umfassten den Lautstärkereger sowie den Umschalter

zwischen Eingang 1 und Eingang 2 des jeweils vom Probanden benutzten Kanals des Kopfhörerverstärkers.

Mit Umschalten der beiden Eingänge konnten die Probanden nun zwischen einer Stereo-Mischung, welche Live-on-Tape für die Fernsehauswertung von Toningenieur Ralf Freudenberg erstellt wurde und der von im Rahmen dieser Arbeit erstellten 3D-Mischung (eine simulierte 9.1 Lautsprecheraufstellung) umschalten.

Während bzw. nach dem circa dreiminütigen Musikstück hatten die Probanden die Möglichkeit ihre Antworten auf die Fragen anzukreuzen. Damit war die Vorführung der Kopfhörersimulation abgeschlossen.

Anschließend wurde ein circa 30 Sekunden dauernder Teil des Musikstückes abgespielt, bei dem zwischen der Stereo- und der 5.1-Abmischung (ebenfalls von Ralf Freudenberg) umgeschaltet wurde.

Für das Abhören der Stereo- sowie der 5.1-Abmischung kamen Lautsprecher der Firma Geithain zum Einsatz. Diese einleitende Vorführung diente dazu, den Probanden noch einmal zu veranschaulichen wie sich ein „normales“ Musik hören anfühlt um eine Referenz zur 3D-Anlage zu bekommen. Im Anschluss wurde an alle Probanden ein neuer Fragebogen verteilt, den sie während der Vorführung des *Iosono* 3D-Systems ausfüllen sollten. Kurz vor der Vorführung wurden die Versuchspersonen unterrichtet, dass sie den Fragebogen zum Kopfhörer nach Hören



Abbildung 31: Fünf Testpersonen hören die binaurale Kopfhörersimulation.

(eigene Fotografie)

der 3D-Lautsprecheranlage noch abändern dürften, da der Test einen Vergleich mit Referenz darstellt.

Bei der Vorführung des 3D-Lautsprechersystems war es äußerst wichtig, dass nur maximal vier Probanden gleichzeitig die Vorführung hörten, damit jene möglichst mittig sitzen konnten, um damit ein optimales Hörerlebnis zu bekommen.

Die Versuchsdauer variierte, je nach Ausführlichkeit des Feedbacks der Probanden, zwischen 30 und 50 Minuten.

Die Fragebögen können im Anhang eingesehen werden.

## 6.9. Auswertung des Fragebogens

Das Balkendiagramm zeigt jeweils die Durchschnittswerte der Testergebnisse an.

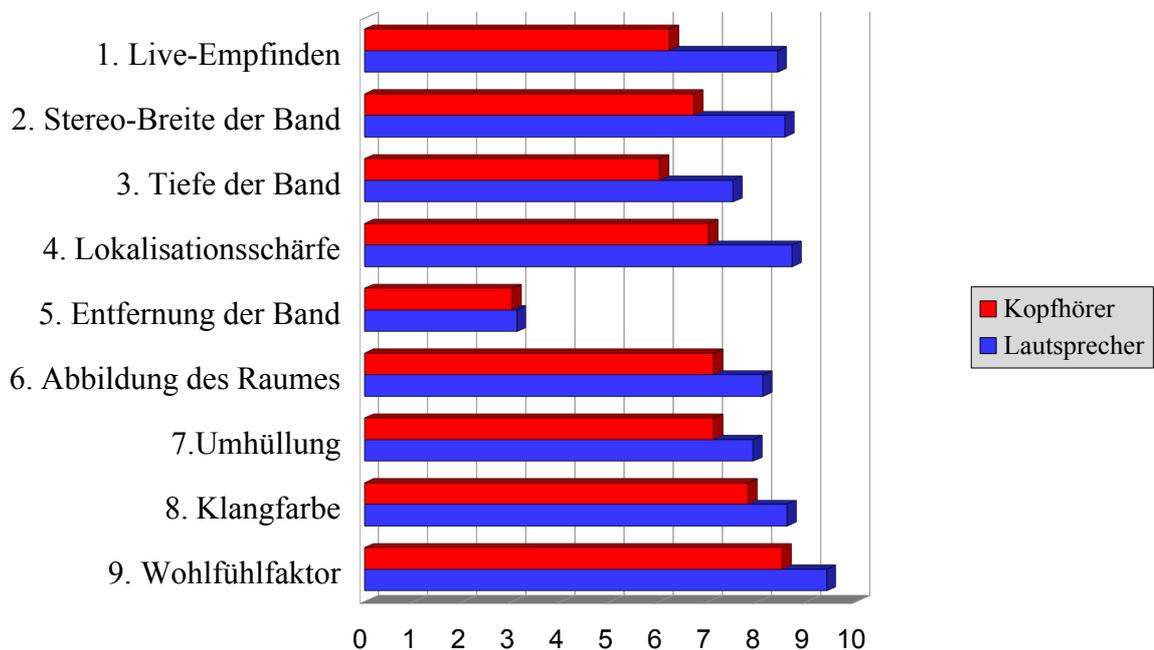


Abbildung 32: Durchschnittswerte der Testergebnisse

(eigene Darstellung)

## 6.10. Diskussion der Ergebnisse

Selbstverständlich kann der Hörversuch aufgrund der Anzahl der Probanden und Versuchsdurchläufe nicht als repräsentativ betrachtet werden, lässt aber die Tendenz erkennen, dass eine 3D-Wiedergabe über Kopfhörer nicht als deutlich schlechter empfunden wird. An diesem Punkt können durch weitere, vertiefende Untersuchungen und daraus abgeleitete Fragestellungen auch über diese Tendenz reichende Erkenntnisse erlangt werden.

Generell wird das echte Lautsprechersystem in allen Kategorien als „besser“ bewertet, was jedoch nicht überrascht, da die Wiedergabe über Kopfhörer lediglich eine Simulation darstellt, die darüber hinaus ein 3D-Format mit weniger Lautsprechern verwendet. Überraschend hingegen ist die Tatsache, dass die Simulation über Kopfhörer im Vergleich zum Lautsprechersystem nicht großartig schlechter abschneidet.

