Hochschule der Medien - Stuttgart
Fakultät 2 - Electronic Media

Bachelorarbeit im Studiengang Audiovisuelle Medien

zur Erlangung des akademischen Grades: Bachelor of Engineering

Aspekte der objektbasierten Rundfunkproduktion

am Beispiel des EU-Forschungsprojekts ORPHEUS im Bayerischen Rundfunk

vorgelegt von Miriam Böhm

Matrikel-Nr.: 28897

am 11. August 2017

Erstprüfer: Prof. Oliver Curdt

Zweitprüfer: M.A. Werner Bleisteiner

Eidesstattliche Erklärung

„Hiermit versichere ich, Miriam Böhm, ehrenwörtlich, dass ich die vorliegende Bachelorarbeit mit dem Titel:

„Aspekte der objektbasierten Rundfunkproduktion am Beispiel des EU-Forschungsprojekts ORPHEUS im Bayerischen Rundfunk“

selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.“

(Ort, Datum) (Miriam Böhm)

Kurzfassung

In den letzten Jahren hat sich das Medienkonsumverhalten durch die Verwendung mobiler Endgeräte sehr verändert. Die Art und Weise, wie Medien produziert werden, ist jedoch gleich geblieben. Rundfunkanbieter versuchen die verschiedenen Plattformen, Wiedergabesysteme und -situationen zu bedienen und nicht zuletzt auch die persönlichen Vorlieben der Konsumenten. Dadurch entsteht eine Vielzahl an Formaten. Das EU-Forschungsprojekt *Orpheus* befasst sich damit, Lösungen für diese Problemstellung zu finden. Der dabei verfolgte Ansatz arbeitet „objektbasiert“ statt wie bisher üblich kanalbasiert. Was darunter zu verstehen ist, welche neuen Möglichkeiten sich dadurch insbesondere für das Medium Hörfunk bieten, welchen Mehrwert der Hörer davon hat und wie dies technisch umgesetzt wird, beschreibt diese Arbeit. Sie stellt *Orpheus* vor und geht dabei auf eine objektbasierte Produktion im Bayerischen Rundfunk ein.

Abstract

In the last few years, the media consumption behaviour has changed considerably due to the use of mobile devices. However, the way media is produced has remained the same. Broadcasters try to cope with the various platforms, playback systems and -situations, and try to serve the personal preferences of the consumers. This creates a large number of formats. The EU research project *Orpheus* is concerned with finding solutions to this problem. It pursues an ‘object-based’ approach, instead of channel-based as usual. This thesis describes the basic principles of ‘object-based’ production, the new possibilities it offers to radio broadcasting, the potential benefits for the listeners and how the technical implementation can be achieved. It also exemplifies an object-based pilot production at Bayerischer Rundfunk.

Inhaltsverzeichnis

Eidesstattliche Erklärung 2

Kurzfassung 3

Abstract 3

Inhaltsverzeichnis 4

Abkürzungsverzeichnis 6

1 Die Notwendigkeit für Veränderungen 7

2 Von kanalbasiert zu objektbasiert 9

3 Orpheus 14

3.1 Projektvorstellung . . . 14

3.1.1 Motivation . . . 15

3.1.2 Ziele . . . 16

3.1.3 Neue Möglichkeiten . . . 16

3.1.4 Anwendungsbeispiel . . . 18

3.2 Projektstruktur . . . 19

3.3 Projektpartner . . . 22

4 Pilot- und Referenzarchitektur 26

5 Speicherung und Archivierung 29

5.1 Broadcast Wave BW64 . . . 29

5.2 Audio Definition Model (ADM) . . . 30

5.3 Archivierung . . . 31

6 Produktion **32**
6.1 Aufnahme . . . 33
6.1.1 Mikrofone . . . 34
6.1.2 Metadaten . . . 35
6.2 Bearbeitung . . . 36
6.3 Hall . . . 37
6.4 Monitoring . . . 38

7 Beispielproduktion - Die Kunst des Geräuschemachens **39**
7.1 Produktionsumgebung und -tools . . . 39
7.2 Stereoproduktion . . . 42
7.3 Objektbasierte Produktion . . . 43
7.4 Fazit . . . 44

8 Sendung **45**

9 Pilotphase 1 **48**
9.1 Phase 1A - Live . . . 48
9.2 Phase 1B - Wie live . . . 50
9.3 Phase 1C - Live encodiertes MPEG-H . . . 51

10 Distribution **53**
10.1 MPEG-DASH . . . 53
10.2 AAC & ADM . . . 54
10.3 MPEG-H . . . 54

11 Wiedergabe **56**

12 Fazit **58**

Danksagung **61**

Abbildungsverzeichnis **62**

Quellenverzeichnis **63**

Anhang **66**

Abkürzungsverzeichnis

AAC	Advanced Audio Coding
AAX	Avid Audio eXtension
ADM	Audio Definition Model
AES	Audio Engineering Society
API	Application Programming Interface
AV	Audio Video
BWF	Broadcast Wave Format
BW64	Broadcast Wave 64 bit
DASH	Dynamic Adaptive Streaming over HTTP
DAW	Digital Audio Workstation
EBU	European Broadcasting Union
EU	Europäische Union
FOA	First Order Ambisonics
HOA	Higher Order Ambisonics
HTML	Hypertext Markup Language
IEC	Internationale elektrotechnische Kommission
IP	Internet Protocol
ISO	Internationale Organisation für Normung
ITU-R	International Telecommunication Union, Radiocommunication Sector
MADI	Multichannel Audio Digital Interface
MPEG	Moving Picture Experts Group
PCM	Puls Code Modulation
TCP	Transmission Control Protocol
VST	Virtual Studio Technology
WP	Work Package
XML	Extensible Markup Language

1 Die Notwendigkeit für Veränderungen

Die Vielfältigkeit der Art und Weise, wie Medien genutzt werden können, hat in den vergangenen Jahren stark zugenommen. Das klassische Radio- oder Fernsehgerät ist längst nicht mehr das einzige Gerät, mit dem Rundfunkinhalte empfangen werden können. Per Internetverbindung ist es möglich, auch mobile Geräte wie Smartphones, Tablets und Laptops zum Konsumieren von Hörfunk und Fernsehen zu nutzen. Mobiles Breitbandinternet und Wifi Hotspots sorgen dafür, dass sich auch die Nutzungsumgebung verändert hat. Es wird nicht mehr nur in den privaten vier Wänden Radio gehört und ferngesehen, sondern auch unterwegs in Bus und Bahn, der Fußgängerzone, auf dem Weg zur oder von der Arbeit. Das Nutzungsumfeld variiert stark. Je nach Ort entscheidet sich auch, wie gehört wird - ob über Lautsprecher oder Kopfhörer. Durch das Internet wurde es auch möglich, dass Rundfunkangebote nicht mehr linear und live zum Zeitpunkt der Ausstrahlung konsumiert werden müssen. Beispielsweise als Podcast können sie onDemand zu jedem beliebigem Zeitpunkt angehört und -gesehen werden. Es kann pausiert und vor- und zurückgesprungen werden. Smartphones und Tablets bieten durch die Installation von Apps viele neue Interaktions- und Individualisierungsmöglichkeiten, die das einfache Radio- und Fernsehgerät nicht haben.

Dieser Fülle an Veränderungen haben sich Rundfunkanstalten bisher noch wenig angepasst und neue sich bietende Möglichkeiten werden nur spärlich genutzt. Mit der Umstellung von Analog- zu Digitaltechnik hat sich die Art und Weise, wie Radio und Fernsehen produziert werden, in den vergangenen Jahrzehnten dennoch nicht grundlegend geändert.

Eine weitere in den letzten Jahren zu beobachtende Veränderung ist der Trend zur Vernetzung von Produktionssystemen. IP-Netzwerke sind mittlerweile überall vorhanden und sind universeller und günstiger als herkömmliche Einzelverbindungen. Auch Audio und Video können so verschickt werden. Das führt auch zu Veränderungen in

der Produktionsinfrastruktur von Rundfunkhäusern. Derzeit gibt es eine große Vielzahl an Schnittstellen (z.B. MADI, AES/EBU, XLR, Koaxial, Siemens-Stecker, BNC, Cinch, usw.), die den Austausch und die Kompatibilität untereinander oft erschweren. Es ist jedoch absehbar, dass es in Zukunft nur noch eine einzige Art Stecker geben wird, den Netzwerkstecker RJ45 mit Cat-Kabel, über den die verschiedenen Signale in unterschiedlichen Protokollen geschickt werden.

Um neue Möglichkeiten und Technologien für den Rundfunk zu entwickeln, auszuprobieren und neue Standards zu definieren, haben sich 10 europäische Rundfunkanstalten, Forschungsinstitute und Firmen zu einer Arbeitsgemeinschaft zusammengeschlossen. Durch die EU im Rahmen von *Horizon 2020* unterstützt, starteten sie das Forschungsprojekt *„Orpheus - Audioerlebnis der nächsten Generation“*. Das Projekt dauert 30 Monate und endet im Mai 2018.

Diese Arbeit stellt *Orpheus* bis zum jetzigen Projektstand vor. Sie zeigt die Herangehensweise an die neuen Technologien und die Ideen, wie diese umgesetzt werden können. Der Fokus liegt dabei hauptsächlich auf dem Medium Radio, da hierauf das Hauptaugenmerk des Projekts liegt. Fernsehen soll gegebenenfalls nur am Rande an geeigneter Stelle erwähnt werden.

In Kapitel 2 wird zunächst der dem Projekt zugrunde liegende neue Denkansatz erläutert.

Kapitel 3 gibt anschließend einen ersten Überblick über das Projekt, dessen Struktur und stellt die Projektpartner vor. Der Kern von Orpheus, die Erstellung einer Referenzarchitektur, wird in Kapitel 4 erklärt. Kapitel 5 bis 11 gehen auf die einzelnen Glieder einer Radioproduktionskette ein - Produktion, Sendung, Speicherung, Distribution und Wiedergabe - und zeigen, welche neuen Technologien und Arbeitsabläufe angedacht sind. Ein Beispiel, wie nach aktuellem Stand der Entwicklungen eine Radioproduktion ablaufen kann, wird dabei in Kapitel 7 anhand einer Featureproduktion des Bayerischen Rundfunks vom April 2017 aufgezeigt. In Kapitel 9 wird eine Live-Sendung beschrieben, die Anfang Juli 2017 im Zuge der Pilotphase 1 bei der BBC in London stattfand.

Abschließend werden in Kapitel 12 die bisherigen Ergebnisse des Projekts zusammengefasst und ein Ausblick auf die noch zu leistende Arbeit gegeben. Zudem wird ein Fazit gezogen, das Vor- und Nachteile aufzeigt.

2 Von kanalbasiert zu objektbasiert

Seit es Ende des 19. Jahrhunderts möglich wurde, Schall zu speichern und zu einem späteren Zeitpunkt wiederzugeben, ist die Tonbearbeitung kanalbasiert. Kanal ist hier gleichbedeutend mit Lautsprecher. Jeder Lautsprecher bekommt ein eigenes, fertig gemischtes Signal. Anfangs gab es nur einen Lautsprecher, aus dem im Laufe der Jahre zwei wurden - aus Mono wurde Stereo. Der Musikproduktion und dem Rundfunk reichte dies lange Zeit völlig aus.

Im Kino jedoch wuchs die Zahl der Lautsprecher weiter an. Zu den beiden Stereolautsprechern für Links und Rechts, kam ein Center-Lautsprecher hinzu, um im großen Kinosaal für eine bessere Ortung von Schallquellen, insbesondere von Sprache, in der Mitte der Leinwand zu sorgen. Bald darauf folgten Surround-Lautsprecher, die seitlich und hinter dem Publikum platziert wurden. Diese waren ursprünglich nur dafür gedacht, einen realistischen Raumklang durch ein diffuses Schallfeld zu vermitteln. Hierbei wird ein sog. „Array“, eine Anordnung aus mehreren (Surround-)Lautsprechern, verwendet. Die Lautsprecher der linken Hälfte bekommen das Links-Surround-Signal und die der rechten Hälfte das Rechts-Surround-Signal. Es kam noch ein weiterer Lautsprecher für die Wiedergabe tiefer Frequenzen hinzu, der Subwoofer, Tieftöner oder auch LFE (Low Frequency Effects) genannt wird. Schließlich etablierte sich diese Lautsprecheranordnung unter dem Namen *5.1 Format*:

- 3 Lautsprecher vor dem Publikum hinter der Leinwand (Links, Center, Rechts)
- 2 Surroundkanäle, bestehend aus mehreren Lautsprechern, seitlich und hinter dem Publikum (Links Surround, Rechts Surround)
- 1 Tieftöner (LFE)

Abbildung 2.1 veranschaulicht diese Anordnung der Lautsprecher im Kinosaal.

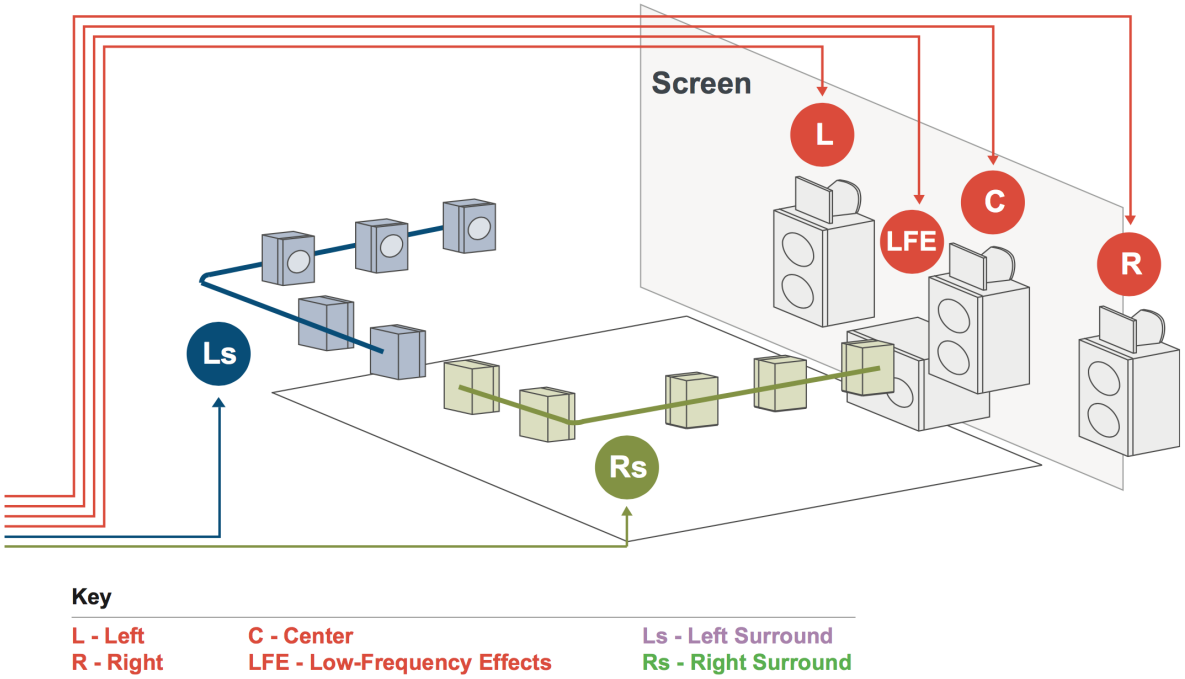

Abb. 2.1: Das 5.1 Format im Kino

Im Zuge der Digitalisierung entstand 2010 noch ein weiteres Format, das 7.1 Format. Hierbei werden die beiden Surroundkanäle jeweils in zwei weitere Kanäle unterteilt - in Seite und Hinten.

Bald begnügte man sich aber nicht mehr damit, nur Raumklang auf den Surroundkanälen wiederzugeben. Schallereignisse, die außerhalb der Leinwand stattfinden, sollten auch von dort zu hören sein. Mit je nur einem, bzw. zwei Kanälen für Links Surround und Rechts Surround war dies aber nicht zufriedenstellend umsetzbar. Es entstanden weitere Formate.

Einen in der Audiowelt völlig neuen Denkansatz stellten im April 2012 die *Dolby Laboratories* vor, als sie „Dolby Atmos“ präsentierten. Ziel ist es, das Mischen, die Distribution und die Wiedergabe von Filmen zu verbessern und ein neues Kinoerlebnis zu schaffen. Es soll erreicht werden, dass es keinen sog. „Sweetspot“ im Kino mehr gibt, einen Sitzplatz, an dem der Klang optimal ist und alle Schallquellen an der richtigen Stelle geortet werden können. Mit Dolby Atmos soll es möglich sein, alle Schallquellen von jedem Sitzplatz aus gleich gut orten zu können.

Erreicht wird das durch eine objektbasierte Mischung. Audiosignale werden hierbei als Objekte betrachtet und dynamisch auf individuelle Lautsprecher verteilt. In der Filmton-

mischung werden Töne, Klänge, Schallereignisse nun nicht mehr wie bisher auf einzelne Kanäle und somit konkrete Lautsprecher gemischt, sondern diese Objekte werden mit Metadaten versehen. Die Objekte können mit einem Panner Plugin in der Digitalen Audio Workstation (DAW) *ProTools* an einer beliebigen Position auf einer gedachten Halbkugel über der Abhörposition platziert werden. Die Metadaten enthalten unter anderem die Informationen zur Lautstärke und zur Position auf dieser Halbkugel. Erst bei der Wiedergabe berechnet ein Renderer, eine Recheneinheit, anhand der Metadaten, auf welche Lautsprecher das Objekt mit welchem Pegel geschickt werden muss, damit es an der gewünschten Position zu hören ist. Hierfür stehen alle sich im Kinosaal befindlichen Lautsprecher zur Verfügung, deren exakte Position dem Renderer vorher bekannt sein muss. Durch das Hinzufügen von Lautsprechern an der Decke soll ein realistisches Klangfeld geschaffen werden, das den Kinobesucher[1] vollständig umhüllt. Dolby Atmos unterstützt bis zu 64 einzelne Lautsprecher, auf die maximal 128 Spuren, sog. „Tracks“, verteilt werden können.