### 6.10.1. Detaillierte Ergebnisbetrachtungen

1. Live-Empfinden: Das Gefühl Live auf dem Konzert zu sein vermittelt die 3D-Kopfhörersimulation deutlich schlechter. Das ist sicherlich auch darauf zurückzuführen, dass die Kopfbewegungen der Probanden nichts am Klangbild ändern und sich die Probanden so nicht im Konzert „umschauen“ können. Wäre eine 3D-Kopfhörersimulation mit Head-Tracking System verwendet worden, wäre das Ergebnis wohl positiver ausgefallen.

Ein weiterer Nachteil des fehlenden Head-Tracking Systems ist, dass einige Probanden Vorn/Hinten-Vertauschungen empfinden, welche entstehen können wenn Schall gleichzeitig von vorne und hinten mit gleicher interauraler Laufzeitdifferenz an beiden Ohren auftrifft. Mit einem Abtasten der Peilbewegungen des Kopfes hätte dieser Effekt minimiert werden können. (vgl. Weinzierl, 2008, S. 589)

2. Stereo-Breite der Band: Auch in der wahrgenommenen Stereo-Breite ist die echte 3D-Anlage besser aufgestellt als der Kopfhörer. Die Kopfhörersimulation ist folglich nicht in der Lage, einen vergleichsweise großen Raum zu erschaffen.

3. Tiefe der Band: Obwohl das Panning keine Informationen über Laufzeiten nutzt, sondern die

Objekte je nach Anordnung im Tiefen-Panorama auf mehr oder weniger Lautsprecher verteilt, wird von den Probanden dennoch eine deutliche Tiefe auf beiden Wiedergabesystemen festgestellt. Vermutlich wurde diese Tiefe empfunden, da die Probanden nicht an eine derart räumliche, weil dreidimensionale Wiedergabe, gewöhnt sind und diese nicht von herkömmlichen Wiedergabesystemen kennen. Beim Kopfhörer wird die Tiefe allerdings von einigen Probanden nur seitlich als gut bewertet, während die vordere Abbildung als flach empfunden wird.

4. Lokalisationsschärfe: Das 3D-Lautsprechersystem weist eine deutliche Ausprägung von Lokalisationseigenschaften auf, wohingegen die 3D-Kopfhörersimulation Probleme in der genauen Darstellung der Klangquellen aufweist. Vor allem die Bläsersektion wird als schlecht lokalisierbar empfunden.

5. Entfernung der Band: Auf beiden Wiedergabesystemen wird die Band als relativ nah wahrgenommen. Wäre man auf dem Konzert, würde das in etwa der ersten Reihe entsprechen. Bei der Kopfhörerwiedergabe wird die Band oft dicht vor dem Kopf, auf Höhe der Stirn wahrgenommen.

6. Abbildung des Raumes: Offensichtlich hat der Kopfhörer gute Ansätze größere Räume abzubilden und das, obwohl dieser sehr dicht auf der Ohren sitzt. Dennoch schneidet das 3D-Lautsprechersystem knapp besser ab.

7. Umhüllung: Die 3D-Lautsprecheranlage schafft es besser den Hörer zu umhüllen. Bei der Kopfhörerwiedergabe funktioniert es schlechter den Hörer in allen Bereichen um den Kopf herum mit Schall zu umgeben. So wird beschrieben, dass es sich so anfühle als höre man mit großen dreidimensionalen Kegeln um die Ohren herum. Die Räumlichkeit vorne sei nicht optimal.

8. Klangfarbe: Beide Systeme geben eine relativ klangtreue Wiedergabe des Konzerts ab. Der Kopfhörer wird im Vergleich als etwas dumpfer empfunden.

9. Wohlfühlfaktor: Der Wohlfühlfaktor kann schon von Prinzip aus nur komplett subjektiv beurteilt werden. So spielt sicherlich auch die Gewöhnung an Kopfhörer eine große Rolle. Probanden, die es laut eigenen Aussagen generell vermeiden, zum Beispiel aus Komfortgründen, Musik über Kopfhörer zu hören, fällt es schwerer sich auf die binaurale Simulation einzulassen.

## 7. Fazit

Aus eigenen Erfahrungen und dem durchgeführten Hörversuch kann abgeleitet werden, dass eine authentischen Raumwiedergabe sowohl von einem echten 3D-Audio-System als auch von einer binauralen Kopfhörersimulation gewährleistet werden kann. Um jedoch in den vollen Genuss einer 3D-Wiedergabe zu kommen, ist die Installation eines auf Lautsprechern basierenden 3D-Systems zu empfehlen. Die hier getestete Kopfhörersimulation wird in den meisten Punkten als nicht deutlich schlechter empfunden, kann jedoch im Vergleich zum Lautsprecher-basierten 3D-System nie vollends überzeugen.

Die Reaktionen der Probanden, die am Hörversuch teilnahmen, waren in Bezug auf die 3D-Audio-Wiedergabe beider Systeme jedoch durchweg positiv, woraus abzuleiten ist, dass durchaus Bedarf an dieser Art der Wiedergabe vorhanden ist.

Eine Lösungsalternative könnte tatsächlich die 3D-Wiedergabe über Kopfhörer darstellen, welche vor allem im Gegensatz zur normalen Stereo-Wiedergabe überzeugt hat. Für die Produktion und Realisierung stehen dafür bereits viele Möglichkeiten zur Verfügung. Von einer Anschaffung eines echten 3D-Systems werden wohl viele Konsumenten aufgrund des deutlich höheren technischen und finanziellen Aufwands absehen.

Die Verbreitung professioneller Anwendungen von 3D-Audio, wie im Kino oder auf Events, wächst allerdings stetig, was darauf hoffen lässt, dass es in Zukunft eine weitere Entwicklung in Richtung 3D-Audio geben wird.

## 8. Literaturverzeichnis

### 8.1. Bücher und andere Veröffentlichungen

Ahnert, Wolfgang, Tennhardt, Hanns-Peter (2008): Raumakustik. In: Weinzierl, Stefan (Hrsg.) (2008): Handbuch der Audiotechnik. Berlin Heidelberg: Springer-Verlag.

Ahrens, Jens (2012): Analytic Methods of Sound Field Synthesis . Springer-Verlag Berlin Heidelberg.

Apple Inc. (2011): Logic Pro 9 – Benutzerhandbuch.

Audioborn (2015): Introducing Auratorium – World's first physics-based real-time 3D audio processor.

Dickreiter, Michael (Hrsg.) (2014a): Analoge Tonspeicherung. In: Dickreiter, Michael (Hrsg.) , Dittel, Volker (Hrsg.), Hoeg, Wolfgang (Hrsg.), Wöhr, Martin (Hrsg.) (2014): Handbuch der Tonstudioteknik. Band 1. 8. Auflage. Berlin: DE GRUYTER.

Dickreiter, Michael (Hrsg.) (2014b): Analoge Tonregieanlagen. In: Dickreiter, Michael (Hrsg.) , Dittel, Volker (Hrsg.), Hoeg, Wolfgang (Hrsg.), Wöhr, Martin (Hrsg.) (2014): Handbuch der Tonstudioteknik. Band 1. 8. Auflage. Berlin: DE GRUYTER.

Friesecke, Andreas (2014): Die Audio Enzyklopädie. Ein Nachschlagewerk für Tontechniker. 2. Auflage. Berlin: DE GRUYTER.

Hoeg, Wolfgang (2014): Beschallung. In: Dickreiter, Michael (Hrsg.) , Dittel, Volker (Hrsg.), Hoeg, Wolfgang (Hrsg.), Wöhr, Martin (Hrsg.) (2014): Handbuch der Tonstudioteknik. Band 1. 8. Auflage. Berlin: DE GRUYTER.

- Kugler, Bernd (2015): 3D-Audio im Ü-Wagen. In: vdt Magazin - 3D-Audio-Produktionen. Heft 3/2015.
- Lee, Hyunkook (2014): Psychoacoustic Considerations in Surround Sound with Height (Psychoakustische Betrachtungen zu Surround Sound mit Höhenabbildung). Tonmeistertagung – VDT International Convention 2014. Köln, Deutschland.
- Lee, Hyunkook (2015): Evaluation of the elevation effect for phantom images. 3rd International Conference on Spatial Audio (ICSA) 2015. Graz, Österreich.
- Lehmann, Katja, Langhammer, Jan (2013): Iosono 3D-Sound. In: Overschmidt, Gordian (Hrsg.), Schröder, Ute B. (Hrsg.) Fullspace Projektion - Mit dem 360°lab zum Holodeck. Berlin Heidelberg: Springer-Verlag.
- Maier, Peter (2008): Studioakustik. In: Weinzierl, Stefan (Hrsg.) (2008): Handbuch der Audiotechnik. Berlin Heidelberg: Springer-Verlag.
- Meyer-Schwarzenberger, Gernot (2014): Audio-Dateiformate. In: Dickreiter, Michael (Hrsg.) , Dittel, Volker (Hrsg.), Hoeg, Wolfgang (Hrsg.), Wöhr, Martin (Hrsg.) (2014): Handbuch der Tonstudioteknik. Band 1. 8. Auflage. Berlin: DE GRUYTER.
- Schätzlein, Frank (Hrsg.) (2005): Ton und Sounddesign beim Fernsehen. In: Schätzlein, Frank (Hrsg.), Segeberg, Harro Sound - Zur Technologie und Ästhetik des Akustischen in den Medien. 1. Auflage. Schüren Verlag GmbH.
- Schröder, Dirk (2016): Auratorium. Interview, geführt von Sascha Coronini am 30.01.2016.
- Schullan, Bernhard, Zuleeg, Ralf, Hoeg, Wolfgang (2014): Beschallung. In: Dickreiter, Michael (Hrsg.) , Dittel, Volker (Hrsg.), Hoeg, Wolfgang (Hrsg.), Wöhr, Martin (Hrsg.) (2014): Handbuch der Tonstudioteknik. Band 1. 8. Auflage. Berlin: DE GRUYTER.
- Slavik, Karl M. (2014): Film- und Fernsehton. In: Dickreiter, Michael (Hrsg.) , Dittel, Volker (Hrsg.), Hoeg, Wolfgang (Hrsg.), Wöhr, Martin (Hrsg.) (2014): Handbuch der Tonstudioteknik. Band 1. 8. Auflage. Berlin: DE GRUYTER.

- Slavik, Karl M., Weinzierl, Stefan (Hrsg.) (2008): Wiedergabeverfahren. In: Weinzierl, Stefan (Hrsg.) (2014): Handbuch der Audiotechnik. Berlin Heidelberg: Springer-Verlag.
- Spikofski, Gerhard (2014): Kunstkopfverfahren. In: Dickreiter, Michael (Hrsg.) , Dittel, Volker (Hrsg.), Hoeg, Wolfgang (Hrsg.), Wöhr, Martin (Hrsg.) (2014): Handbuch der Tonstudioteknik. Band 1. 8. Auflage. Berlin: DE GRUYTER.
- SPL electronics GmbH (2003): Atmos 5.1 Surround Miking System – Benutzerhandbuch.
- Stoll, Gerhard, Hartmann, Christian (2014): Tonaufnahme und Tonwiedergabe. In: Dickreiter, Michael (Hrsg.) , Dittel, Volker (Hrsg.), Hoeg, Wolfgang (Hrsg.), Wöhr, Martin (Hrsg.) (2014): Handbuch der Tonstudioteknik. Band 1. 8. Auflage. Berlin: DE GRUYTER.
- Theile, Günther (2014): Tonaufnahme und Tonwiedergabe. In: Dickreiter, Michael (Hrsg.) , Dittel, Volker (Hrsg.), Hoeg, Wolfgang (Hrsg.), Wöhr, Martin (Hrsg.) (2014): Handbuch der Tonstudioteknik. Band 1. 8. Auflage. Berlin: DE GRUYTER.
- Theile, Günther, Wittek, Helmut, (2012): 3D Audio Natural Recording. Tonmeistertagung 2012. Köln, Deutschland.
- Theile, Günther (2001): Multichannel Natural Music Recording - Based on Psychoacoustic Principles. 19. AES International Conference. Schloss Elmau, Deutschland.
- von Schultendorff, Gernot, Höpfner, Rainer, Batke, Jan-Mark (2015): Real-diffuse enveloping sound reproduction. ICSA 2015. Graz, Österreich.
- Weinzierl, Stefan (Hrsg.) (2008): Handbuch der Audiotechnik. Berlin Heidelberg: Springer-Verlag.

## 8.2. Internetquellen

Audioborn (2015b): Auratorium. Pfad <http://www.auratorium.de/> Stand 17.12.2015.

Auro Technologies (2015): Auro-3D Home Theater Setup - Installation Guidelines. Pfad [http://www.auro-3d.com/wp-content/uploads/documents/Auro-3D-Home-Theater-Setup-Guidelines\\_lores.pdf](http://www.auro-3d.com/wp-content/uploads/documents/Auro-3D-Home-Theater-Setup-Guidelines_lores.pdf) Stand 07.01.2016.