Ein veranschaulichendes Beispiel für ein Klangobjekt ist das Geräusch eines Hubschraubers. Der Mischtonmeister möchte, dass dieser Hubschrauber rechts oben zu hören ist. Somit platziert er das Audiosignal des Hubschraubers dort an der gewünschten Position. Dem Renderer im Kinosaal ist anschließend bekannt, wo sich welche Lautsprecher befinden. Er berechnet die Signale für die in Frage kommenden Lautsprecher so, dass der Hubschrauber genau an der gewünschten Position rechts oben zu hören ist.

Weil es aber sinnvoll ist, bestimmte Signale weiterhin über das herkömmliche Lautsprecherarray wiederzugeben, ist Dolby Atmos eine Mischung aus kanalbasierter und objektbasierter Arbeitsweise. Raumklang (Hall), Atmosphären und Musik können als sogenannte „Beds“ weiterhin kanalbasiert über 5.1 (7.1, 9.1, usw.) wiedergegeben werden und nur einzelne Objekte, wie etwa der zuvor beschriebene Hubschrauber, werden objektbasiert auf bestimmte Positionen im Raum verteilt. Somit sind 10 Tracks (bei 9.1) für die Beds reserviert, die über die bisherigen Formate (5.1, 7.1, 9.1, usw.) ausgegeben werden. Auf diese Weise wird auch die Abwärtskompatibilität für bisherige Soundsysteme in den Kinos gewährleistet, die nicht auf Dolby Atmos umgerüstet sind. Die übrigen 118 Tracks stehen dann für Objekte zur Verfügung.[2]

[1] Anmerkung: der Leserlichkeit halber wird in dieser gesamten Arbeit auf die Gleichstellung der Geschlechter verzichtet und nur die männliche Form verwendet

[2] vgl. Dolby Laboratories (2014) S. 2-7

Um Dolby Atmos nutzen und dekodieren zu können, ist die Rendering & Mastering Unit (RMU) von Dolby erforderlich. Diese basiert auf dem Dolby AC-4 Codec, bei dem es sich um ein proprietäres Format von Dolby handelt.

Dolby bietet neben Kino auch Geräte für den Rundfunkbereich an, wie etwa für Liveübertragungen von Fußballspielen. Doch auch hier ist wieder die Dolby-eigene Hardware notwendig, sowohl für die Signalwege vor Ort im Übertragungswagen, als auch für den Empfang beim Endnutzer.

Aus diesem Grund zögern europäische Rundfunkanstalten die Systeme von Dolby zu nutzen, da man hier an einen Hersteller gebunden und somit abhängig von ihm wäre. Frühere Erfahrungen haben gezeigt, dass es in der Produktion durch die Verwendung von herstellergebundenen Formaten zu Problemen kommen kann, wenn sich in der Produktiosstruktur etwas ändert oder weiterentwickelt. Für die Distribution zum Konsumenten kann jedoch durchaus auf solche Formate zurückgegriffen werden.

Im Mai 2017 veranstaltete die europäische Rundfunkunion EBU einen Workshop zum Thema „Object-Based Audio“ in Genf. Den Teilnehmern wurde hier unter anderem auch das Orpheus-Projekt vorgestellt. Bei einer Umfrage unter den Teilnehmern stimmten 70% dafür, dass es äußerst wichtig sei, offene Standards zu verwenden.[3]

Der Grundgedanke von objektbasiertem Audio stellt auch die Basis für das Projekt Orpheus dar, in dem europäische Rundfunkanstalten, Forschungsinstitute und Hersteller gemeinsam eigene Lösungen, größtenteils unter Verwendung offener Standards, für einen objektbasierten Produktionsbetrieb entwickeln.

Kanalbasiert im Rundfunk bedeutet bisher, dass Programminhalte (Sendungen, Features, Beiträge, Musik, usw.) in Stereo, seltener auch in 5.1, produziert werden. Die einzelnen Audiosignale der Mikrofone laufen durch ein Mischpult, werden im Panorama verteilt und es entsteht die fertige, kanalbasierte Stereo- oder 5.1-Mischung. Diese wird gesendet und kommt beim Hörer auch in Stereo oder 5.1 an. Abbildung 2.2 stellt diesen Ablauf vereinfacht in einem Schaubild dar.

Objektbasiert bedeutet nun, dass zum Audiosignal noch Metadaten hinzukommen, die Information u.a. über Lautstärke und Position im Raum enthalten. Erst beim Hörer wird das Abhörformat gewählt (Binaural für Kopfhörerwiedergabe, Stereo, 5.1, etc.) und die Mischung berechnet, gerendert. Diese Produktionsweise zeigt Abbildung 2.3.

[3] vgl. Homepage Orpheus (2017) - „Object-Based Audio“ Seminar der EBU

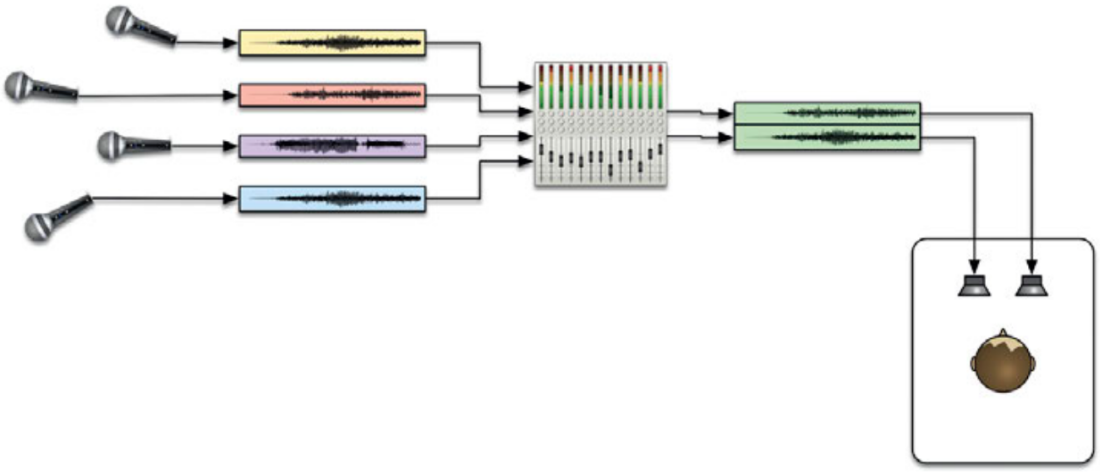

Abb. 2.2: Kanalbasierte Produktion für eine Stereomischung

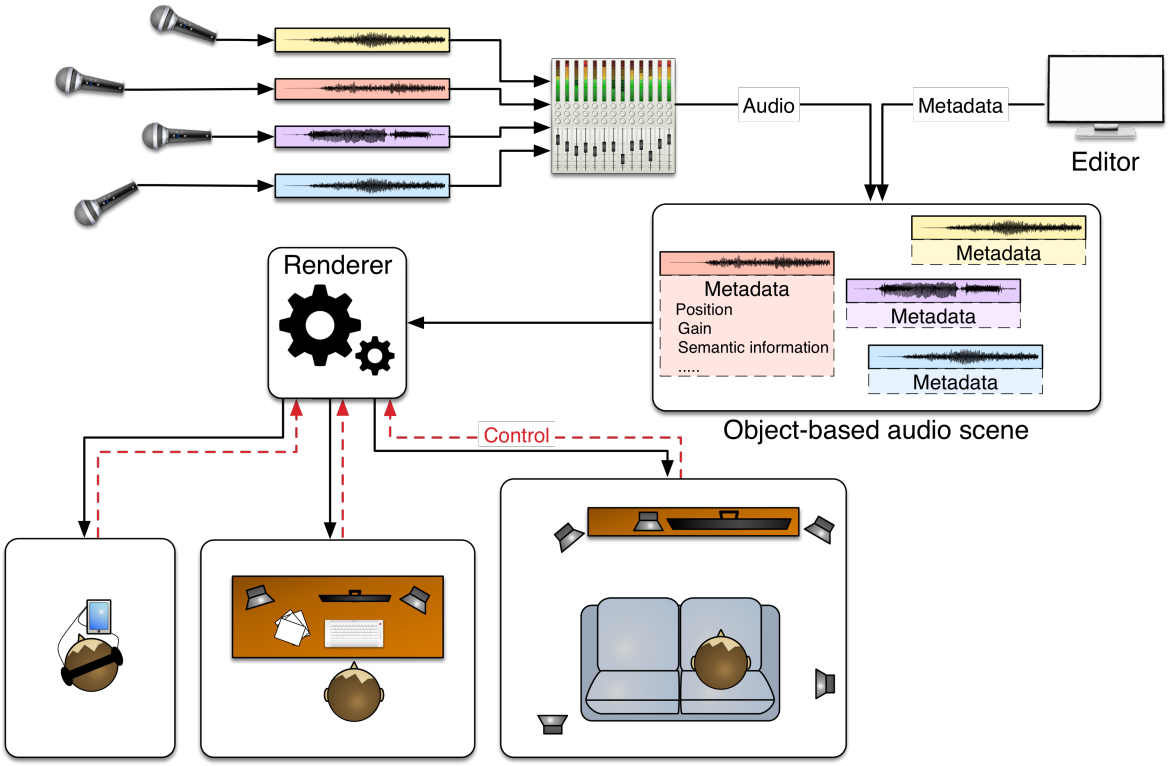

Abb. 2.3: Objektbasierte Produktion, bei der das Format erst beim Hörer entsteht

3 Orpheus

3.1 Projektvorstellung

Am 1. Dezember 2015 startete das von der EU im Rahmen von *Horizon 2020* unterstützte Forschungsprojekt „Orpheus - Audioerlebnis der nächsten Generation“. Das Projekt hat eine Laufzeit von 30 Monaten und endet im Mai 2018. Der vollständige Name lautet:

Orpheus - object-based audio experience
Object-based b**r**oadcasting - For Euro**p**ean Leaders**h**ip in Next **G**eneration A**u**dio Experience**s**

Ziel ist die Entwicklung einer komplett objektbasierten Produktionskette für Rundfunkanstalten. Durch die Nutzung vieler neuer Möglichkeiten, die objektbasiertes Audio bietet, soll eine Anpassung an die heutige Mediennutzung erfolgen und ein erheblicher Mehrwert für den Konsumenten entstehen.

Im Rahmen des Projekts soll untersucht werden, ob und wie aktuell existierende Rundfunkinfrastrukturen hin zu objektbasierter Produktion verändert werden können. Es soll eine Referenzarchitektur entwickelt, umgesetzt und exemplarisch getestet werden, die eine komplette objektbasierte Produktionskette abbildet. Durch die Nutzung und Entwicklung offener Standards und Formate soll eine einfache und kosteneffiziente Integration in bestehende Produktions- und Distributionsinfrastruktur ermöglicht werden. Die gewonnenen Erkenntnisse und Entwicklungen sollen bei entsprechenden Standardisierungsgremien, wie der ITU und der EBU, vorgelegt werden.

Für den Endkonsumenten soll ein neues Audioerlebnis geschaffen werden, das viele Möglichkeiten zur individuellen Anpassung und auch zur Barrierefreiheit bietet.

Öffentlich einsehbare Zwischenstände des Projekts sind als sog. „Public Deliverables“ auf der Homepage von Orpheus, www.orpheus-audio.eu, unter „Publications“ zu finden und dienen als Hauptliteraturquelle für diese Arbeit. Sie sind in die 6 Arbeitspakete des Projekts aufgeteilt. Updates informieren über die aktuellen Zwischenstände der jeweiligen Arbeitspakete und sind mit Unterpunkten (.2, .3, usw.) gekennzeichnet, z.B. Deliverable No. *D3.3 Object-based capture.*

3.1.1 Motivation

Der Gedanke von objektbasiertem Audio ist zwar nicht mehr ganz neu, neu ist jedoch die ganzheitliche Herangehensweise, eine komplett objektbasierte Rundfunkproduktionskette zu durchdenken und zu entwickeln: Aufnahme, Produktion, Speicherung, Archivierung, Sendung & Ausspielen, Distribution und Wiedergabe. Auf diese Weise kann das volle Potenzial von objektbasiertem Audio ausgeschöpft werden.

Kern von Orpheus ist die Erstellung einer Referenzarchitektur. Diese soll als generelle Richtlinie dienen, um einen objektbasierten Produktionsablauf in vorhandene Produktionsinfrastruktur von Rundfunkanstalten zu integrieren. Dabei sollen unterschiedliche Optionen zur Wahl stehen, damit individuell auf die Infrastruktur einer jeden Sendeanstalt eingegangen werden kann und es nicht nur „die eine richtige Lösung“ gibt.

Diese Referenzarchitektur wird anhand einer Pilotarchitektur erstellt. Diese Pilotarchitektur ist zwar allgemein gültig, wurde aber speziell an die Infrastruktur des Projektpartners BBC angepasst, da hier die Pilotphase 1 stattfand. Weil dies das erste Mal überhaupt ist, dass eine komplett objektbasierte Produktionskette geschaffen wird, gibt es noch keine Erfahrungswerte, auf die zurückgegriffen werden kann. Es muss sich zeigen, ob die getroffenen Überlegungen und technischen Entwicklungen auch in der Praxis umsetzbar sind und die gewünschten Ergebnisse erzielt werden können.

Dieser Pilot dient somit nicht nur Demonstrations- und Bewertungszwecken, sondern es soll auch herausgefunden werden, wie die Anforderungen an einen objektbasierten Produktionsalltag sind. Die Erfahrungen und Erkenntnisse aus Pilotphase 1 fließen anschließend wiederum in die Referenzarchitektur ein, die schließlich Schnittstellen und Komponenten liefert, die für eine objektbasierte Produktionskette nötig sind.[4]

[4] vgl. *D2.2 Interim Reference Architecture Specification And Integration* (2017) S. 7-9

3.1.2 Ziele

Das Projekt Orpheus verfolgt einen ganzheitlichen Lösungsansatz, indem durch die Erstellung der Referenzarchitektur die komplette Produktionskette abgebildet wird. Zusätzlich soll durch das Projekt der Übergang zu einer objektbasierten Audiozukunft in Europa geschaffen und der europäische Mediensektor wettbewerbsfähiger gemacht werden. Das Potenzial des Projekts soll sich auch auf Märkte außerhalb Europas auswirken, wie beispielsweise auf Nordamerika, Asien und Australien.

Mit dem objektbasierten Ansatz von Orpheus besteht auch das Potenzial, kosten- und aufwandseffizienter zu arbeiten. Derzeit ist es gängige Praxis, mehrere Versionen desselben Inhalts zu erstellen: für verschiedene Ausspielformate (meist Stereo und 5.1), unterschiedliche Fassungen etwa für schwerhörige oder blinde Menschen oder gar komplett neue Mischungen in einer anderen Sprache. Dies bedeutet einen zeitlichen und finanziellen Mehraufwand. Im Rahmen von Orpheus ist es möglich, eine einzige objektbasierte Version zu erstellen und zu übertragen, die dann individuell angepasst werden kann. Erst der Konsument entscheidet, welche Unterversion und welches Format er hört, ob Stereo, 5.1 oder Binaural über Kopfhörer. Sind von einer Sendung mehrsprachige Versionen vorhanden, kann auch zwischen diesen gewählt werden.[5]

3.1.3 Neue Möglichkeiten

Durch diese objektbasierte Herangehensweise sollen sich dem Nutzer 4 neue Möglichkeiten bieten. Diese sind wie folgt:

1. Zugänglichkeit
Hiermit soll eine barrieresenkende Wirkung erzielt und es Menschen mit Behinderungen wie Schwerhörigkeit oder Blindheit erleichtert werden, Rundfunkinhalte zu nutzen. Mit einem Regler kann etwa das Mischungsverhältnis zwischen Sprache und Hintergrundgeräuschen angepasst und somit die Sprachverständlichkeit erhöht werden. In einem Live-Transkript des Gesprochenen kann der Inhalt zusätzlich nachgelesen werden.

[5] vgl. *D6.1 Dissemination Plan And Innovation Roadmap* (2016) S. 9 f

2. Immersion

Immersion bedeutet das vollständige Eintauchen in ein virtuelles Geschehen, in diesem Fall das Klangerlebnis. Der Hörer hat z.B. durch Kopfhörer oder eine mehrkanalige Surround-Anlage die Möglichkeit, in den Klang einzutauchen und vollständig von ihm umgeben zu sein. Es entsteht ein realitätsnahes, dreidimensionales Erlebnis für die Ohren.

3. Interaktivität

Möglichkeiten zur Interaktivität werden derzeit im Rahmen der Pilotphase 2 erarbeitet. Eine der Überlegungen ist nonlineares Storytelling mit variabler Länge. Hierbei soll es dem Hörer möglich sein, beispielsweise nur eine Kurzfassung einer Sendung, eines Features, zu hören. Anhand von Cuepunkten, die während der Produktion definiert werden, werden lediglich die essenziellen Kernpunkte ohne Ausschmückungen oder Nebenhandlungen erzählt.