Auro Technologies (2016a): Auro-3D Creative Tool Suite. Pfad <http://www.auro-3d.com/professional/creative-tools/> Stand 10.01.2016

Auro Technologies (2016b): Auro Technologies - Auro-3D Authoring Tools - User Guide. Pfad <http://www.auro-3d.com/wp-content/uploads/documents/Auro-3D-Authoring-Tools-Documentation.pdf> Stand 10.01.2016

Auro Technologies, Barco (2015): AUROMAX - Next generation Immersive Sound System. Pfad [http://www.auro-3d.com/wp-content/uploads/documents/AuroMax\\_White\\_Paper\\_24112015.pdf](http://www.auro-3d.com/wp-content/uploads/documents/AuroMax_White_Paper_24112015.pdf) Stand 06.01.2016.

Auro Technologies (2011): Auro-3D Octopus Codec – Principles behind a revolutionary codec. Pfad <http://www.auro-3d.com/wp-content/uploads/documents/Auro3D-Octopus-White-Paper-v2-7-20111117.pdf> Stand 16.01.2016.

Avid Technology (2016): Pro Tools Software comparison table. Pfad <http://www.avid.com/DE/products/pro-tools-software-comparison-table> Stand 31.01.2016

Barco Audio Technologies (2015) Iosono for Events & Entertainment. Pfad <http://www.iosono-sound.com/events-entertainment/> Stand 22.12.2015.

Barco Audio Technologies (2015): Barco Audio Technologies - Great sound is our business and passion. Pfad [www.barco.com/~/\\_media/Downloads/Brochures/I-L/IOSONO/BARCO%20Audio%20Technologies%20-%20General%20pdf.pdf](http://www.barco.com/~/_media/Downloads/Brochures/I-L/IOSONO/BARCO%20Audio%20Technologies%20-%20General%20pdf.pdf) Stand 07.01.2016.

Bauhaus Dessau (2016): zdf@bauhaus – Konzertreihe auf der Bauhausbühne Dessau. Pfad <http://www.bauhaus-dessau.de/zdfbauhaus-16.html>) Stand 07.01.2016.

Denon (2016): DTS:X. Pfad <http://www.denon.com/Pages/DTSX.aspx> Stand 04.01.2016.

Dolby Laboratories (2014): Dolby Atmos Next-Generation Audio for Cinema - White Paper. Pfad <http://www.dolby.com/us/en/professional/cinema/products/dolby-atmos-next-generation-audio-for-cinema-white-paper.pdf> Stand 07.01.2016.

Dolby Laboratories (2015a): Dolby Atmos. Pfad <http://www.dolby.com/us/en/brands/dolby-atmos.html> Stand 16.12.2015.

Dolby Laboratories (2015b): Dolby Atmos Home Theater. Pfad <http://www.dolby.com/us/en/technologies/home/dolby-atmos.html> Stand 16.12.2015.

Dolby Laboratories (2015c): Dolby Atmos on the Go. Pfad <http://www.dolby.com/us/en/technologies/mobile/dolby-atmos.html> Stand 16.12.2015.

Dolby Laboratories (2015d): Dolby Atmos Cinema Sound. <http://www.dolby.com/us/en/technologies/cinema/dolby-atmos.html> Stand 16.12.2015.

DTS (2013): DTS Technology for Home Theater. Pfad [http://www.dts.com/~media/eff9b713e9db4e02b615026dbed50e30/DTS\\_Home\\_Theater\\_NeoX\\_infosheet.pdf](http://www.dts.com/~media/eff9b713e9db4e02b615026dbed50e30/DTS_Home_Theater_NeoX_infosheet.pdf) Stand 07.01.2016.

DTS (2016a): Neo:X Receivers: Pfad <http://www.dts.com/consumers/entertainment-audio/neox.aspx#product> Stand 25.01.2016

DTS (2016b): DTS Headphone: X. Pfad <http://www.dts.com/professionals/sound-technologies/headphonex.aspx#features> Stand 04.01.2016.

Eckel, Gerhard, Rumori, Martin (2014): StiffNeck: The Electroacoustic Music Performance Venue in a Box,. Pfad [http://iem.at/~eckel/download/CoS\\_icmc2014.pdf](http://iem.at/~eckel/download/CoS_icmc2014.pdf) Stand 07.01.2016.

Elen, Richard (2001): Ambisonics - The Surround Alternative. Pfad  
<http://www.ambisonic.net/pdf/ambidvd2001.pdf> Stand 07.01.2016.

Jack Audio (2016): Jack Audio Connection Kit. Pfad <http://www.jackaudio.org/> Stand  
26.01.2016

Kompendium, Infotip Service GmbH (2016b): Positionen der Ebenen bei NHK 22.2. Pfad  
[http://kompendium.infotip.de/files/wdb/GRAFIK/2100\\_AUDIO/2130\\_3D-Audio/ABB\\_2130\\_04\\_06\\_Pos\\_seitl.gif](http://kompendium.infotip.de/files/wdb/GRAFIK/2100_AUDIO/2130_3D-Audio/ABB_2130_04_06_Pos_seitl.gif) Stand 11.01.2016.

Kompendium, Infotip Service GmbH (2016a): Surround Sound und 3D-Audio. Pfad  
<http://kompendium.infotip.de/3d-audio.html> Stand 05.01.2016.

Neumann (2016): Neumann KU100. Pfad  
[http://www.neumann.com/img/photosGraphics/Zooms/KU100\\_Z.jpg](http://www.neumann.com/img/photosGraphics/Zooms/KU100_Z.jpg) Stand 07.01.2016.

New Audio Technology (2016a): Spatial Audio Designer – Mix-Modul. Pfad  
<http://newaudiotechnology.com/wp-content/uploads/2013/05/SAD-2015.jpg> Stand  
07.01.2016.

New Audio Technology, (2016b): Spatial Audio Designer. Pfad  
<http://newaudiotechnology.com/de/produkte/spatial-audio-designer/> Stand 07.01.2016.

Sonic Emotion (2016): Sonic Emotion Absolut 3D. Pfad <http://www2.sonicemotion.com/> Stand  
05.01.2016.