Eine weitere Überlegung ist, dass es dem Hörer möglich sein soll, bestimmte Elemente einer Sendung, wie etwa Musiktitel, austauschen zu können und damit eine Anpassung an den individuellen Musikgeschmack vorzunehmen.

4. Personalisierbarkeit

Hiermit kann der Hörer das Rundfunkprogramm an seine individuellen Vorlieben anpassen. Bei einer Fußballübertragung kann zum Beispiel gewählt werden, in welcher Fankurve man (virtuell) sitzen möchte. Durch einen Schieberegler kann zwischen Heim- oder Gastmannschaft gewählt werden. Beim Fernsehen bietet es auch die Möglichkeit, nur die Stadionatmosphäre zu hören und den Kommentator ganz abzuschalten.

Eine weitere Möglichkeit der Personalisierbarkeit ist die Wahl der Sprache. Ist eine Sendung in mehreren Sprachen verfügbar, kann die bevorzugte Sprache gewählt werden.

Objektbasierte Produktionen können auch an die jeweilige Abhörsituation angepasst und die Sprachverständlichkeit erhöht werden. Durch einen Regler kann das Verhältnis von Vordergrund (meist Sprache) zu Hintergrund (Geräusche, Musiken, oder Atmosphären, kurz Atmos) verändert oder auch der Grad der Kompression erhöht werden, wenn sich der Hörer gerade in einer lauten Umgebung, wie etwa dem Auto oder öffentlichen Verkehrsmitteln, befindet.[6]

[6]Gespräch über neue Anwendungsmöglichkeiten - mit W. Bleisteiner (2017)

3.1.4 Anwendungsbeispiel

Ein mögliches Szenario aus dem Alltag könnte wie folgt aussehen:

Emma sitzt morgens zu Hause im Wohnzimmer und isst in Ruhe ihr Frühstück. Über ihre 5.1 Surround-Anlage mit integriertem Receiver für objektbasiertes Audio hört sie Radio. Der Receiver nutzt das WLAN und erstellt ein 5.1 Format mit voller Programmdynamik.

Anschließend macht sie sich auf den Weg zur Arbeit. Zu Fuß geht sie zur nahegelegenen U-Bahn Station und nutzt eine App auf ihrem Smartphone, um die Radiosendung weiter anhören zu können. Die App erkennt, dass Emma über Kopfhörer hört und erstellt ihr eine Binauralfassung. Auch merkt die App, dass Emma sich nicht in einem WLAN befindet, sondern mobile Daten nutzt, und wählt ein dateneffizienteres Format, das weniger Datenvolumen verbraucht.

Beim Einsteigen in die U-Bahn merkt Emma, dass sie kaum noch etwas vom Radioprogramm versteht, da es im Hintergrund sehr laut ist. In der App dreht sie nun am Regler für Sprachverständlichkeit. Der Kompressionsgrad wird erhöht und Emma kann der Radiosendung wieder gut folgen.

Mitten im U-Bahn-Tunnel ist das mobile Internet sehr schwach, so dass der Empfang wenige Minuten lang unterbrochen ist. Als wieder Internetempfang vorhanden ist, springt Emma in der Zeitleiste der App ein Stück zurück und kann ab dem Zeitpunkt weiterhören, ab dem die Verbindung unterbrochen war. Als ein Musiktitel gespielt wird, der nicht einer ihrer Lieblingstitel ist, springt sie wieder zum Echtzeitempfang vor.

Als Emma in der Arbeit ankommt und ihr ruhiges Büro betritt, schaltet sie ihr Radiogerät an. Dieses gibt Stereo als Ausgabeformat über die beiden integrierten Lautsprecher wieder.

3.2 Projektstruktur

Das Projekt ist in 6 Arbeitspakete gegliedert, die sogenannten „Workpackages“ (WP). Diese sind:

- WP1 - Projektmanagement
- WP2 - Pilot- und Referenzarchitektur
- WP3 - Produktion
- WP4 - Speicherung und Archivierung
- WP5 - Bereitstellung und Darstellung
- WP6 - Verbreitung, Innovationsmanagement, Verwertung und Standardisierung

Die Arbeitspakete sind wiederum in Teilaufgaben, sog. „Tasks“ (T), aufgeteilt, wie in Abbildung 3.1 zu sehen ist. Jedem Arbeitspaket und jeder Teilaufgabe sind die einzelnen Projektpartner zugeordnet, so dass jeder Partner sich auf seinem Spezialgebiet in das Projekt einbringen kann.

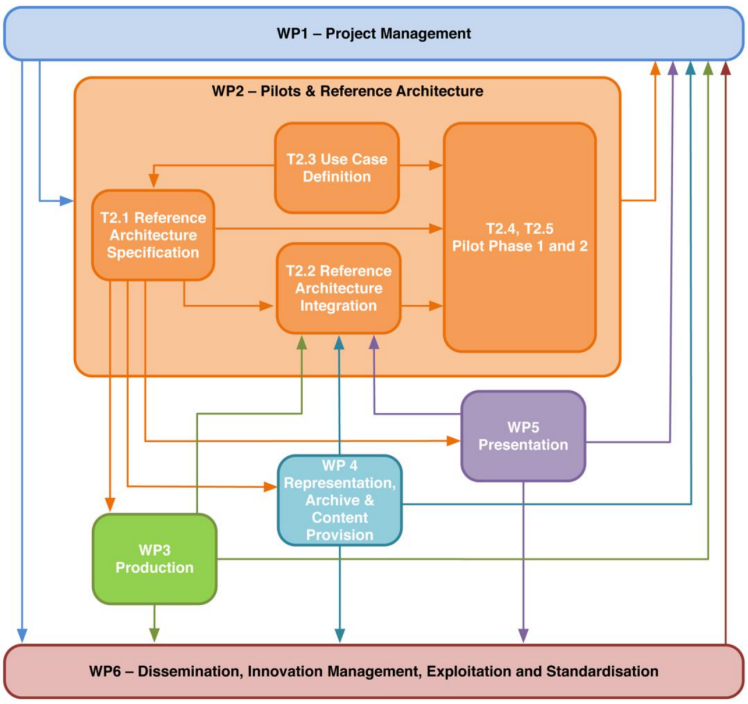

Abb. 3.1: Grafische Übersicht der Projektstruktur

WP1 bildet den administrativen Rahmen für das Projekt und kümmert sich um das Management. Hier wird sichergestellt, dass Zeitpläne und Budget eingehalten und Berichte zum aktuellen Arbeitsfortschritt rechtzeitig fertiggestellt werden. Zudem werden hier Tools für die Kommunikation der Partner untereinander und zur Europäischen Komission bereitgestellt.

WP2 beinhaltet den Kern des Projekts und ist somit das umfangreichste Arbeitspaket. Hier werden Entscheidungen hinsichtlich der Spezifikationen der Pilot- und Referenzarchitektur getroffen. Die Architektur wird implementiert und als Pilotphase 1 und Pilotphase 2 durchgeführt und getestet. Ergebnisse und Entwicklungen aus anderen Arbeitspaketen fließen in dieses Arbeitspaket mit ein und werden ebenfalls in den Piloten getestet, überprüft und weiterentwickelt.

Pilotphase 1 fokussiert sich auf eine lineare Live-Sendung in Form eines Hörspiels für unterschiedliche Plattformen. Es bietet dem Hörer die Möglichkeit selbst zu entscheiden, aus der Sicht welchen Charakters er die Geschichte erleben möchte.

Pilotphase 2 entwickelt Funktionen für nicht-live-Inhalte (onDemand) und schafft ein nicht-lineares Nutzererlebnis (User Experience). Es werden zeitbasierte Funktionen getestet und es wird überprüft, wie Sendungen an die Wünsche der Hörer angepasst werden können.

In **WP3** werden Tools für die Produktion (Aufnahme, Bearbeitung und Mischung) definiert, entwickelt und getestet. Es entstehen Beispielproduktionen, anhand derer die bisherigen Entwicklungen für eine komplett objektbasierte Produktion getestet werden können.

WP4 beschäftigt sich mit der Speicherung und Archivierung von objektbasierten Audioinhalten, sowie deren Übertragung zum Konsumenten. Es wird untersucht, welche Formate und Codecs sich für diese Zwecke eignen und diese gegebenenfalls (weiter-) entwickelt.

WP5 kümmert sich um die Entwicklung von Tools und Algorithmen für die Bereitstellung und Darstellung von objektbasiertem Audio. Es werden Renderer für unterschiedliche Plattformen als Software- und Hardwarelösungen entwickelt, mit denen der Konsument die objektbasierten Inhalte nutzen kann.

WP6 kümmert sich abschließend um Verbreitung, Innovationsmanagement, Verwertung und Standardisierung. Es geht darum, die im Rahmen des Projekts entstandenen Innovationen und Entwicklungen an eine breitere Öffentlichkeit zu tragen und in Medienunternehmen und Rundfunkanstalten einzuführen. Die entstandenen Entwicklungen sollen bei Standardisierungsgremien wie der EBU und der ITU-R vorgelegt werden und einen wichtigen Beitrag zu offenen Standards leisten.[7]

[7]Gespräch über die Aufgaben innerhalb der Arbeitspakete - mit W. Bleisteiner (2017)

3.3 Projektpartner

Fraunhofer IIS

Das Fraunhofer IIS (Institut für integrierte Schaltungen) ist Teil der Fraunhofer-Gesellschaft zur Förderung der angewandten Forschung e. V. und hat seinen Sitz in Erlangen, Deutschland. Es ist Europas größte Organisation für angewandte Forschung in Bereichen wie Audio & Multimedia, Kommunikation & Wissen und Gesundheit & Umwelt. Im Bereich Audiocodierung und Multimediatechnologien ist es die weltweit führende Forschungseinrichtung.[8]

Das Fraunhofer IIS ist beteiligt an großen Standardisierungsgremien wie der EBU, der ITU und MPEG. Der vom Fraunhofer IIS entwickelte mp3-Codec ist weltweiter Standard und aus dem Audiobereich nicht mehr wegzudenken. Die Technologie vom Fraunhofer IIS ist Teil des MPEG-H 3D Audio Standards (siehe Kapitel 10.3). Dieser ermöglicht eine Personalisierung und Nutzerinteraktivität von objektbasiertem Audio.

Innerhalb von Orpheus ist das Fraunhofer IIS für die Koordination des gesamten Projekts zuständig, insbesondere aber für die technische Koordination, und übernimmt zusammen mit Eurescom das Projektmanagement. Hauptaufgaben sind die Speicherung und Archivierung in WP4. Dabei kümmert sich das Fraunhofer IIS maßgeblich um T4.1 (Codecs und Formate). Auch im späteren Standardisierungsprozess spielt das Fraunhofer IIS eine tragende Rolle.[9]

Eurescom

Eurescom, European Institute for Research and Strategic Studies in Telecommunications GmbH, ist ein international führendes Dienstleistungsunternehmen für Projektmanagement mit Sitz in Heidelberg, Deutschland. Speziell im Telekommunikationsbereich hat Eurescom jahrelange Erfahrung und auch mit EU geförderten Informations- und Kommunikationstechnologie-Projekten.[10]

Innerhalb von Orpheus stellt Eurescom das allgemeine Projektmanagement und unterstützt das Fraunhofer IIS in WP1 beim Projektmanagement. Eurescom ist zuständig für Verwaltungsaufgaben wie beispielsweise das Bereitstellen von Kommunikationswerkzeugen und das Organisieren von Sitzungen. Eurescom regelt das Zusammenstellen von Informationen für regelmäßige Finanzberichte an die Europäische Kommission.[11]

[8] vgl. Homepage Fraunhofer IIS (2017)
[9] vgl. *D6.1 Dissemination Plan And Innovation Roadmap* (2016) S. 27
[10] vgl. Homepage Eurescom (2017)
[11] vgl. *D6.1 Dissemination Plan And Innovation Roadmap* (2016) S. 27 f

BBC | R&D

Die BBC, British Broadcasting Corporation, ist Großbritanniens öffentlich-rechtlicher Rundfunkanbieter für Radio, Fernsehen und Internet mit Hauptsitz in London, United Kingdom. Die Abteilung R&D, Research and Development, ist weltweit führend in der Technologieentwicklung und im Testen von Prototypen durch engen Austausch mit den Programmmachern und dem stetigen Erheben von Nutzeranforderungen.

Die BBC setzt die in Orpheus entwickelten Ideen im Rahmen der Pilotphase 1 in einer realen Produktionsumgebung um und testet diese hinsichtlich ihrer praktischen Anwendung. Kernkompetenzen der BBC liegen dabei im Bereich Audio und User Experience.[12]

IRT

Das IRT, Institut für Rundfunktechnik GmbH, ist ein Forschungsinstitut für Rundfunk- und Multimediatechnologien und befindet sich in München, Deutschland. Es ist ein gemeinschaftliches Tochterunternehmen der öffentlich-rechtlichen Rundfunkanstalten von Deutschland, Österreich und der Schweiz und arbeitet an der stetigen Weiterentwicklung und Qualitätsverbesserung von Radio und Fernsehen. Dabei setzt das IRT auf offene Standards. Bereits seit Jahrzehnten beschäftigt es sich mit der Forschung und Entwicklung zu immersivem Audio, beginnend mit Kunstkopfaufnahmen, der Forschung an binauralen Raumsynthesesystemen und in den vergangenen Jahren der erfolgreichen Forschung an Prototypen zu objektbasiertem Audio.

Das IRT kümmert sich innerhalb von Orpheus um die Erforschung von Lautheitsmetern in einer objektbasierten Umgebung und um die Entwicklung eines Renderers. Des Weiteren leitet das IRT WP2 und T4.2.[13]

Elephantcandy

Elephantcandy ist ein Privatunternehmen aus Amsterdam in den Niederlanden. Es entwickelt mobile Apps für Musik & Audio und Audioanalyse und -rendering Komponenten für private wie auch öffentliche Unternehmen.

Für Orpheus konzeptioniert und entwickelt Elephantcandy eine App, die objektbasiertes Audio in hoher Qualität für unterschiedliche mobile Endgeräte rendern kann.[14]

[12] vgl. *D6.1 Dissemination Plan And Innovation Roadmap* (2016) S. 28
[13] vgl. *D6.1 Dissemination Plan And Innovation Roadmap* (2016) S. 29
[14] vgl. *D6.1 Dissemination Plan And Innovation Roadmap* (2016) S. 29 f

Trinnov Audio

Trinnov Audio ist ein Hersteller von Audioprozessoren aus Neuilly-Plaisance, Frankreich. Hauptanliegen des Unternehmens ist es, die Audioqualität in jedweder Abhörumgebung zu verbessern, sei es zu Hause, im Kino, im Konzertsaal oder im Studio.

Innerhalb von Orpheus bringt Trinnov Audio sein umfassendes Wissen im Bereich Raumklang mit ein und entwickelt neue Technologien für objektbasiertes Audio.[15]

b<>com

b<>com ist ein Forschungsinstitut in Cesson-Sévigné, Frankreich. Sein Schwerpunkt liegt im Bereich von Hypermedien, wie beispielsweise UHD, VR und 3D Audio, sowie Hochgeschwindigkeitsnetzwerken & Sicherheit. Das *Advanced Media Coding Lab* erforscht neue Aspekte von 2D und 3D im Bildbereich wie auch von räumlichem Klang im Tonbereich. Ziel ist es, deren Realismus und Wirkung zu verbessern. Des Weiteren untersucht das *Uses & Acceptability Lab* User Experience, Ergonomie und digitales Design.[16]

Im Rahmen von Orpheus beschäftigt sich b<>com hauptsächlich mit der Entwicklung von Mikrofonarrays und deren Signalverarbeitung für die Aufnahme von dreidimensionalen Audioszenen im Ambisonics Format und im Higher Order Ambisonics Format (HOA; siehe Kapitel 6.1).[17]

Ircam

Ircam, Institut de Recherche et Coordination Acoustique/Musique, ist ein Forschungszentrum für Musik und Akustik in Paris, Frankreich. Es ist dem Centre Pompidou unter der Schirmherrschaft des französischen Kultur- und Kommunikationsministeriums angegliedert. Aktuell forscht Ircam an Analyse, Synthese und wahrnehmungsorientierter Verarbeitung von 3D Raumimpulsantworten, die mit Mikrofon- und Lautsprecherarrays gemessen werden.