Sonic Emotion, 2013: Sonic Wave I The pro listening experience. Pfad  
[http://www2.sonicemotion.com/wp-content/uploads/2014/02/BROCHURE\\_SonicWaveI\\_5\\_November\\_2013\\_FINAL\\_Lowres.pdf](http://www2.sonicemotion.com/wp-content/uploads/2014/02/BROCHURE_SonicWaveI_5_November_2013_FINAL_Lowres.pdf). Stand 07.01.2016.

SPL electronics GmbH (2016): Atmos 5.1 Controller & Brauner ASM 5. Pfad  
[http://spl.info/fileadmin/user\\_upload/produkte-w/atmos5.1/Karussel/Atmos\\_Hauptbild.jpg](http://spl.info/fileadmin/user_upload/produkte-w/atmos5.1/Karussel/Atmos_Hauptbild.jpg)  
Stand 07.01.2016.

Van Daele , Bert, Van Baelen, Wilfried (2012): Productions in Auro 3D - Professional workflow and costs. Pfad <http://www.auro-3d.com/wp-content/uploads/documents/Professional-Workflow-White-Paper-v0-6-20120228.pdf> Stand 07.01.2016.

Wittek, Helmut (2004): Räumliche Wahrnehmung von virtuellen Quellen bei Wellenfeldsynthese. Pfad [http://www.hauptmikrofon.de/HW/Wittek\\_TMT2004\\_Paper\\_final.pdf](http://www.hauptmikrofon.de/HW/Wittek_TMT2004_Paper_final.pdf) Stand 06.01.2016.

# 9. Anhang

## 9.1. Anhangsverzeichnis

Fragebogen.....	ii
Iosono SAW: Alle Audio-Objekte.....	iii
Iosono SAW: Band.....	iii
Iosono SAW: Atmo.....	iii
Iosono SAW: Schlagzeug.....	iii
Iosono SAW: Effekte.....	iv
Iosono SAW: Gesang.....	iv
Nuendo SAW-Session.....	iv
New Audio Technology SAD: Alles.....	v
New Audio Technology SAD: Gesang.....	v
New Audio Technology SAD: 9.1.....	v
New Audio Technology SAD: Band.....	v
New Audio Technology SAD: Atmo.....	vi
New Audio Technology SAD: Effekte.....	vi
Pro Tools SAD-Session: Routing.....	vi
CD-Rom.....	vi

# Fragebogen

## **3D-Audio: Kopfhörer-/ Lautsprecherwiedergabe**

**1. Live-Empfinden:** Ich befinde mich mitten unter den Zuschauern.  
Von 1 (trifft überhaupt nicht zu) bis 10 (trifft vollkommen zu).

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

**2. Stereo-Breite der Band:** Distanz von links nach rechts.  
Von 1 (sehr schmal) bis 10 (sehr breit).

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

**3. Tiefe der Band:** Entfernung zwischen vorderen und hinteren Musikern.  
Von 1 (sehr flach) bis 10 (sehr tief).

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

**4. Lokalisationsschärfe:** Die Positionen der einzelnen Musiker sind deutlich zu unterscheiden.  
Von 1 (nicht lokalisierbar) bis 10 (sehr gut lokalisierbar).

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

**5. Entfernung der Band:** Abstand zwischen Hörer und Band.  
Von 1 (sehr nah) bis 10 (weit entfernt).

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

**6. Abbildung des Raumes:** Der Konzertraum wird realistisch abgebildet.  
Von 1 (trifft überhaupt nicht zu) bis 10 (trifft vollkommen zu).

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

**7. Umhüllung:** Ich fühle mich vollkommen vom Schall umgeben.  
Von 1 (trifft überhaupt nicht zu) bis 10 (trifft vollkommen zu).

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

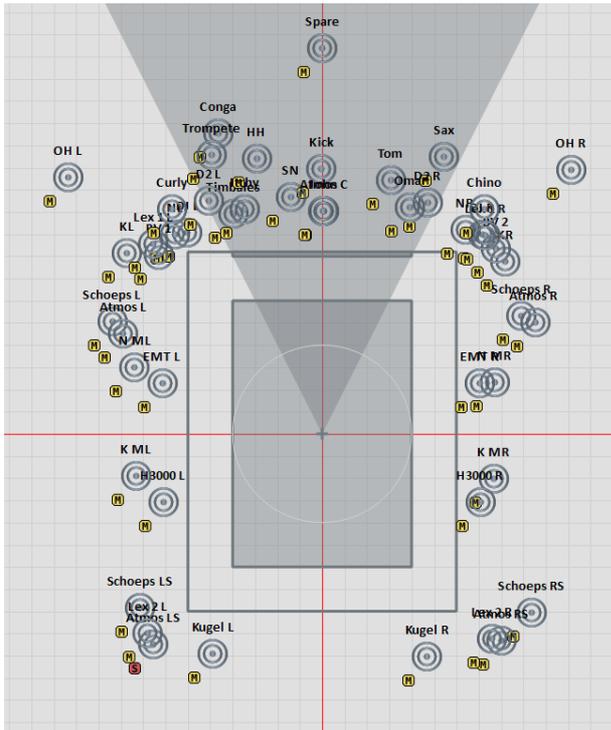
**8. Klangfarbe:** Das Klangspektrum des Konzerts wird exakt abgebildet.  
Von 1 (verfälschte Abbildung) bis 10 (exakte Abbildung).

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

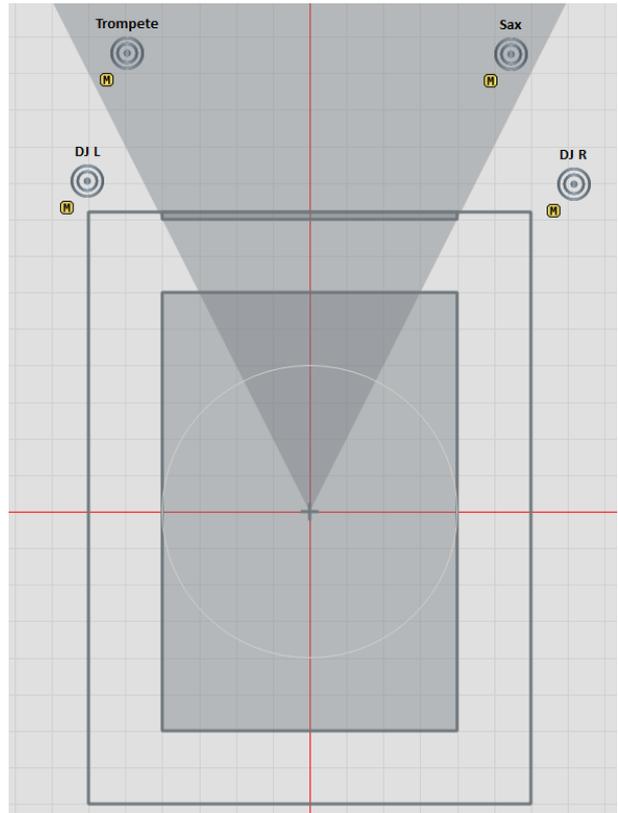
**9. Wohlfühlfaktor:** Von 1 (ich fühle mich überhaupt nicht wohl) bis 10 (Ich fühle mich sehr wohl).

1	2	3	4	5	6	7	8	9	10
---	---	---	---	---	---	---	---	---	----

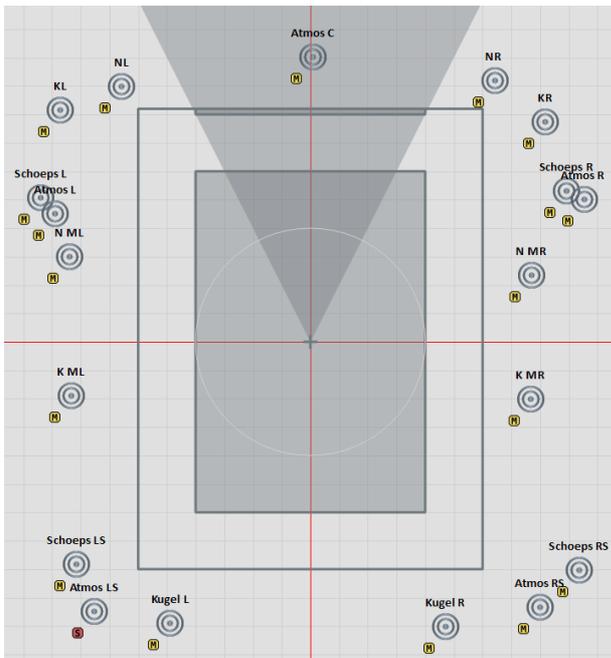
Iosono SAW: Alle Audio-Objekte



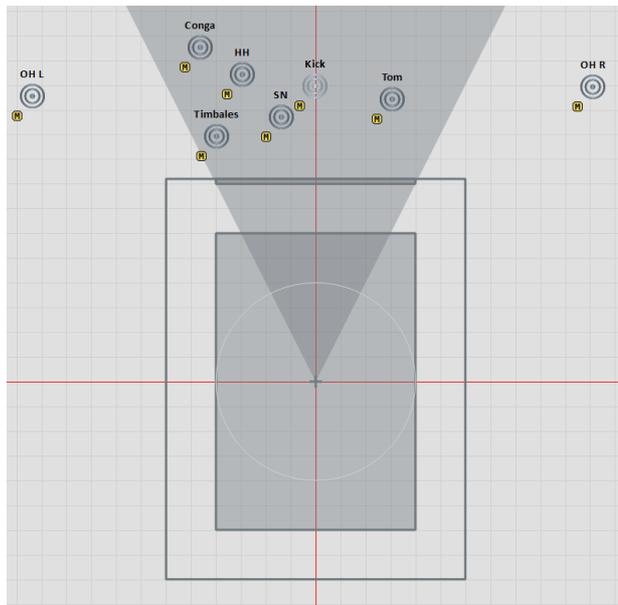
Iosono SAW: Band



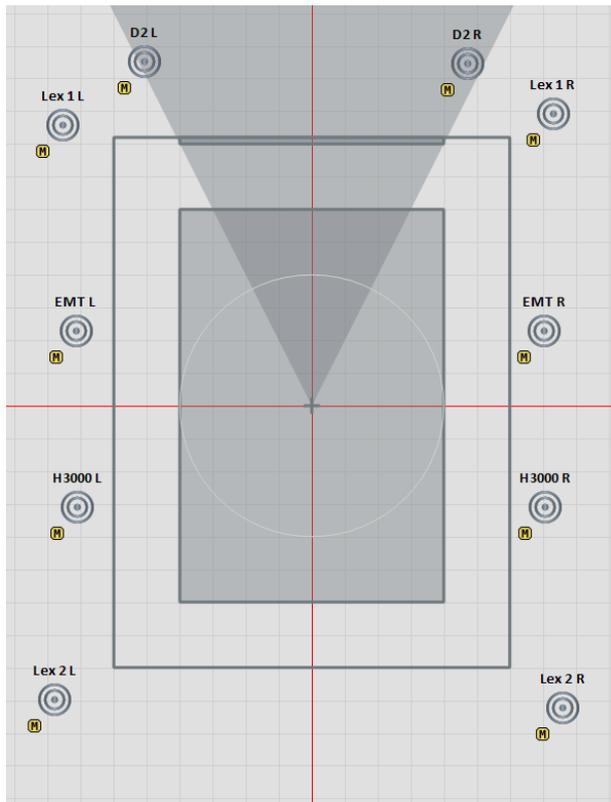
Iosono SAW: Atmo



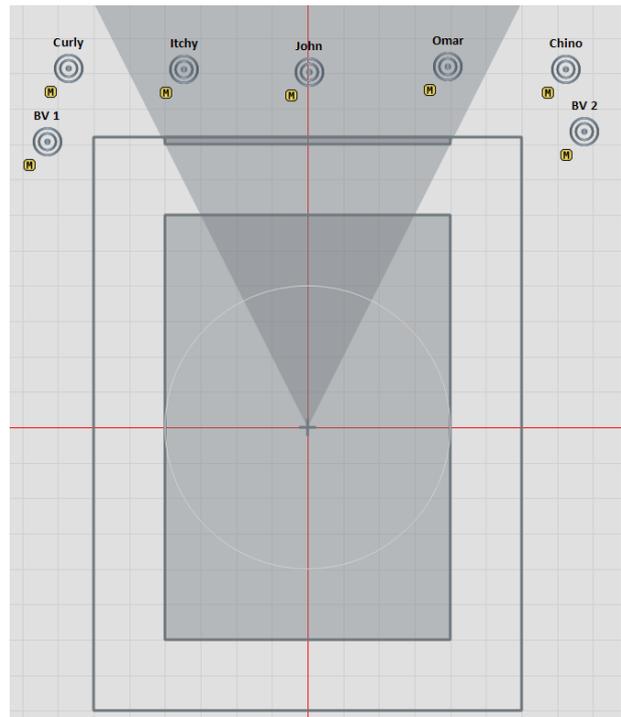
Iosono SAW: Schlagzeug



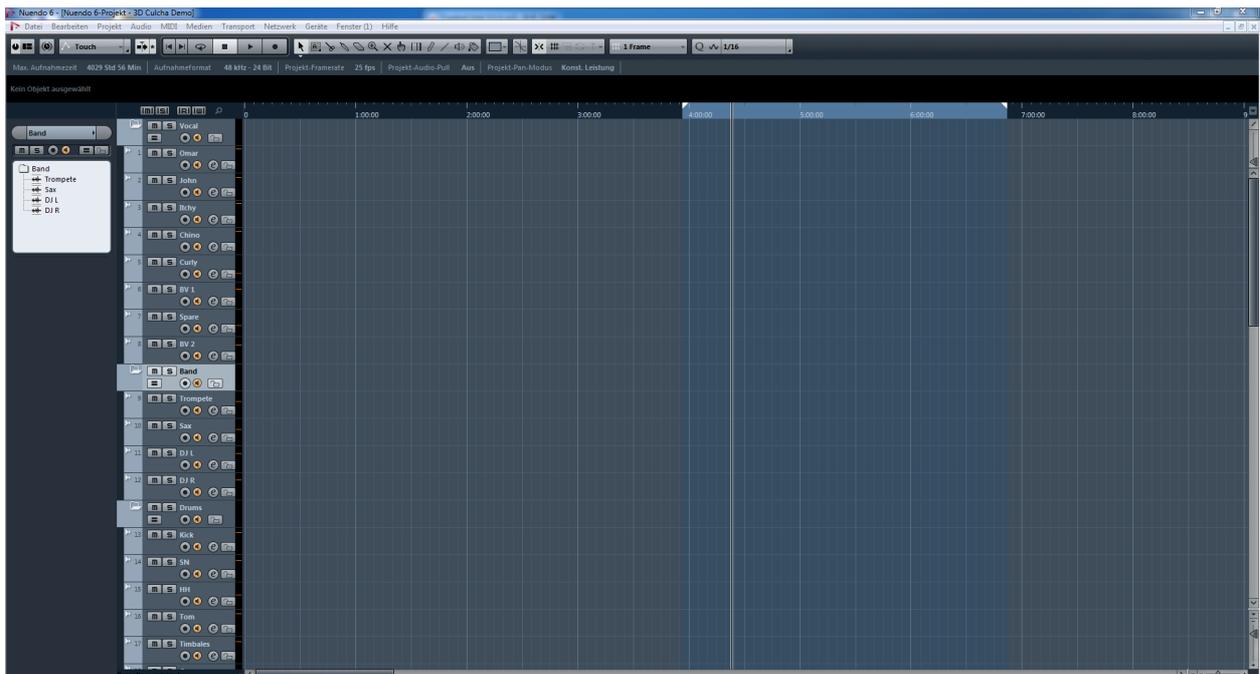
## Iosono SAW: Effekte



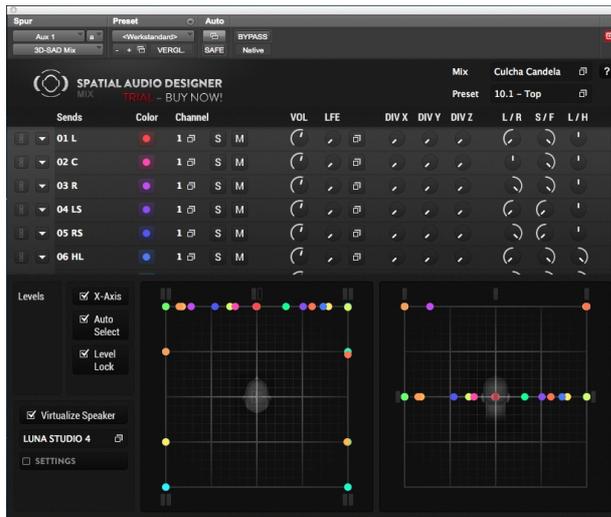