Für Orpheus kümmert sich Ircam um die Entwicklung von Hall. Des Weiteren arbeitet es an der Weiterentwicklung des ADM-Metadatenformats (siehe Kapitel 5.2) und an Produktionstools für ADM.[18]

[15] vgl. *D6.1 Dissemination Plan And Innovation Roadmap* (2016) S. 30
[16] vgl. Homepage b<>com (2017)
[17] vgl. *D6.1 Dissemination Plan And Innovation Roadmap* (2016) S. 31
[18] vgl. *D6.1 Dissemination Plan And Innovation Roadmap* (2016) S. 31 f

Bayerischer Rundfunk (BR)

Der BR ist die öffentlich-rechtliche Rundfunkanstalt des Freistaats Bayern mit Hauptsitz in München, Deutschland. Der BR war immer an vorderster Front dabei, wenn es darum ging, neue Technologien anzuwenden: die Einführung von Stereoübertragungen in den 1960er Jahren, Kunstkopfproduktionen in den 70ern, digitales Aufnehmen und Übertragen in den 80er und 90er Jahren und seit den 2000er Jahren interaktive Hybridfernseh- und Radioangebote.[19]

Im Orpheus Projekt testet der BR die verschiedenen, von den Projektpartnern entwickelten, Tools und wendet sie in exemplarischen Produktionen an. Im Rahmen von Orpheus sind bisher folgende Produktionen entstanden:

- eine etwa 5-minütige Fußballreportage
- eine mehrsprachige, 15-minütige Kurzfassung des einstündigen Features „Donnerschacht und Frösche in Schnapsgläsern - Die Kunst des Geräuschemachens“ (siehe Kapitel 7)
- ein 15-minütiges Erklärstück über „Objektbasiertes Audioerlebnis“
- eine ca. 5-minütige Quartettaufnahme

Diese Produktionen demonstrieren Möglichkeiten wie die Regelung von Vorder- und Hintergrund, Immersion, Sprachauswahl und das Mitlesen via Transkript.[20]

Magix

Magix Software GmbH stellt Multimedia-Software her und hat seinen Hauptsitz in Berlin, Deutschland. Sowohl für den Privatgebrauch als auch für den professionellen Bereich bietet Magix Programme zur Erstellung, Gestaltung, Archivierung und Präsentation von Fotos, Videos, Grafiken, Webseiten und von Musik. Viele Rundfunkanstalten und Musikstudios in Europa und weltweit nutzen die DAW *Sequoia* im Bereich (Klassik-) Musikproduktion und im Mastering.[21]

Die Aufgabe von Magix innerhalb von Orpheus ist es, eine DAW bereitzustellen, die objektbasiert arbeitet und die nötigen Anforderungen erfüllt. Hierfür wird *Sequoia* dahingehend weiterentwickelt und getestet.[22]

[19] vgl. *D6.1 Dissemination Plan And Innovation Roadmap* (2016) S. 32
[20] vgl. *D5.3 Document On Methodology Of Evaluation Of User Experience* (2017) S. 20
[21] vgl. Homepage Magix (2017)
[22] vgl. *D6.1 Dissemination Plan And Innovation Roadmap* (2016) S. 32 f

4 Pilot- und Referenzarchitektur

Eine für das ganze Projekt zentrale Rolle hat die Entstehung der Pilot- und Referenzarchitektur, die in WP2 angesiedelt ist. Hier werden Schnittstellen definiert, Rückmeldungen anderer Arbeitspakete kommen an und Entscheidungen für die genaue Umsetzung der Piloten werden getroffen.

Ziel der Referenzarchitektur ist es, Schnittstellen, Anforderungen und Spezifikationen einer komplett objektbasierten Produktionskette zu definieren. Soweit es möglich ist, werden dabei offene Standards verwendet. Die Referenzarchitektur soll künftig anderen Rundfunkhäusern als Richtlinie für die Implementierung objektbasierter Formate, Tools und Komponenten in ihre eigene Produktionsinfrastruktur dienen. Die aus dem Orpheusprojekt durch Tests gewonnenen Erfahrungen fließen in die Definition der Referenzarchitektur mit ein und ebenso in die im Rahmen des Projekts entwickelten Produktionstools.

Eine erste Variante der Referenzarchitektur ist die Pilotarchitektur. Die daraus gewonnenen Erfahrungen werden genutzt, um die Referenzarchitektur zu verbessern. Auch gibt die Pilotarchitektur einen ersten Aufschluss darüber, inwieweit Rundfunkanstalten diese neue Technologie für ein Rund-um-die-Uhr-Programm übernehmen können. Da sich aber der gesamte bisherige Arbeitsprozess grundlegend ändert und noch wenig Erfahrungswerte und Wissen im Umgang mit objektbasiertem Audio vorhanden sind, stellt diese Aufgabe eine große Herausforderung dar.

Die Pilotarchitektur ist zwar allgemein gültig, jedoch ist sie an die Infrastruktur der BBC angepasst, da hier die Pilotphase 1 stattfand. Sie ist unterteilt in Makroblöcke, die die einzelnen Glieder einer Produktionskette abbilden:

Aufnahme - Vorproduktion und Mischung - Sendestudio - Distribution - Wiedergabe

Die Makroblöcke sind unabhängig voneinander und bestehen aus weiteren kleinen Elementen, wie Abbildung 4.1 zeigt. Sie entsprechen in etwa den Arbeitspaketen.[23]

[23] vgl. *D2.2 Interim Reference Architecture Specification And Integration* (2017) S. 7-11

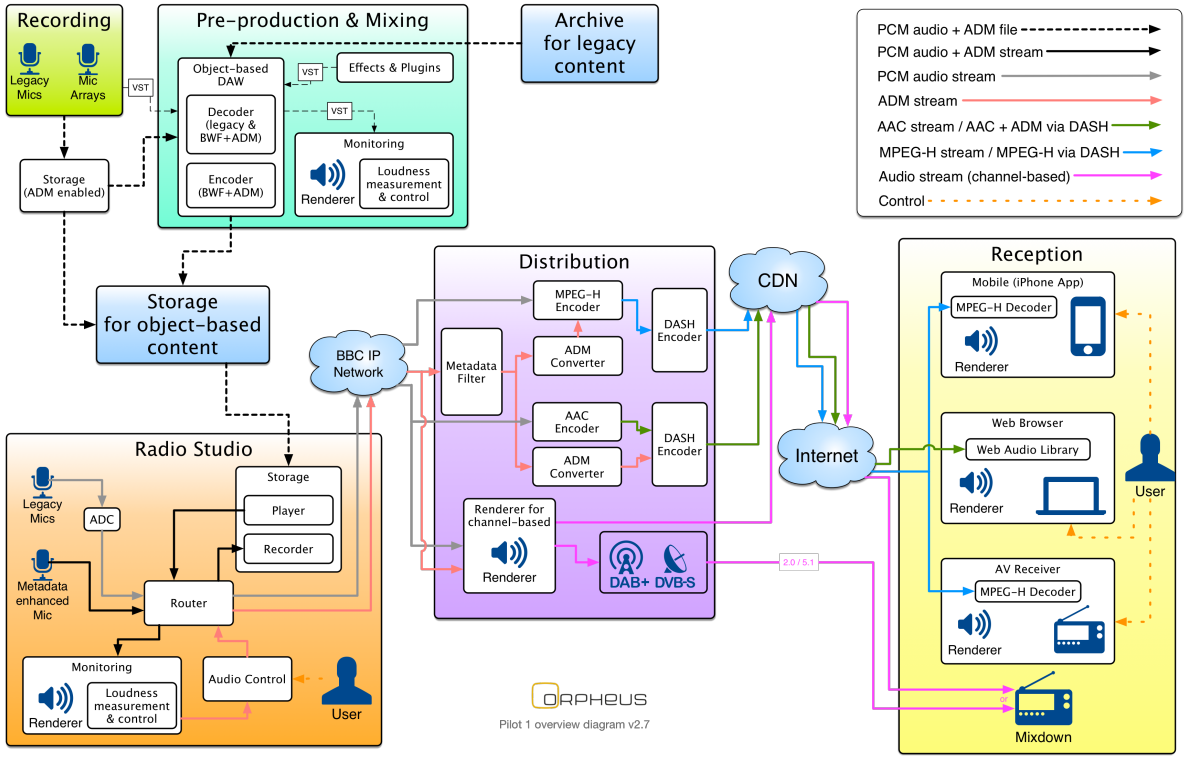

Abb. 4.1: Grafische Übersicht über die Makroblöcke der Pilotarchitektur

Im **Aufnahme**-Makroblock geht es darum herauszufinden, welche Anforderungen an die Infrastruktur und die Tools für die Aufnahme gestellt werden. Es gibt drei Aufnahmearten: kanalbasiert, szenenbasiert und objektbasiert. Diese werden in Kapitel 6 näher erläutert.[24]

Da Audiosignale nicht nur live sind, sondern auch für die Vorproduktion aufgezeichnet werden müssen, ist der Aufnahme-Makroblock eng mit **Vorproduktion und Mischung** verknüpft. Die Audiosignale werden im Broadcast Wave 64 bit Format (BW64; siehe Kapitel 5.1) zusammen mit den dazugehörigen ADM Metadaten in einer DAW aufgezeichnet, dort bearbeitet, ausgesteuert, gemischt und exportiert. Im Rahmen von Orpheus wird dafür die DAW Sequoia verwendet, die für diesen Zweck weiterentwickelt wird, um mit Metadaten umgehen und diese bearbeiten zu können. Des Weiteren müssen die neu entwickelten Plugins für objektbasiertes Audiomaterial eingebunden werden können. Auch muss es möglich sein, herkömmliches Audiomaterial (z.B. aus Archiven) um Metadaten zu ergänzen und somit für eine objektbasierte Produktion nutzen zu können.

[24] vgl. *D2.2 Interim Reference Architecture Specification And Integration* (2017) S. 11

An diesen Makroblock sind die beiden kleineren Elemente *Speicherung* und *Archivierung* eng angegliedert.[25]

Im **Sendestudio** werden einzelne Signalquellen (Mikrofonsignale, vorproduzierte Beiträge, Musiken, Leitungen, etc.) kombiniert und es entsteht das Radioprogramm. In einem objektbasierten Umfeld kann auch von einer Komposition, statt von einer Mischung gesprochen werden, denn die Audiosignale bleiben getrennt. Sie sind aber mit Metadaten versehen, die dem Renderer im Produktionsstudio oder im Endgerät des Hörers vorgeben, wie das fertige Signal zusammengesetzt werden muss.[26]

Der Makroblock **Distribution** befasst sich damit, wie ausgestrahlte Inhalte vom Sendestudio zum Hörer übertragen werden. Das Produktionsformat (BW64 + ADM) muss in ein Format gewandelt werden, das sich besser als Streamingformat eignet (z.B. AAC und MPEG-H; siehe Kapitel 10). Hauptübertragungskanal ist das Internet. Um jedoch abwärtskompatibel zu bleiben, kommen auch weitere Distributionswege in Betracht.[27]

Im letzten Makroblock, **Wiedergabe**, wird überlegt, wie der Empfang, die Personalisierung und die Wiedergabe von objektbasiertem Inhalt funktioniert und welche Hardware- und Softwarelösungen es dafür geben kann. Drei Ansätze werden verfolgt: Wiedergabe in einer App auf dem Smartphone, Wiedergabe im Web-Browser und Wiedergabe über einen AV-Receiver.[28]

[25] vgl. *D2.2 Interim Reference Architecture Specification And Integration* (2017) S. 12-15
[26] vgl. *D2.2 Interim Reference Architecture Specification And Integration* (2017) S. 15-19
[27] vgl. *D2.2 Interim Reference Architecture Specification And Integration* (2017) S. 19
[28] vgl. *D2.2 Interim Reference Architecture Specification And Integration* (2017) S. 23

5 Speicherung und Archivierung

Mit der Speicherung und Archivierung von objektbasiertem Audiomaterial beschäftigt sich Arbeitspaket WP4. Hier wird bestimmt, welche Formate innerhalb von Orpheus relevant sind und zum Einsatz kommen.

Ein wichtiger Aspekt bei der Speicherung von Audiomaterial ist die Wiederverwendbarkeit für künftige Produktionen. Daher ist es zwingend notwendig, die Audioqualität zu erhalten und diese nicht durch eine kaskadierte Reduzierung der Bitrate zu verschlechtern. Da Speicher immer preiswerter werden und Netzwerke größere Bandbreiten haben, ist es theoretisch nicht mehr nötig, während der Produktion reduzierte Bitraten zu verwenden.[29]

5.1 Broadcast Wave BW64

Das Broadcast Wave 64 bit Format (BW64) ist der Nachfolger des im Rundfunk üblichen Broadcast Wave Formats (BWF). Durch die Verwendung von 64 bit sind Dateigrößen von mehr als 4 GB möglich, wodurch eine große Anzahl an Mehrkanaldateien und Objekten gespeichert werden kann. Des Weiteren können ADM Metadaten mit abgespeichert werden. In einem BW64 Container können unterschiedliche Audioformate verpackt werden. Das innerhalb von Orpheus relevante Format ist PCM mit 24 bit und einer Abtastrate von 48 kHz.

Für den Austausch unter den Projektpartnern werden BW64 und ADM Metadaten verwendet. BW64 wurde in Sequoia implementiert, ebenso wie die ADM Tools vom Ircam.[30]

[29] vgl. *D3.4 Implementation And Documentation Of A Live Object-based Production Environment* (2016) S. 15

[30] vgl. *D4.2 Interim Report Object-based Radio* (2017) S. 10

5.2 Audio Definition Model (ADM)

Das innerhalb von Orpheus verwendete Modell zur Beschreibung von Metadaten ist das Audio Definition Model (ADM), da es aktuell das zuverlässigste und vollständigste Modell ist. AMD wurde im Juni 2015 als Recommendation ITU-R BS.2076-0 spezifiziert und im Juni 2017 ergänzt. Das Datenmodell ist flexibel und erlaubt kanal-, szenen- und objektbasiertes Audio unabhängig von der Dateigröße oder der Komplexität des Audios zu beschreiben.

In der ITU-R Spezifikation wird XML als beschreibende Sprache verwendet, die auf dateibasierten Gebrauch abzielt wie etwa BW64. Dadurch wird das Audio zusammen mit seinen beschreibenden Metadaten als eine einzige, zusammengehörige Datei betrachtet. Für Streaminganwendungen, bei denen Audio und Metadaten in Echtzeit erstellt und übertragen werden, eignet es sich nicht, sondern nur für fertig erstellte Dateien. Die ITU-R arbeitet aktuell an der Entwicklung einer seriellen Version von ADM, mit der Echtzeitstreaming möglich sein soll.

ADM bietet eine Vielzahl an Parametern und Eigenschaften. Einige sind für das Orpheusprojekt nicht relevant, andere wären, besonders im Hinblick auf Interaktion, wünschenswert, doch sind sie noch nicht vorhanden. Sie müssen erst noch spezifiziert werden.

Mit dem Parameter *typeDefinition* kann beispielsweise zwischen *Objects*, objektbasiertem Audio, oder zwischen *DirectSpeakers*, kanalbasiertem Audio, unterschieden werden.

Für *Objects* gibt es wiederum weitere beschreibende Eigenschaften wie *Position*, *Gain*, *Diffuse* und *jumpPosition*. Für alle diese Eigenschaften ist definiert, welche Werte einstellbar sind, bei *Position* etwa Winkel zwischen -180.0 und 180.0 oder bei *Gain* Werte von 0.0 bis unendlich. Es ist auch möglich, Standardwerte festzulegen.[31]

[31] vgl. *D4.2 Interim Report Object-based Radio* (2017) S. 10-12

5.3 Archivierung

Für Rundfunkanstalten ist es extrem wichtig, ein Langzeitarchiv zu haben, in dem Audio-(und Video-)material über Jahrzehnte hinweg gespeichert werden kann. Zum einen wird gesendetes Material gesichert und zum anderen wird Archivmaterial für künftige Produktionen verwendet (beispielsweise Geräusche, Atmos und Musiken). Es ist daher unbedingt notwendig, auf dieses Material ohne Verlust an Informationen und Qualität zugreifen zu können. Um das zu garantieren, sollten Produktions- und Archivierungsformate folgende Anforderungen erfüllen:

- öffentlich und frei verfügbar dokumentiert sein
- nicht proprietär sein oder von proprietären Geräten abhängen
- nicht von einem bestimmten Betriebssystem oder Gerätetyp abhängig sein
- ein unkomprimiertes und nicht datenreduziertes Format sein
- technische Metadaten enthalten
- mind. 48 kHz Abtastrate und eine Bittiefe von 24 bit unterstützen

Innerhalb von Orpheus einigten sich die Projektpartner auf BW64 als Containerformat, das PCM kodiertes, unkomprimiertes und nicht datenreduziertes Audio und ADM Metadaten enthält. Es ist ein kostenloses und frei verfügbares Format, das es zudem erlaubt, die technischen Metadaten zusammen mit den Audiodaten direkt zu speichern. Solange die Dateigröße unterhalb von 4 GB liegt, ist es auch abwärtskompatibel zu dem bisher verwendeten BWF Format.

Abwärtskompatibilität ist ein wichtiges Anliegen im Rahmen von Orpheus. Es muss möglich sein, neue Technologien einfach und unkompliziert in ein vorhandenes System mit bereits existierendem Inhalt zu integrieren.[32]

[32] vgl. *D4.2 Interim Report Object-based Radio* (2017) S. 23 f

6 Produktion

Für das objektbasierte Vorproduzieren von Beiträgen, Sendungen, Features usw. ändert sich die bisherige Arbeitsweise schon bei der Aufnahme. An eine DAW werden neue Anforderungen gestellt. Sie muss ein lineares Mehrkanalformat (wie etwa BW64 inklusive ADM Metadaten) erstellen, speichern, bearbeiten und exportieren sowie bereits existierende objektbasierte Dateien importieren können. Metadaten müssen editierbar sein und für bisheriges, nicht-objektbasiertes Material neu erstellt werden können (z.B. für Geräusche, Atmos und Musiken aus Archiven).