## Iosono SAW: Gesang



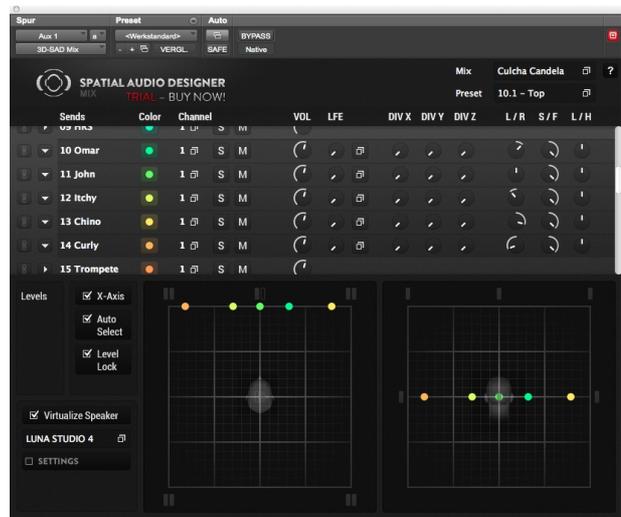
## Nuendo SAW-Session



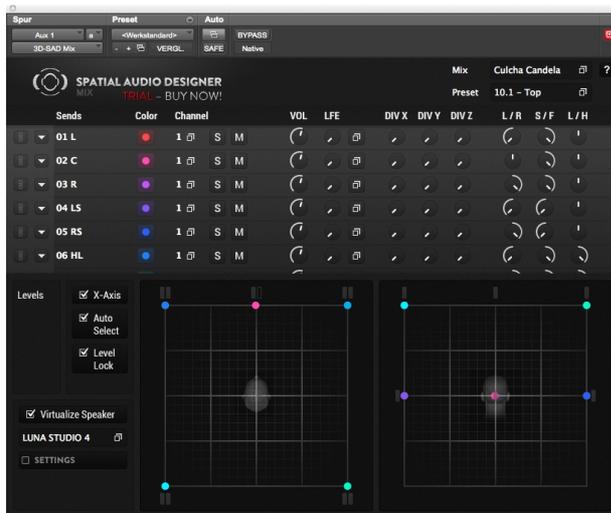
## New Audio Technology SAD: Alles



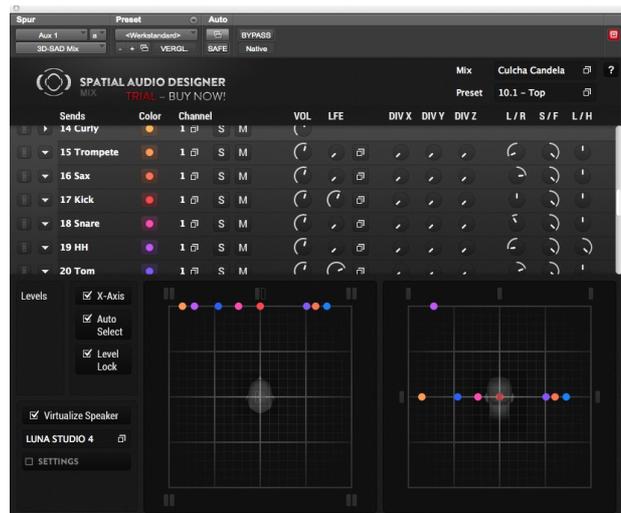
## New Audio Technology SAD: Gesang



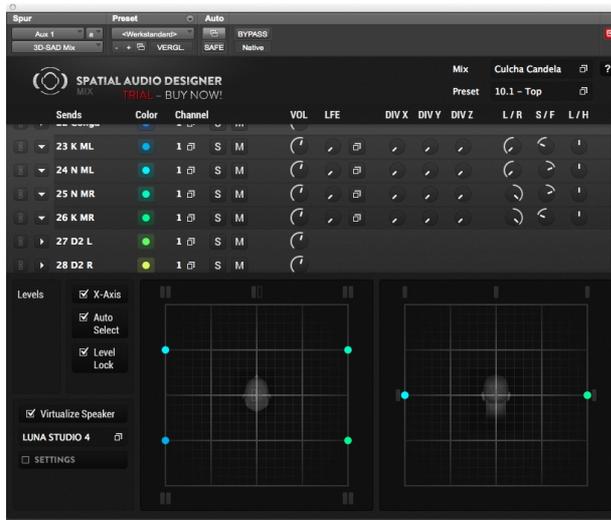
## New Audio Technology SAD: 9.1



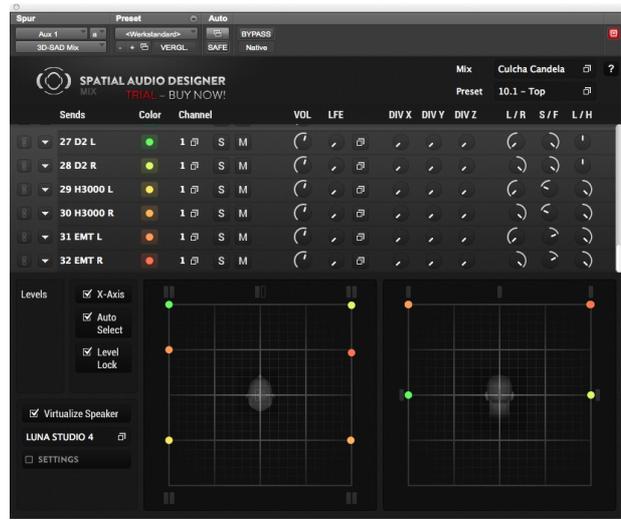