Auch muss es möglich sein, neue Plugins & Effekte, die objektbasiert arbeiten und innerhalb von Orpheus entwickelt werden, in die DAW einzubinden. Die typischen genutzten Effekte sind Hall, Equalizer und dynamische Effekte (Kompressor & Limiter und Expander & Gate). Für Letztere ergeben sich einige Einschränkungen. Besonders Kompressoren und Limiter werden in der DAW gerne in Summenwege eingebunden. Objektbasiert ist das jedoch nicht möglich, da die Effekte nur auf einzelne Objekte angewendet werden können.[33]

Mit dieser Thematik beschäftigt sich der Makroblock Vorproduktion & Mischung. Task T3.1 kümmert sich um die Entwicklung von Mikrofonarrays und deren Signalverarbeitung. Im Rahmen von Task T3.2 werden entsprechende DAW-Erweiterungen entwickelt und implementiert. Innerhalb von Orpheus entwickelte Plugins werden im Virtual Studio Technology (VST)-Format erstellt, da diese Schnittstelle die entsprechend hohe Busbreite (Kanalanzahl) ermöglicht. Im Unterschied dazu sind beim ProTools AAX-Pluginformat bei Bussen derzeit nur 10 Kanäle möglich.[34]

[33] vgl. *D2.2 Interim Reference Architecture Specification And Integration* (2017) S. 12-14
[34] Gespräch über die Verwendung von VST Plugins innerhalb von Orpheus - mit W. Bleisteiner (2017)

6.1 Aufnahme

Jede Radiosendung beginnt mit der Aufnahme von Audiosignalen. Für objektbasierte Audioproduktionen bedarf es auch hierfür neuer Methoden und Geräte. Um in der Produktion alle Freiheiten zu haben, sollten die Signale entweder in Mono oder Stereo so direkt wie möglich aufgenommen werden oder als 3D Audioszene mit den entsprechenden Metadaten für Gain, Position und Halleinstellungen. Für die Aufnahme von 3D Audioszenen werden Programme benötigt, die die Signale von Mikrofonarrays in Mehrkanal- oder Higher Order Ambisonics (HOA) Audioszenen umwandeln. Diese Programme werden von den Projektpartnern b<>com, Ircam und Fraunhofer IIS entwickelt.[35]

In einer objektbasierten Audioproduktionsumgebung gibt es folgende drei Verfahren:

Kanalbasiert werden Audiosignale in der DAW über einen Panoramaregler direkt einzelnen Lautsprecher zugeordnet, z.B. bei Stereo oder 5.1.

Szenenbasiert bilden die Signale ein dreidimensionales Klangfeld ab. Dabei entsprechen einzelne Signale aber nicht einzelnen Lautsprechern. Ein Beispiel hierfür sind Higher Order Ambisonics (HOA) Signale, kurz Ambisonics genannt. Bei dem *First Order Ambisonics B-Format* (FOA) werden 4 Kanäle aufgezeichnet, aus denen sich die Anteile für die einzelnen Lautsprecher errechnen lassen. Bei *Second Order* sind es 9, bei *Third Order* 16 Kanäle, usw. Je höher die Anzahl an Kanälen, desto präziser ist die räumliche Auflösung des Schallfelds.[36]

Objektbasiert wird ein Audiosignal als Audioobjekt bezeichnet. Die Raumposition dieses Objekts wird in den Metadaten gespeichert. Erst bei der Wiedergabe wird durch einen Renderer berechnet, auf welchem Lautsprecher welches Signal in welcher Lautstärke zu hören ist.[37]

[35] vgl. *D3.3 Object-based Capture* (2017) S. 3
[36] vgl. Wikipedia - Ambisonics (2017)
[37] vgl *D3.3 Object-based Capture* (2017) S. 8

6.1.1 Mikrofone

Auch in der objektbasierten Produktion ist es wichtig, eine klare Trennung der einzelnen Signale zu haben. Hierfür müssen die Mikrofone möglichst nah an der Schallquelle platziert werden. Somit erhält man ein „trockenes“ Signal, das wenig bis keine Rauminformationen enthält. Später wird dann die gewünschte Rauminformation als Hall hinzugefügt, etwa bei einem Hörspiel oder in der Musikproduktion. Hall kann entweder aufgezeichnet oder künstlich erzeugt werden.[38]

Für eine dreidimensionale Aufnahme, beispielsweise von Hall in hoher Qualität, ist es von Vorteil, kleine und leicht zu handhabende Mikrofonanordnungen, sog. „Mikrofonarrays“, zu benutzen. Ein Algorithmus wandelt die aufgezeichneten Mikrofonsignale anschließend in unterschiedliche Wiedergabeformate um, die für eine 3D Wiedergabe geeignet sind.

Das Fraunhofer IIS hat dafür einen speziellen 3D Audio Capturing Algorithmus entwickelt und in ein Echtzeit VST Plugin integriert. Das Plugin analysiert die Eingangssignale eines horizontalen, kreisförmigen Mikrofonarrays mit 8 Mikrofonkapseln (siehe Abbildung 6.1) und erzeugt ein Audiosignal im gewünschten 2D oder 3D Wiedergabeformat. Die Analyse des Eingangssignals geschieht anhand bestimmter Parameter, die das Signal beschreiben. Deshalb ist es sehr einfach möglich das Ausgangssignal zu verändern und es beispielsweise rotieren zu lassen.[39]

Gängige Mikrofone zum Aufnehmen von 3D Audioszenen sind Kugelarrays, bei denen die Mikrofonkapseln auf einer Kugeloberfläche angebracht sind. Ein Beispiel hierfür ist das *EigenMike* von *mh acoustics* (siehe Abbildung 6.2).

Auch die mit Kugelarrays aufgenommenen Signale müssen in ein Format gewandelt werden, mit dem es sich leichter arbeiten lässt, wie etwa das HOA Format. Die Qualität der herstellereigenen Algorithmen für diese Formatwandlung wird aber teilweise hochqualitativen Ansprüchen nicht gerecht. b<>com arbeitet daran, die Signalwandlung zu optimieren, um eine möglichst hohe Audioqualität zu erhalten.[40]

[38] vgl. *D3.3 Object-based Capture* (2017) S. 8
[39] vgl. *D3.3 Object-based Capture* (2017) S. 10-12
[40] vgl. *D3.3 Object-based Capture* (2017) S. 12-14

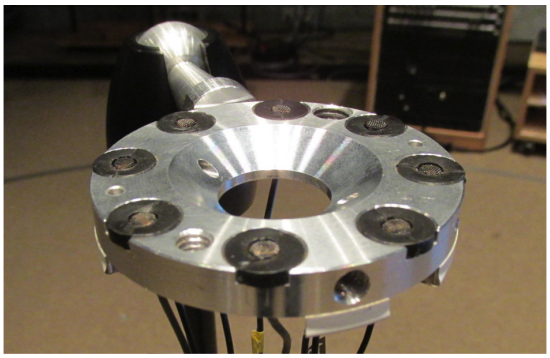

Abb. 6.1: Zylindrisches Mikrofonarray

Abb. 6.2: Kugelmikrofonarray *EigenMike*

6.1.2 Metadaten

Für die Bestimmung der Richtung, aus der ein Schallereignis kommt, hat b<>com einen Algorithmus entwickelt, der anhand der Signale eines Kugelarrays als Hauptmikrofon (Main) und einem Stützmikrofon (Spot) in direkter Nähe der Schallquelle die richtungsbestimmenden Winkel für Azimut und Elevation berechnen kann. Dieser *MainSpot Algorithmus* wurde in ein Echtzeit VST Plugin integriert, lieferte jedoch keine stabilen Ergebnisse, wenn beispielsweise Hall auftrat. Zudem funktioniert der Algorithmus nur für HOA Signale erster Ordnung.

Ein neuer Algorithmus wurde entwickelt, der bessere Resultate erzielt und auch keine Einschränkungen in der HOA Ordnung hat. In Tests bei Musikaufnahmen zeigte sich, dass der Algorithmus stabil läuft, sobald alle Musiker spielen und an jedem Stützmikrofon ein Signal anliegt. Die errechnete Richtung wich nur um wenige Grad von der tatsächlichen Position ab. Da der Rechenaufwand jedoch um einiges höher ist, als bei dem *MainSpot Algorithmus*, wird dieser neue Algorithmus vorerst nicht als VST Plugin implementiert werden, sondern als eigenständiges Programm laufen, damit die Position einer Schallquelle als ADM Metadaten exportiert werden kann.[41]

[41] vgl. *D3.3 Object-based Capture* (2017) S. 16-18

6.2 Bearbeitung

Szenenbasierte (Ambisonics-) Aufnahmen müssen in der Gesamtzahl ihrer Spuren (4, 9, 16, usw.) betrachtet werden. Daher braucht es andere Tools für das Bearbeiten und Aussteuern solchen Materials. b<>com hat zwei VST Plugins entwickelt, mit denen HOA-Szenen visualisiert und teils auch bearbeitet werden können.

Das *HoaScope* führt eine räumliche Analyse des HOA-Signals durch und stellt Position und Pegel farblich dar, wie in Abbildung 6.3 zu sehen ist. Die Farbe gibt Aufschluss über die akustische Energie mit der dazugehörigen Richtung, aus der der Schall kommt.

Der *Sound Field Editor* visualisiert das HOA-Signal ähnlich wie das *HoaScope*, verfügt aber zusätzlich noch über Kontrollpositionen. Die obere Hälfte von Abbildung 6.4 zeigt die Energie des Schallfeldes. Die eingezeichneten Punkte sind die Kontrollpositionen, die das Plugin aus dem Quellsignal heraus schätzt. Für diese Quellen kann eine Vorverstärkung (Gain), eingestellt werden, wie in der Abbildung 6.4 unten rechts zu sehen ist. Der Rechenaufwand für dieses Plugin ist relativ gering, allerdings können die Schallquellen nicht sehr genau getrennt werden. In komplexeren Szenen kann es so zu hörbaren Artefakten kommen.[42]

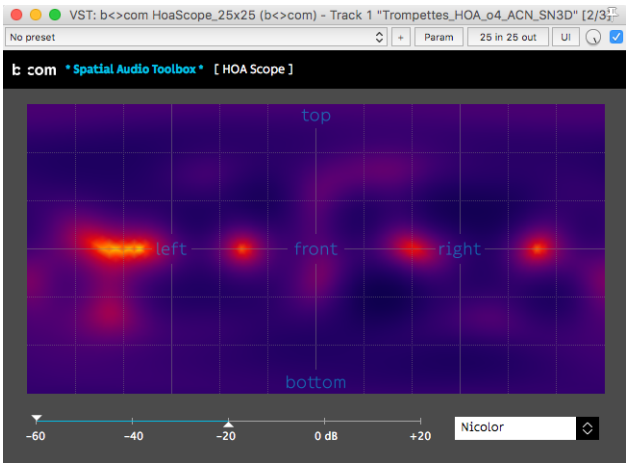

Abb. 6.3: HoaScope Plugin

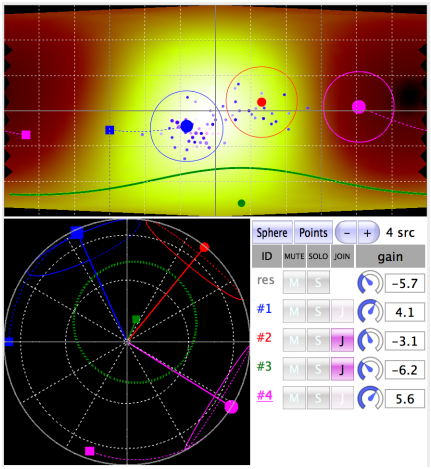

Abb. 6.4: Sound Field Editor Plugin

[42] vgl. *D3.3 Object-based Capture* (2017) S. 19-22

6.3 Hall

Die Verwendung von Hall ist für den gesamten Audioproduktionsbereich von enormer Wichtigkeit. Objektbasierte Hallalgorithmen und deren Einbindung werden vom Ircam untersucht und entwickelt. Dies ist eine sehr umfangreiche Aufgabe, die viele Herausforderungen mit sich bringt.

Es gibt unterschiedliche Arten von Hallmodellen. Manche davon können bereits in objektbasierte Metadatenmodelle implementiert werden. Beim ADM Metadatenmodell fehlt jedoch die explizite Beschreibung von Hallparametern, von Distanz- und Tiefenberechnung und von diffuser Kanal-/Objektberechnung.

Es gibt zwei Herangehensweisen von Metadatenmodellen an Hall - signalbasiert oder parametrisch. Im Fall der parametrischen Beschreibung wird der Hall erst durch den Renderer berechnet. Das hat den Nachteil, dass die Qualität des Halls stark vom verwendeten Renderer abhängt. Somit würde in der Produktion die Kontrolle über den Hall verloren gehen, da nicht simuliert werden kann, wie das Signal beim Hörer klingt.

Die signalbasierte Herangehensweise überträgt Hallsignale als eigene Audiokanäle und eignet sich für kanal- und szenenbasierte Produktionen. Auf diese Art wird auch bisher bei kanalbasierten Produktionen gearbeitet, bei denen der Hallanteil über Raummikrofone aufgezeichnet wird.[43]

Beim Abhören der objektbasierten Mischung des Features „Donnerschacht und Frösche in Schnapsgläsern - Die Kunst des Geräuschemachens“ über 40 Lautsprecher im *3D Audio Lab* des IRT zeigte sich, dass die während des binauralen Mischens über Kopfhörer als optimal empfundenen Hallanteile für die Wiedergabe über Lautsprecher nicht geeignet waren und um ca. 5 dB abgesenkt werden mussten. Das verdeutlicht, dass es Vorteile hat, den Hall als einzelnes Objekt zur Verfügung zu haben. Der Renderer kann somit getrennt vom Audiosignal darauf zugreifen und könnte bei Wiedergabe über Lautsprecher den Hall entsprechend der jeweilig vorhandenen Raumakustik anpassen. Ein dafür benötigtest Einmesssystem ist beispielsweise im Trinnov AV-Receiver *Altitude32* integriert (siehe Kapitel 11).

Die Entwicklung von Hallalgorithmen, die sich für objektbasierte Produktionen eignen, dauert noch an.

[43] vgl. *D3.2 Implementation And Documentation Of Reverberation For Object-based Broadcast* (2016) S. 17

6.4 Monitoring

Während der Produktion ist es wichtig, den Inhalt abzuhören und auch das Erlebnis des Hörers simulieren zu können. Hierfür müssen unterschiedliche Lautsprecheranordnungen zur Verfügung stehen und es muss auch die Möglichkeit gegeben sein, Binaural zu hören.

In Sequoia wurde der Renderer vom Fraunhofer IIS implementiert. Für die Referenzarchitektur soll jedoch ein Basisrenderer verwendet werden. Diese befindet sich aktuell noch in der Entwicklungs- und Standardisierungsphase der ITU-R.

Es ist auch erforderlich, die Audiosignale auf ihren Aussteuerungspegel und ihre Lautheit hin überprüfen zu können. Innerhalb von Orpheus wird ein Algorithmus verwendet, der Lautheit und True Peak nach EBU R.128 ermittelt. Diese Pegel können für jedes Audioobjekt einzeln berechnet werden, aber auch für den Ausgang des Renderers, um den Gesamtpegel überprüfen zu können.

Das IRT arbeitet an der Entwicklung eines objektbasierten Lautheitsmeters. Eine große Herausforderung besteht darin, dass während der Produktion noch nicht bekannt ist, mit welchem Renderer das Signal für welche Lautsprecheraufstellung generiert wird. Daher ist es wünschenswert, dass die Lautheitsmessung anhand einer Audioszene geschieht und nicht von Renderingmethode und Lautsprecherposition abhängig ist. Bis es ein wirklich objektbasiertes Lautheitsmeter gibt, werden innerhalb von Orpheus die Objekte in Kanäle umgerechnet und anschließend deren Lautheit gemessen.

Es werden mögliche Varianten zur Entwicklung eines objektbasierten Algorithmus zum Berechnen der Lautheit untersucht. Tests, bei denen nur die Objekte an sich gemessen wurden, ohne deren Position zu berücksichtigen und ohne eine Gewichtung vorzunehmen, lieferten bereits brauchbare Ergebnisse. Allerdings müssen noch weitere Tests mit realistischem Material durchgeführt werden, um wirklich aussagekräftige Ergebnisse zu erzielen. Diese Arbeit dauert noch an.[44]

[44] vgl. *D3.5 Specification And Implementation Of Reference Audio Processing For The Use In Content Creation And Consumption Based On Novel Broadcast Auality Standards* (2017) S. 19-24

7 Beispielproduktion - Die Kunst des Geräuschemachens

Anfang April 2017 wurde im BR das 53-minütige Radio-Feature „Donnerschacht und Frösche in Schnapsgläsern - Die Kunst des Geräuschemachens“ produziert. Es eignete sich gut für eine objektbasierte Umsetzung im Rahmen von Orpheus, weil es durch die große Anzahl der darin vorkommenden Geräusche viele kreative Gestaltungsmöglichkeiten bot. Der Münchner Geräuschemacher Max Bauer erzählt darin die Geschichte seines Berufs. Er beschreibt die Herausforderungen seiner täglichen Arbeit und wie sich durch die Wandlung von Analog zu Digital, die in den 1990er Jahren stattfand, sein Arbeitsalltag verändert hat.

7.1 Produktionsumgebung und -tools

Ein Produktionsteam besteht von technischer Seite her aus einem Toningenieur und einem Tontechniker. Der Ingenieur kümmert sich um das Signalrouting, die Aussteuerung und den Klang der Signale, sowie um die Plugins & Effekte und das Mischen. Der Techniker ist für die Aufnahme, das Editieren und das Platzieren der einzelnen Elemente an der gewünschten Stelle zuständig. Redaktionell wird ein Feature von einem Regisseur und einem Assistenten begleitet. In diesem Fall führten die Autoren selbst Regie.

Die im BR für Featureproduktionen übliche DAW ist *ProTools HD* der Firma *Avid*. Alle Arbeitsprozesse für solch umfangreiche Produktionen sind an ProTools angepasst und seit Jahren gängige Praxis.