## New Audio Technology SAD: Band



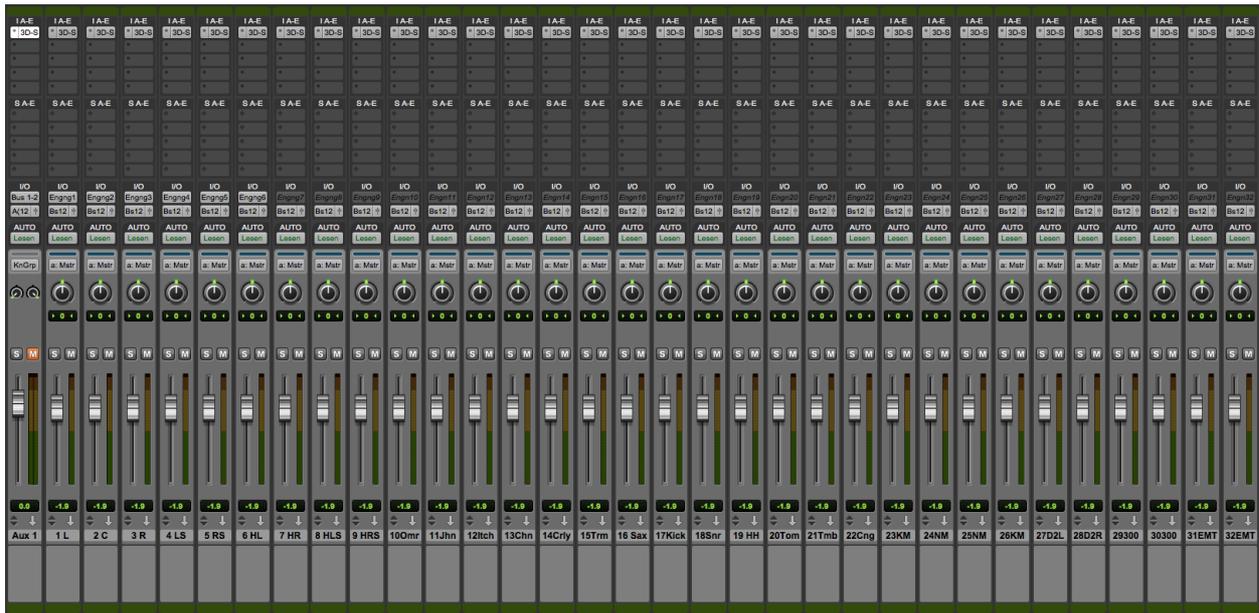
## New Audio Technology SAD: Atmo



## New Audio Technology SAD: Effekte



## Pro Tools SAD-Session: Routing



## CD-Rom

Die beiliegende CD-Rom beinhaltet neben den Grafiken des Anhangs auch die binaurale Simulation des Musikstücks im WAV-Format. Weiter befindet sich ein Einblick ins Tonstudio sowie die gesamte Arbeit im PDF-Format auf der CD-Rom.