Zunächst sollte das Feature in einem Zeitraum von 5 Tagen in Stereo produziert werden, wie es für eine Featureproduktion üblich ist. Anschließend sollte in 3 Tagen eine objektbasierte Fassung erstellt werden, die dann in unterschiedlichen dreidimensionalen Formaten wiedergegeben werden kann. Die für eine objektbasierte Produktion nötigen

DAW Erweiterungen gab es zum Produktionszeitpunkt noch nicht, daher wurde ein vom IRT entwickeltes Prototyp-Experimental-System verwendet - das Mischsystem EDDIE und der Renderer HOG. Beides läuft unter Linux auf einem externen Rechner, der über MADI mit ProTools verbunden wurde. In Abbildung 7.1 ist die Anbindung an die Infrastruktur von Studio 9 des BRs schematisch aufgeführt.

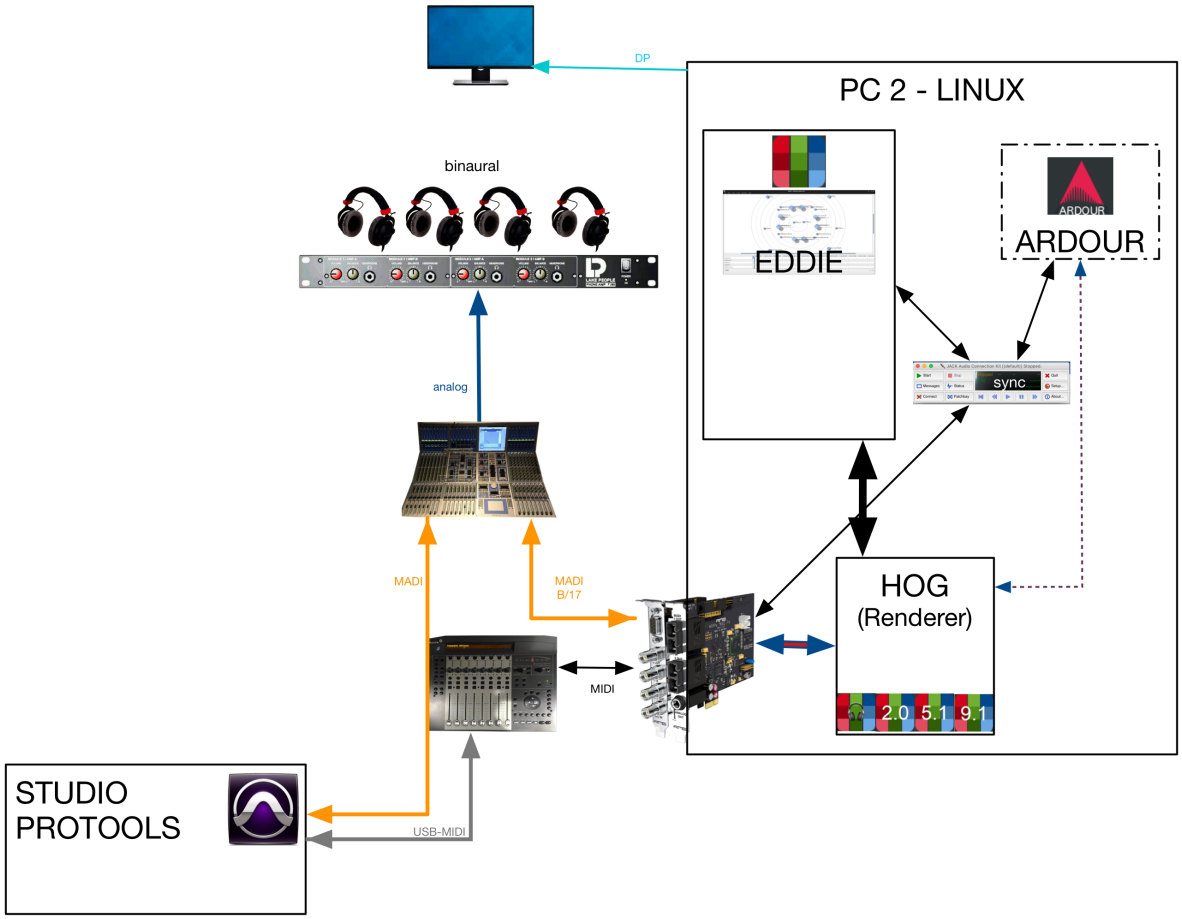

Abb. 7.1: EDDIE-HOG Anbindung an das Studio 9 des BRs

In EDDIE können Objekte rund um den Hörer herum platziert werden, was in Abbildung 7.2 zu sehen ist. Alle 32 möglichen Eingänge werden, wie man es aus DAWs kennt, als Spuren dargestellt. In der Abbildung 7.2 sind drei dieser Spuren als *Frosch 1*, *Frosch 2*, und *Frosch 3* zu erkennen. In den Spuren lässt sich nun über Keyframes definieren, wann ein Objekt anfängt und aufhört und an welcher Position im Raum es sich befinden soll. Mit der Maus lassen sich die Objekte, in der Abbildung als blaue Kreise dargestellt, rund um den Mittelpunkt, der den Hörer darstellt, platzieren. Durch

die Keyframes können die Objekte auch animiert werden. Befindet sich das zweite Keyframe an einer anderen Position als das erste, wird zwischen den beiden Keyframes eine Bewegung interpoliert.

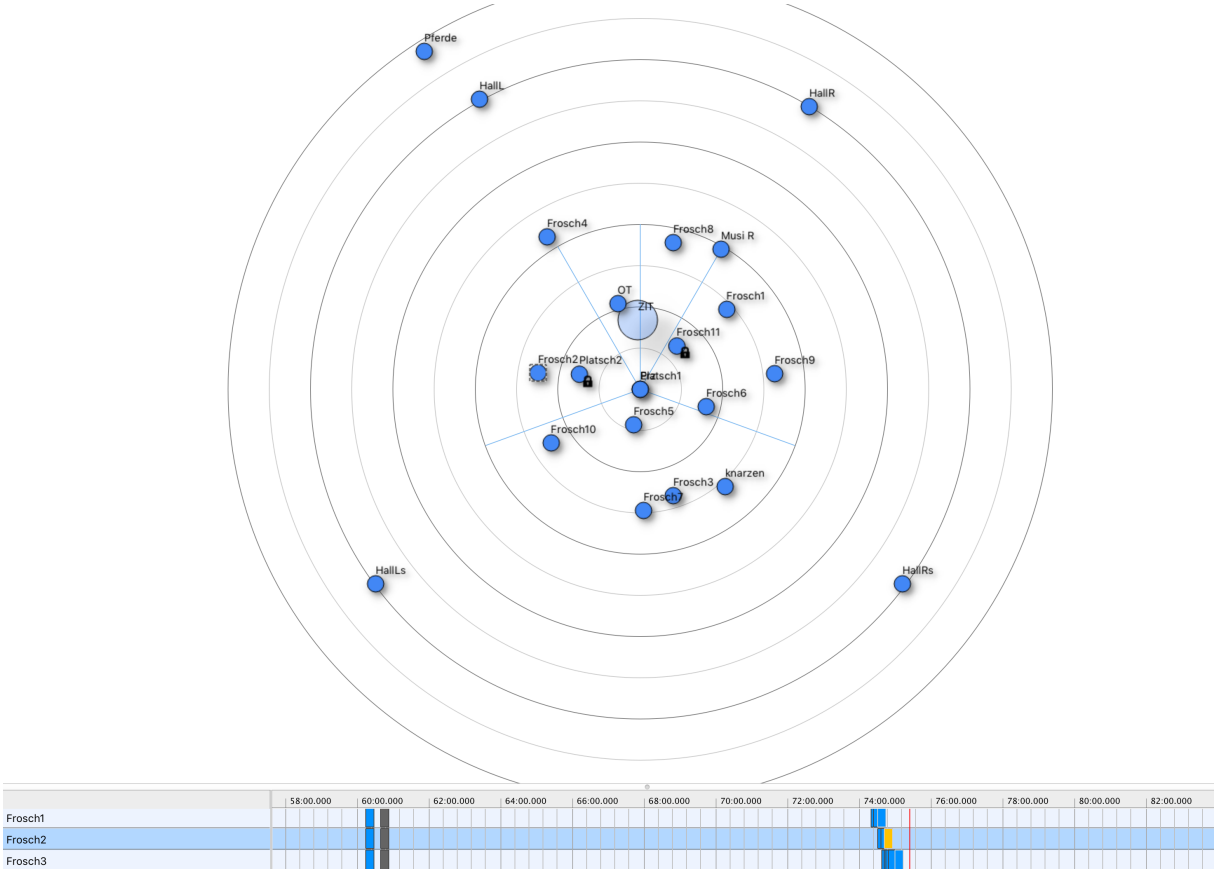

Abb. 7.2: Prototyp-Experimental-System EDDIE: hier die „Froschszene“ vom Ende des Features

HOG ist die Rendering Einheit. Die Mischung aus EDDIE kann in unterschiedliche Formate gerendert werden, wie beispielsweise in Binaural, Stereo oder 5.1.

Aufgrund dieser EDDIE-HOG Anbindung als Outboard-System und ProTools als Zuspieler, ist diese Produktion keine native objektbasierte Produktion, sondern eher ein Hybrid, da das vorhandene, nicht objektbasierte Produktionssystem auf diese Weise erweitert wurde.

Diese Produktionsweise war bereits in den Wochen zuvor anhand kleinerer Produktionen getestet worden. Es wurde z.B. ein Quartett mit unterschiedlichen Mikrofonen aufgenommen (Kunstkopf, AMBEO, herkömmliche Mikrofonierung mit Haupt- und Stützmikrofonen). Mit EDDIE wurde gemischt und diese Mischung mit HOG in unterschiedliche Formate gerendert. Die Ergebnisse wurden anschließend miteinander verglichen, etwa die gerenderte Binauralfassung mit der nativen Binauralfassung des Kunstkopfs.

Diese Quartettaufnahme dient ebenfalls als Beispiel für die Erprobung der neuen Möglichkeiten, die objektbasierte Produktionen dem Hörer bieten. Sie wird im Rahmen der Pilotphase 1 verwendet.

7.2 Stereoproduktion

Bei der Featureproduktion konnte auf die Erfahrungen aus der Quartettproduktion zurückgegriffen werden. Jedoch war dies die erste Produktion in solch einem Umfang. Schon im Vorfeld musste überlegt werden, wie die anschließende objektbasierte Mischung möglichst effizient umgesetzt werden kann. Eine Schwierigkeit, die es zu lösen galt, ist die Tatsache, dass EDDIE-HOG bisher 32 Eingangskanäle besitzt, die ProTools Session jedoch viel umfangreicher werden würde als nur 32 Spuren. Der Tontechniker dieses Features versuchte sich eine sinnvolle Arbeitsweise zu überlegen. Die Session musste klar strukturiert sein. Jeder Sprecher sollte eine eigene Audiospur bekommen, was allein schon 9 Spuren waren. Außerdem sollten alle zu verwendenden Effekte direkt aus dem selben ProTools-System kommen. Bei Feature- und Hörspielproduktionen ist es sonst üblich, ein weiteres ProTools-System nur als Zuspieler für Effekte zu verwenden.

Stereo-Hallräume sollten durch ein Mehrkanal-Hall-Plugin über eine Quad-Summe in 4 Kanäle umgewandelt und als statische Objekte an EDDIE-HOG geschickt werden.

Des Weiteren war angedacht, die einzelnen Mono- und Stereospuren über Post-Fader geschaltete Send-Wege dynamisch auf die 32 Ausgänge an EDDIE-HOG zu schicken. Anpassungen der Mischung sollten nur als Automation der Sends vorgenommen werden. Auf diese Weise sollte eine einzige ProTools Session erhalten bleiben, in der sich sowohl die Stereofassung als auch die objektbasierte Fassung befinden. Die Mischung in EDDIE-HOG sollte anschließend über einen Rückweg in ProTools aufgezeichnet werden.

Mit diesen Überlegungen im Hinterkopf verlief die Stereoproduktion weitestgehend reibungslos. Jedoch wuchs die Anzahl der Spuren auf 70 - 80 an, was weit mehr ist als sonst üblich.

7.3 Objektbasierte Produktion

In der nun folgenden objektbasierten Mischung mit EDDIE-HOG stellte sich schnell heraus, dass die bisherigen Überlegungen zwar gut, aber nicht ausreichend waren. Die Reduzierung der ProTools Audiospuren auf insgesamt nur 32 Spuren mit automatisierten Send-Wegen stellte sich als sehr komplex und damit unübersichtlich heraus.

Ein weiteres Problem war die Mischung aus Mono-, Stereo- und Quad-Spuren, da die Eingänge in EDDIE-HOG nur einkanalig sind.

Somit wurde die komplette ProTools Session umstrukturiert. Mehrere Audiospuren wurden zu einer Subgruppe zusammengefasst und fest auf einen EDDIE-Eingang geschickt, bzw. zwei Monoeingänge bei einer Stereo-Subgruppe. Beispielsweise wurde für alle Zitatoren eine Mono-Subgruppe gebildet, da diese nie gleichzeitig sprechen. Durch diese Neuordnung war ein effizienteres und übersichtlicheres Arbeiten möglich.

Während des Mischens in EDDIE wurde es nötig, nicht nur Lautstärkeanpassungen vorzunehmen, sondern auch weitere Parameter wie Hallanteile, Filter und dynamische Effekte zu ändern. Der ursprüngliche Plan, nur eine ProTools Session zu haben, konnte durch die Umstrukturierung und die zusätzlichen Änderungen jedoch nicht eingehalten werden. Die ProTools Session für die objektbasierte Fassung hatte am Ende ca. 110 Spuren.

Anfang Mai konnte diese Mischung im *3D Audio LAB* des IRT in München mit knapp 40 Surround-Lautsprechern angehört werden. Das Ergebnis erfüllte die Erwartungen, auch wenn an wenigen Stellen kleine Anpassungen notwendig waren.

Für Mitte Juni wurden noch weitere Produktionstermine angesetzt, in denen eine 15-minütige Kurzfassung des Features erstellt und zusätzliche Sprachaufnahmen für eine englische und eine französische Fassung gemacht wurden. Hierfür wurde die ProTools Session in ihrer Spuranzahl reduziert. Abbildung 7.3 zeigt einen Screenshot dieser Session, aus dem die Spuren ersichtlich sind.[45]

[45] Gespräch über die objektbasierte Produktion - mit W. Bleisteiner (2017)

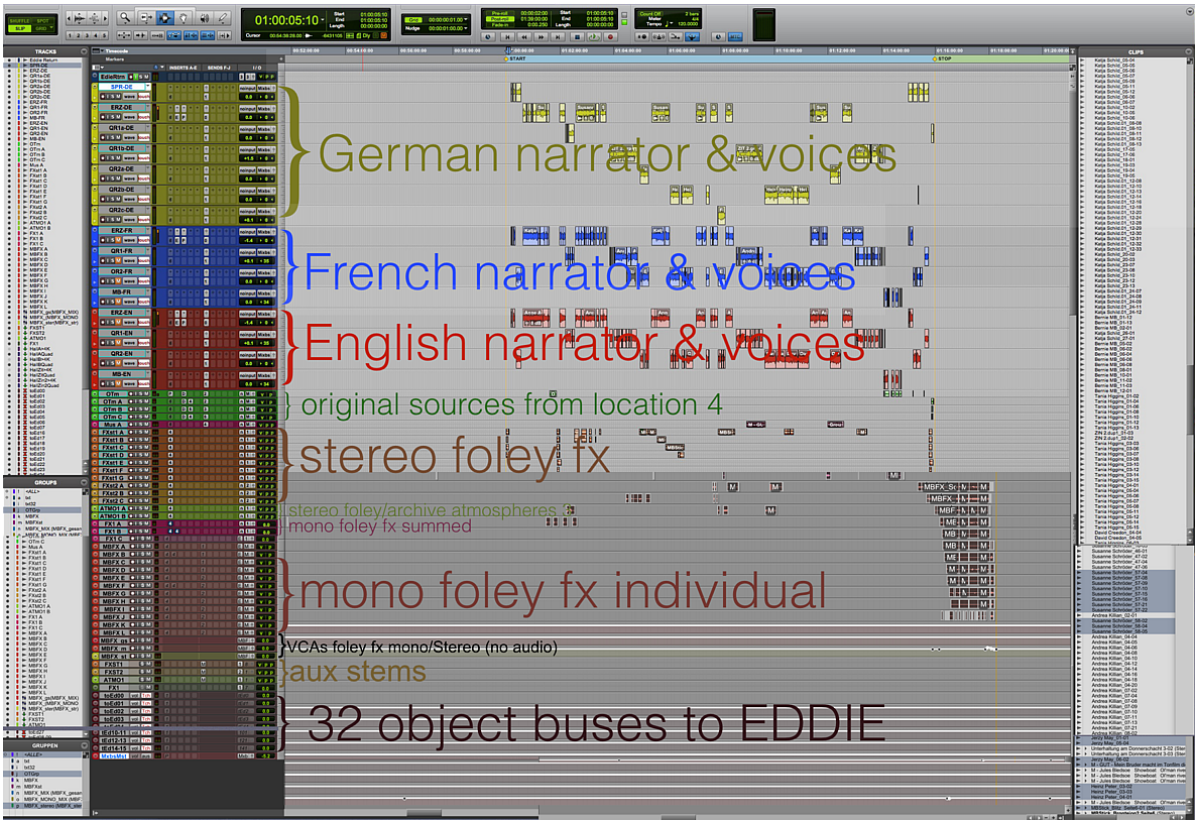

Abb. 7.3: ProTools Session der mehrsprachigen Kurzfassung des Features

7.4 Fazit

Die Erfahrung mit dieser Produktion zeigt, dass ein objektbasierter Arbeitsablauf auf diese Weise zunächst experimentell ist. Das EDDIE-HOG System ist nicht für einen regulären, professionellen Gebrauch gedacht und diente hier zunächst als „Proof-of-Concept". Mit dem System lassen sich gute Ergebnisse erzielen, aber es hat viele Einschränkungen und birgt auch einige Fehler. Manche davon ließen sich während der Produktion durch Entwickler des IRT beheben, indem sie den Programmcode von EDDIE-HOG bearbeiteten und Erweiterungen durchführten. Beispielsweise schrieben sie für die Kurzfassung des Features einen Code, der es ermöglichte, die Kürzungen sehr schnell in EDDIE zu übertragen, ohne die Objekte & Keyframes manuell verschieben zu müssen.[46]

Die jedoch größte Herausforderung dieser Produktion war aber die Umstellung in Objekten zu denken statt wie bisher in Spuren.

[46] Gespräch über die Erstellung der Kurzfassung - mit W. Bleisteiner (2017)

8 Sendung

BBC R&D hat ein IP-Studio gebaut, um zu untersuchen, welche Vorteile eine netzwerkbasierte Produktionsinfrastruktur bieten kann. Vernetzung über IP wird künftig sehr stark zunehmen und auch Audio kann zusammen mit Metadaten problemlos IP-basiert verschickt werden. Zwar ist es möglich die bisherige Produktionsinfrastruktur nachzubessern, jedoch ist dies mit erhöhtem Aufwand und erheblichen Mehrkosten verbunden. Somit wurde im IP-Studio von vornherein auf eine Anbindung über Ethernet gesetzt.

Aufnahme und Kontrolle
Einzelne Signalquellen werden erfasst und zu einer Komposition zusammengestellt, die die Radiosendung bildet. Diese Quellen können die Mikrofone im Sendestudio sein, Leitungen von Extern (z.B. aus einem Fußballstadion), Telefonanrufe, vorproduzierte Beiträge, Jingleelemente, Musik, usw. In einer bisherigen kanalbasierten Produktion laufen diese Signale in ein Mischpult, aus dem dann die fertige (Stereo-)Mischung kommt, die anschließend gesendet wird und so beim Hörer ankommt. Objektbasiert spricht man nun von einer Komposition, da die Signale getrennt bleiben. Anhand der zusätzlichen Metadaten berechnet der Renderer im Studio und beim Hörer die Mischung.

Statt eines Mischpults braucht es auch hierfür einen Controller, ob graphisch, physikalisch oder eine Kombination aus beidem. Hier werden die Metadaten während der Sendung eingestellt (Gain, Position, Vordergrund-/Hintergrunddefinition, usw.). Es gibt aber auch Informationen (Sendungsmetadaten), die bereits vor der Sendung über ein spezielles Interface zur Bearbeitung von Metadaten eingegeben werden können (Sendungstitel, Sendungsbeschreibung, Produktionsteam, Hinweise zur Homepage und/oder sozialen Netzwerken, usw.). Diese Informationen zu bearbeiten ist meist Aufgabe der für die Sendung verantwortlichen Redaktion. Gibt es ein Manuskript zu der Sendung, ist es auch möglich, dieses in den Metadaten mitzuübertragen. Gibt es kein Skript, ist angedacht, ein Live-Transkript durch einen Konverter erstellen zu lassen, der Sprache in Text umwandelt. Ein solcher Service von Google wird derzeit getestet.[47]

[47] vgl. *D2.2 Interim Reference Architecture Specification And Integration* (2017) S. 16f

Monitoring

Wie schon in der Vorproduktion ist es auch während des Sendens wichtig, die Signale überwachen, abhören und aussteuern zu können. Siehe hierfür Kapitel 6.4.

Aufzeichnung und Wiedergabe

Aktuelle Tools erlauben bisher nur die Aufzeichnung von Audio. Um aber auch die Metadaten zu erfassen und zu speichern, wird ein „Sequence Store“ verwendet, der Audio und Metadatenstrom in eine Datenbank zusammen mit einem Zeitstempel aufzeichnet.

Um den Sequence Store bedienten zu können, wird das von der BBC entwickelte Media Access API (Application Programming Interface) verwendet. Für Pilotphase 1 wurde das API dahingehend erweitert, dass es BW64 und ADM Dateien lesen, in einen Audio- und Metadaten-Stream konvertieren und in den Speicher importieren kann. Eine Erweiterung ist auch für den Export geplant. Diese soll das Exportieren von live aufgezeichnetem Material als BW64 und ADM ermöglichen.[48]

AES67

AES67 ist ein 2013 als „AES standard for audio application of networks - High-performance streaming audio-over-IP interoperability“ veröffentlichtes Format für IP-basierte Infrastruktur. An dessen Standardisierung war BBC R&D aktiv mit beteiligt. Es ermöglicht die Kompatibilität zwischen unterschiedlichen IP-basierten Audionetzwerksystemen (z.B. Dante, Ravenna, Livewire, u.a.). Dabei ist es selbst kein neues Protokoll, sondern es gibt lediglich an, wie existierende Protokolle miteinander arbeiten müssen, um ein funktionierendes System zu bilden.[49]

NMOS

In ein Sendestudio sind Geräte unterschiedlicher Hersteller integriert, die zusammen arbeiten können müssen. Networked Media Open Specifications (NMOS) sind eine Gruppe von Spezifikationen der Advanced Media Workflow Association (AMWA). Durch die Verwendung von Netzwerkverbindungen können die Geräte an ein gemeinsames System angeschlossen werden und leicht untereinander kommunizieren. Mit Hilfe der NMOS-Spezifikationen ist es möglich, Inhalte und Geräte in diesem Netzwerk zu identifizieren, sie in Beziehung zueinander zu setzen und zeitliche Informationen hinzuzufügen.[50]

[48] vgl. *D2.2 Interim Reference Architecture Specification And Integration* (2017) S. 18
[49] vgl. *D4.2 Interim Report Object-based Radio* (2017) S. 13
[50] vgl. *D3.4 Implementation And Documentation Of A Live Object-based Production Environment* (2016) S. 8 f

UMCP

Das Universal Media Composition Protocol (UMCP) ermöglicht das Zusammenfügen unterschiedlicher Medienquellen zu einem einzigen neuen Stück Medienoutput. Auch hier wird wieder von einer Komposition gesprochen. Denn UMCP ist dabei nur die Anleitung, die vorgibt, was zu welchem Zeitpunkt getan werden muss, damit das gewünschte Endprodukt dabei herauskommt.

Mehrere Anwender können sich dabei auf derselben Komposition anmelden und diese bearbeiten. So können etwa von einer Station aus Änderungen in Echtzeit eingegeben werden, die anschließend gleich vom Renderer umgesetzt werden können.[51]

[51] vgl. *D4.2 Interim Report Object-based Radio* (2017) S. 14-16

9 Pilotphase 1

Ziel der Pilotphase 1 im Rahmen von Orpheus ist es, eine vollständig objektbasierte Audioproduktionskette anhand der zuvor definierten Pilotarchitektur abzubilden. Hierin sollen alle Bereiche einer Radioproduktion eingebunden werden: Aufnahme, Vorproduktion, Sendung, Distribution und Wiedergabe.

Der erste Pilot bestand aus einer Live-Radiosendung und deren Ausstrahlung, die am 6.7.2017 im IP-Studio der BBC in London stattfand. Hier konnte die überlegte Architektur und die entwickelten Geräte getestet werden. Es sollte Aufschluss darüber gewonnen werden, welche Herangehensweisen funktionieren und welche nicht, welche Schwächen es zu beheben gilt und welche weiteren Forschungen und Entwicklungen notwendig sind.[52]

Die Pilotphase 1 gliedert sich insgesamt in drei Phasen. Dadurch soll die Komplexität des Implementierungs- und Integrierungsprozesses bewältigt und die Abwärtskompatibilität zur existierenden Infrastruktur gewährleistet werden.

9.1 Phase 1A - Live

Die Pilotsendung ist ein ca. 10-minütiges Hörspiel mit dem Titel „The Mermaid's Tears“ und wurde speziell für dieses Projekt entwickelt. Es spielt in der Wohnung einer Mutter, Lesley, die Besuch von zwei Polizisten, Dee und Bill, bekommt.

Mit diesem Hörspiel sollen die Möglichkeiten aufgezeigt werden, die sich durch eine objektbasierte Produktion bieten. Der Hörer kann den Charakteren folgen, wenn diese sich durch die Wohnung bewegen. Er kann auch die Perspektive ändern und zu einem anderen Charakter wechseln.

An geeigneten Stellen tauchen Fotos und Bilder auf, die die Handlung unterstützen, wie etwa der Text einer SMS.

[52] vgl. *D2.3 Interim Pilot Progress Report* (2017) S. 19

Abbildung 9.1 zeigt die Bewegungen der Charaktere durch die Wohnung. Kamerasymbole markieren die Stellen, an denen Bilder auftauchen sollen. Lautsprechersymbole kennzeichnen markante Geräusche.

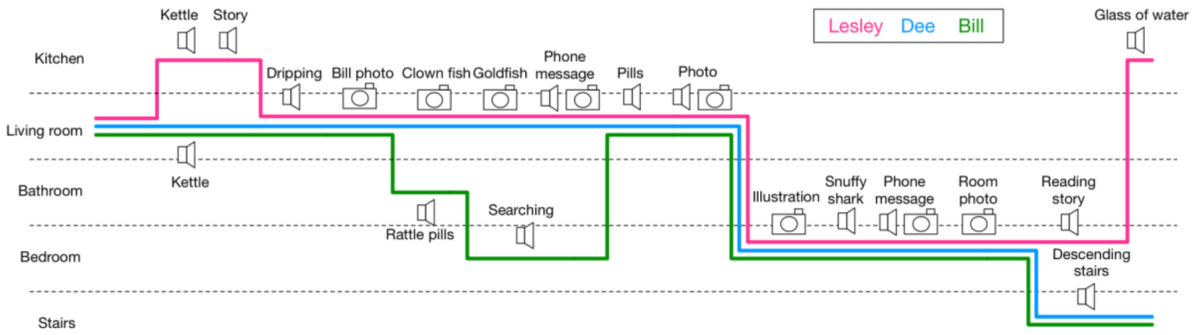

Abb. 9.1: Bewegung der Charaktere im Hörspiel

Für diese objektbasierte Produktion werden einige Geräte benötigt. Mit einem *Vorproduktionsinterface* werden Metadaten erfasst und gespeichert. Hierin können der Titel und eine Beschreibung der Sendung eingegeben sowie die einzelnen Audioquellen beschriftet werden.

Mittels des *Pipeline Configurators* wird das Routing der Audio- und Metadatenströme im IP-Studio gesteuert. Die Objekte können beispielsweise zu einem Renderer, einem Aufzeichnungsgerät und zu einem Pegelmesser geschickt werden.

Das *Live Produktionsinterface* entspricht dem Mischpult in herkömmlichen Radiostudios. Es steuert die Vorverstärkung, die Position und die Beschriftung der Audioquellen und spielt vorproduzierte Audioclips und Bilder ab. Des Weiteren können hiermit die Interaktivitätsmöglichkeiten der Audiostreams festgelegt werden.

Abschließend wird über das *Distributionskontrollinterface* die Verteilung von Audio und Metadaten vom Studio an das *Content Distribution Network* geregelt.

Das Live-Hörspiel aus Phase 1A wurde nur einem kleinen Kreis an Testhörern zur Verfügung gestellt, soll aber noch auf BBC Taster[53] einem größeren Publikum zugänglich gemacht werden.

Der Hörer hat die Möglichkeit, das Format zu wählen (Kopfhörer oder Stereolautsprecher), er kann entscheiden, welchem Charakter er beim Hörspiel folgen möchte und er sieht die Bilder, die während des Hörspiels abgespielt werden.

[53]Plattform zum Testen experimenteller Formate und Technologien: www.bbc.co.uk/taster

Da der MPEG-H Encoder zum Zeitpunkt der Sendung noch nicht vollständig in das IP-Studio eingebunden war, wurde daher eine andere Lösung mit einer AAC Encodierung gewählt, bei der zusätzlich die ADM Metadaten synchron mitübertragen werden. Der Empfang erfolgt mit einem HTML5 Web-Browser.[54]

Die Produktion des Hörspiels im IP-Studio funktionierte sehr gut. Das Hörspiel konnte komplett objektbasiert aufgezeichnet werden, wurde jedoch aufgrund technischer Herausforderungen den Testhörern noch nicht vorgespielt. Momentan sind die Projektpartner noch mit der Auswertung beschäftigt, weshalb hier an dieser Stelle noch keine weiteren Ergebnisse präsentiert werden können.[55]

9.2 Phase 1B - Wie live

Phase 1B stellt eine Zwischenlösung dar. Die Sendung aus Phase 1A wird offline in einen MPEG-H Stream encodiert und anschließend „wie live“ ausgestrahlt. Somit ist das Nutzererlebnis das selbe wie bei einer echten Live-Übertragung. Auf diese Weise kann auch die Distribution von MPEG-H über MPEG-DASH (siehe Kapitel 10.1) auf unterschiedlichen nicht Browser-basierten Plattformen (iOS App und AV-Receiver - siehe Kapitel 11) erprobt werden. Der Inhalt muss vom Produktionsformat BW64 + ADM in das Distributionsformat MPEG-H umgewandelt werden, was einige Hürden mit sich bringt.

Um weitere Funktionen aufzuzeigen, die objektbasiert möglich sind, wird Phase 1B um weitere Inhalte erweitert. Denkbar ist eine Klassikmusikaufnahme, die als HOA-Szene aufgezeichnet wird. Diese soll durch immersiven Klang beim Hörer den Eindruck erwecken, er wäre mitten im Geschehen und wäre selbst der Dirigent.

Ein weiteres Demonstrationsobjekt ist ein Fußballspiel, bei dem das Verhältnis zwischen Vordergrund (Kommentator) und Hintergrund (Stadionatmosphäre) eingestellt werden kann.

Eine weitere Produktion ist „The Turning Forest“. Hierbei handelt es sich um ein ca. 7-minütiges Hörspiel, das von der BBC in BW64 und ADM produziert wurde. Der BR hat hierfür eine deutsche Übersetzung des Erzählertextes erstellt. Der deutsche Erzähler wurde mit ProTools aufgenommen. Anhand einer Datei, die ausschließlich den englisch-

[54] vgl. *D2.3 Interim Pilot Progress Report* (2017) S. 8-14
[55] Gespräch über die Durchführung von Pilotphase 1 - mit W. Bleisteiner (2017)

sprachigen Erzähler enthielt, wurde der deutsche Text genau an den gleichen Stellen platziert wie der Englische. Als der Schnitt fertig war, wurde die Datei mit exakt der gleichen Länge abgespeichert, so dass die BBC beide Sprachvarianten problemlos in das Hörspiel einbauen konnte. Mit „The Turning Forest“ kann die Sprachauswahl im Rahmen des Pilots getestet werden.

Neben der Sprachwahl sollen noch weitere Funktionen erprobt werden, wie etwa das Zurückspringen in der Zeitleiste. Falls der Hörer inhaltlich etwas verpasst hat, kann er an den Anfang des aktuell laufenden Kapitels springen, um den Abschnitt nachzuhören. Anschließend kann er dem laufenden Programm weiter folgen und an die aktuell ausgestrahlte Position vorspringen.

Des Weiteren kann die Lautheit durch die Regulierung des Dynamikumfangs angepasst werden. Befindet sich der Hörer in einer lauten Umgebung, wird die Dynamik durch verstärkte Kompression eingeschränkt und die Lautheit somit erhöht.

Der Hörer hat die Möglichkeit, das Wiedergabeformat zu wählen. Zur Auswahl stehen Binaural für Kopfhörerwiedergabe, Stereo und Mehrkanal-Surround. Der Renderer erstellt die jeweils gewünschte Variante.

Phase 1B ist noch nicht für ein breites Publikum gedacht, sondern soll nur einem kleinen, von den Projektpartnern ausgewählten Kreis an Testhörern zur Verfügung stehen. Es sollen so viele Funktionen und Aspekte wie möglich evaluiert werden, wie etwa die Klangqualität, die Bedienoberflächen und die Klarheit des Konzepts.[56]

9.3 Phase 1C - Live encodiertes MPEG-H

In dieser Phase sind nun alle Komponenten vollständig integriert und die komplette Sendestrecke kann erprobt werden. Der MPEG-H Encoder ist in der Lage, die ADM Metadaten in die MPEG-H Datenstruktur umzuwandeln und in Echtzeit zu encodieren. Audio und Metadaten werden als objektbasiertes Format auf die Ausspielwege verteilt. Des Weiteren wird eine bereits gerenderte Version für bisherige, nicht objektbasierte Geräte, bereitgestellt. Eine tatsächliche Live Übertragung ist jetzt möglich.

Es soll zusätzlich noch weitere Inhalte geben. Diese werden anhand der Erfahrungen aus den Phasen 1A und 1B erst noch produziert.

[56] vgl. *D2.3 Interim Pilot Progress Report* (2017) S. 15-17

Die Funktionen in Phase 1C sind die Gleichen wie in Phase 1B. Zusätzlich soll eine Moderation die einzelnen Elemente der Sendung verbinden und dabei über das Projekt und die neue Technologie informieren.

In Phase 1C soll ein Fachpublikum angesprochen werden, das dem Programm mit der iOS App folgen kann. Die gewonnenen Erfahrungen aus den Phasen 1A, 1B und 1C sollen anschließend umgesetzt werden. Für die Veröffentlichung müssen die Bedienoberflächen klar verständlich und so einfach wie möglich gestaltet sein, da die Mehrheit der Hörer noch keine Erfahrungen im Bereich dieser neuen Audiotechnologie hat.

Der AV-Receiver von Trinnov wird zum Zeitpunkt von Phase 1C noch nicht im Handel erhältlich sein. Gegebenenfalls wird lediglich bei Fachmessen ein Prototyp ausgestellt werden.

Es ist geplant, Phase 1C bei der *International Broadcasting Convention* (IBC) im September 2017 in Amsterdam zu präsentieren.[57]

[57] vgl. *D2.3 Interim Pilot Progress Report* (2017) S. 18

10 Distribution

Wie bereits in Kapitel 5 erläutert, wird innerhalb von Orpheus für die Speicherung von Audio und Metadaten während der Produktion BW64 und ADM verwendet. Für die Distribution zum Konsumenten wurden zwei datenreduzierte Formate gewählt. Diese sind zum einen AAC (+ ADM Metadaten) für den Empfang mit einem HTML5 Web-Browser und zum anderen MPEG-H für den AV-Receiver und die iOS App. Beide Übertragungsformate können gleichzeitig verbreitet werden. Sie nutzen das Internet als Übertragungskanal und MPEG-DASH als Übertragungsprotokoll. [58]

10.1 MPEG-DASH

DASH steht für „Dynamic Adaptive Streaming over HTTP“ und ist ein von MPEG entwickelter internationaler Standard. Es ermöglicht eine adaptive Qualität beim Streaming von Medieninhalten über das Internet, indem es die Bitrate dynamisch an schwankende Netzwerkbedingungen anpasst.

Der Inhalt wird in kleine Sequenzen, in „Segmente“, aufgeteilt, wovon jedes zwischen 2 und 10 Sekunden Wiedergabezeit enthält. Als Container für das Audiomaterial können unterschiedliche Codecs verwendet werden (z.B. mp4, AAC oder MPEG-H). Wird etwa mp4 verwendet, entspricht jedes Segment einem mp4-Fragment. Der Inhalt wird in unterschiedlichen Datenraten codiert und ist somit anpassungsfähig. Der DASH-Client kann beim Herunterladen oder Wiedergeben automatisch die geeignetste Datenrate wählen. Bei schwankenden Netzwerkbedingungen werden somit Aussetzer verhindert.

Da MPEG-DASH das Übertragungsprotokoll HTTP/TCP verwendet, kann es problemlos in existierende Infrastruktur integriert und über das Internet übertragen werden. Somit ist es für eine hohe Anzahl an Nutzern leicht zugänglich.[59]

[58] vgl. *D4.2 Interim Report Object-based Radio* (2017) S. 20 und S. 25
[59] vgl. *D4.2 Interim Report Object-based Radio* (2017) S. 25 f

10.2 AAC & ADM

AAC (Advanced Audio Coding) ist ein von MPEG entwickelter und von ISO/IEC standardisierter Audiocodec, der auch Mehrkanalcodierung unterstützt. AAC ist datenrateneffizient, streamingfähig, weit verbreitetet und wird beispielsweise in vielen Mediatheken verwendet. Der Codec an sich ist kanalbasiert, aber mit einer Funktionserweiterung von JavaScript ist ein objektbasiertes Streaming der ADM Metadaten zusammen mit JavaScript Object Notation (JSON) möglich. Dies kann mit dem WebAudio API im Browser gerendert werden. So können maximal 32 Objekte als Kanäle übertragen werden.[60]

10.3 MPEG-H

MPEG-H 3D Audio ist von Anfang an daraufhin entwickelt worden, objektbasiert zu arbeiten und Audio und Metadaten zu enthalten. Zudem ist es bitrateneffizienter und weniger komplex als die Browser-basierte Lösung.

MPEG-H ist in erster Linie ein Audiokompressionsformat, das als *ISO/IEC 23008-3 (MPEG-H Part 3)* spezifiziert ist. Es wurde speziell dafür entwickelt, die Anforderungen an Next Generation Audio (NGA), Audio der nächsten Generation, zu erfüllen. Daher kann es kanalbasierte, szenenbasierte (HOA) und objektbasierte Signale transportieren. Dies bietet dem Nutzer die Möglichkeiten der Personalisierbarkeit, Interaktivität und Immersion.

MPEG-H unterstützt bis zu 128 Kanäle und 128 Objekte, die auf maximal 64 Lautsprechern wiedergegeben werden können. Für Radio- und TV-Anwendungen ist es aber nicht notwendig, diesen vollen Umfang auszunutzen. Es gibt Profile, die die Komplexität auf bestimmte Werte beschränkt, die für Rundfunkanwendungen sinnvoll sind.
Im Rahmen von Orpheus wird *MPEG-H 3D Audio Low Complexity Profile at Level 3* verwendet. Hierin wird beispielsweise die Abtastrate auf 48 kHz beschränkt, obwohl 96 kHz möglich wären.

Die Metadatenmodelle von MPEG-H und ADM sind zwar ähnlich, weichen aber dennoch im Detail stark voneinander ab. Beispielsweise ist die Anzahl an Objekten, die

[60] vgl. *D4.2 Interim Report Object-based Radio* (2017) S. 26 ff

mit ADM gleichzeitig gerendert werden können, theoretisch unbegrenzt, bei MPEG-H ist diese Anzahl auf 128 Objekte beschränkt. In der Praxis sind es bisher noch weniger. 32 Objekte können verarbeitet werden, jedoch je nach Renderer nur 16 gleichzeitig. Zu dieser Einschränkung kommt es zum Beispiel bei der iOS App, da die Rechenleistung des Smartphones aktuell nicht ausreicht. Somit ist es wichtig, eine gute Übersetzung von ADM Metadaten hin zu MPEG-H Metadaten zu entwickeln.[61]

[61] vgl. *D4.2 Interim Report Object-based Radio* (2017) S. 21 und S. 29 ff

11 Wiedergabe

Für die Wiedergabe sind innerhalb von Arbeitspaket WP5 maßgeblich Trinnov und Elephantcandy verantwortlich. Die Herausforderungen hierbei sind einerseits technische (Wie gelangt das objektbasierte Audio zum Endgerät und wie wird es gerendert?) und andererseits ergonomische sowie konzeptionelle. Die neuen Möglichkeiten, die sich durch eine objektbasierte Produktion bieten, müssen einfach und klar dargestellt werden, damit der Hörer das neue Angebot gut nutzen kann.[62]

Web Browser: AAC + ADM

Weil MPEG-H noch nicht auf herkömmlichen Browsern läuft, wird ein AAC Audiostream verwendet, der zusätzlich die nötigen Metadaten als ADM enthält. Als Renderer wird der aufkommende WebAudio API Standard genutzt.

Wie Tests zeigten, ist es möglich, MPEG-H in einen Browser zu implementieren, doch aufgrund lizenzrechtlicher Belange vom Fraunhofer IIS wird dieser Ansatz derzeit nicht weiterverfolgt.[63]

Trinnov Hardware AV-Receiver: MPEG-H

Der *Trinnov Altitude32* AV-Receiver ist in der Lage, mit unterschiedlichen Audioformaten umzugehen, sowohl objekt- oder kanalbasiert, als auch 2D oder 3D. Es können bis zu 32 Lautsprecher angeschlossen werden. Der AV-Receiver hat ein integriertes Einmesssystem. Durch die Messung einer Impulsantwort werden die Hallanteile eines Raums erfasst und können somit beim Rendern eines lautsprecherbasierten Abhörformats entsprechend berücksichtigt werden.

Der MPEG-DASH Receiver und MPEG-H Decoder & Renderer Library wurden integriert. Erste Tests waren erfolgreich, die Arbeit dauert jedoch noch an.[64] [65]

[62] vgl. *D5.2 Implementation And Documentation Of Intermediate Version Of Object-based Renderers And User Interfaces* (2017) S. 7

[63] vgl. *D5.2 Implementation And Documentation Of Intermediate Version Of Object-based Renderers And User Interfaces* (2017) S. 11 f

[64] vgl. Homepage Trinnov (2017)

[65] vgl. *D5.2 Implementation And Documentation Of Intermediate Version Of Object-based Renderers And*

iOS App: MPEG-H

Mittels MPEG-DASH soll die App den in MPEG-H encodierten objektbasierten Audiostream erhalten. Derzeit können mp4 und AAC Audiostreams empfangen und wiedergegeben werden. Der mp4 Container wird auch für MPEG-H codierte Streams verwendet. An der Integrierung von MPEG-DASH Streams mit dem MPEG-H Decoder wird noch gearbeitet.

Bisher wurde hauptsächlich die iOS App gestaltet und entwickelt. Später können das Konzept, das Design und weitere Ergebnisse dieses Prototyps auf andere Plattformen angewendet werden. [66]

Abbildung 11.1 zeigt das Design des aktuellen Prototyps der iOS App für iPhone.

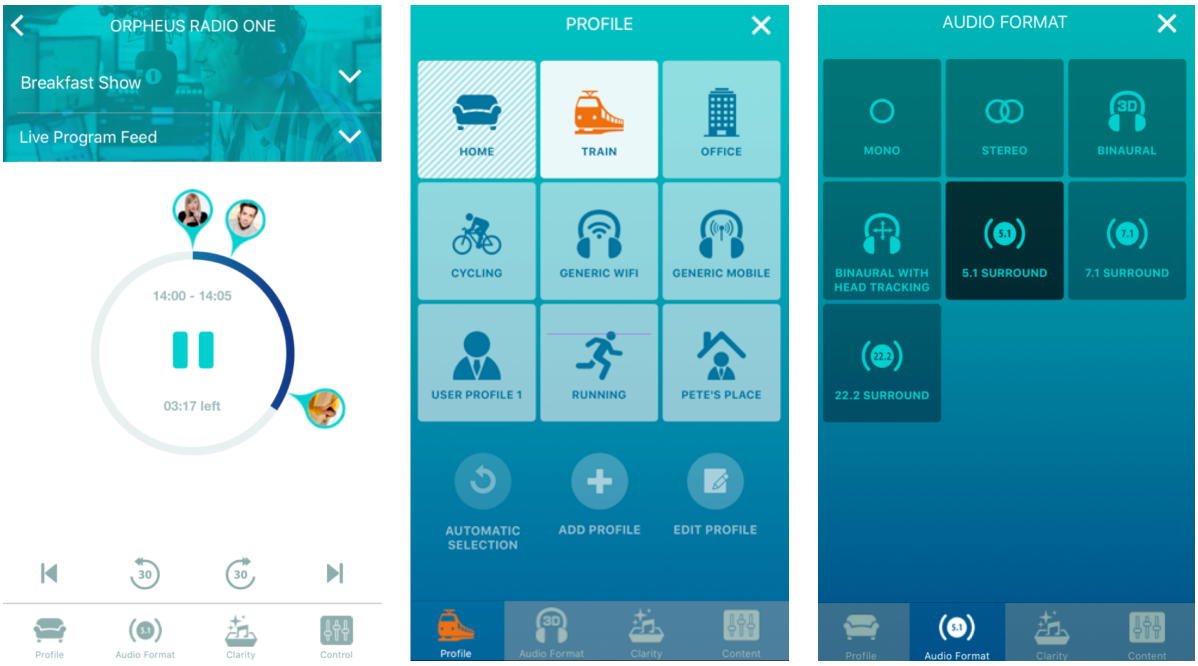

Abb. 11.1: Prototyp iOS App: Navigieren in der Zeitleiste, Auswahl eines Profils, Auswahl des Abhörformats

User Interfaces (2017) S. 14 f

[66] vgl. *D5.2 Implementation And Documentation Of Intermediate Version Of Object-based Renderers And User Interfaces* (2017) S. 16 f

12 Fazit

Etwas mehr als die Hälfte der Projektlaufzeit von Orpheus ist bereits vergangen. Viele Dinge, die anfangs nur als Idee auf dem Papier existierten, wurden umgesetzt. Es sind Produktionstools entstanden, die das Aufnehmen, Analysieren, Bearbeiten und Mischen von objektbasiertem Audio ermöglichen. Die DAW Sequoia von Magix wurde weiterentwickelt. Sie kann nun mit ADM Metadaten umgehen, ermöglicht objektbasiertes Mischen und sie hat einen Renderer integriert. Es wurden Inhalte produziert, die sich zum Erproben der entwickelten Architektur eignen und bei deren Erstellung viele neue Erkenntnisse gewonnen wurden. Das vollständig IP-basierte Ausspielen im Radiostudio der BBC mit objektbasierten Funktionen wurde in Pilotphase 1 erfolgreich getestet. Objektbasierte Inhalte können als AAC + ADM und als MPEG-H Streams zur Verfügung gestellt werden. Das Dekodieren und Wiedergeben objektbasierter Inhalte wurde in Web-Browser, App und AV-Receiver integriert und erfolgreich getestet.

Im September 2017 sollen viele der bisherigen Ergebnisse auf der IBC in Amsterdam einem breiten Fachpublikum präsentiert werden.

Eine große Aufgabe für die kommenden Projektmonate ist die Definition, Ausarbeitung und Durchführung der Pilotphase 2. Es sollen Möglichkeiten der Interaktion und Funktionen für nicht-live Inhalte entstehen. Es werden zeitbasierte Funktionen getestet und es wird überprüft, wie Sendungen an die Wünsche der Hörer angepasst werden können.

Die Entwicklung und Optimierung von Tools und Algorithmen ist derzeit noch im Gange und setzt sich über die nächsten Monate fort.

Gegen Ende des Projekts stehen diverse Tests an. Die entwickelten Tools, die iOS App und auch die Erweiterungen von Sequoia werden daraufhin geprüft, ob alle technischen Funktionen einwandfrei ausführbar und die graphischen Oberflächen anwenderfreundlich gestaltet sind.

Abschließend werden die Erkenntnisse aus der gesamten Projektlaufzeit in die Definition der Referenzarchitektur einfließen und diese zusammen mit den Neuentwicklungen

bei entsprechenden Standardisierungsgremien, wie der ITU und der EBU, vorgelegt. Die aus Orpheus gewonnenen Erkenntnisse und Entwicklungen müssen nach Abschluss des Projekts an andere Rundfunkanstalten herangetragen werden, so dass diesen ein Umstieg auf objektbasiertes Produzieren möglich ist und somit der Übergang zu einer neuen Zukunft der Rundfunklandschaft eingeläutet werden kann.

Diese Vision könnte frühestens in 5 Jahren Realität werden, wenn Rundfunkanbieter von den Vorteilen objektbasierten Produzierens überzeugt werden können und sukzessive auf die neue Technologie umrüsten.

Dennoch bleibt eine Problematik, die es zu lösen gilt. Bei vielen Innovationen im Rundfunkbereich hat sich gezeigt, dass Rundfunkanstalten neue Technologien relativ zögerlich einführen. Sie wissen nicht, ob und wie das Publikum diese annimmt und scheuen daher die eventuell damit verbundenen neuen Arbeitsabläufe. Hersteller von Endgeräten warten jedoch auf Inhalte, bevor sie bereit sind, neue Geräte zu bauen. Die hierfür nötigen großen Investitionen stellen finanzielle Risiken dar.

Als Tontechnikerin beim Radio bleiben nach eingehender Beschäftigung mit dem Thema des „objektbasierten Rundfunks“ einige Bedenken. Den gesamten Produktionsablauf so wie in Orpheus dargestellt auf objektbasiert umzustellen, wird eine sehr große Aufgabe sein. Objektbasierte Produktionen werden nicht über Nacht eingeführt, sondern dieser Prozess wird sich über einen Zeitraum von mehreren Jahren hinziehen. Der Umfang dieser Umstellung ist in etwa vergleichbar mit dem Wechsel von Analog zu Digital. Das verdeutlicht, wie anpassungsfähig und offen für Veränderungen Medienschaffende sein müssen.

Bisher sind noch wenige Erfahrungen auf dem Gebiet der objektbasierten Rundfunkproduktion vorhanden. Diese werden erst im Zuge von Orpheus nach und nach gesammelt. Mitarbeiter müssen geschult und diese Erfahrungen weitergegeben werden. Bis die neuen Arbeitsabläufe anschließend zur Routine geworden sind, wird es einige Zeit dauern.

Ein weiterer bedenklicher Askept bei objektbasierten Produktionen ist, dass der Hörer die Möglichkeit hat in die Mischung einzugreifen und diese zu verändern. Daher ist es nötig in der Produktion genau zu definieren, welche Parameter vom Hörer um welchen Betrag verändert werden können, so dass das Erstellen einer komplett neuen Mischung oder gar das Extrahieren einzelner Audiosignale ausgeschlossen ist.

Unklar und momentan nicht absehbar ist, inwieweit Avid und andere DAW-Hersteller mitziehen und ihre DAWs, wie etwa ProTools, weiterentwickeln werden. Derzeit ist ProTools in der Lage, ADM Metadaten zu lesen. Ebenso ist objektbasiertes Mischen möglich. Doch dies funktioniert nur mit der Integrierung von Dolby Atmos. ProTools ist in der Rundfunklandschaft weit verbreitet und ist beim BR die DAW, die unter anderem zum Produzieren von Hörspielen und Features verwendet wird. Sequoia ist dagegen überwiegend in der (Klassik-)Musikproduktion im Einsatz. Es wird sich zeigen, ob sich das ändern wird und künftig Hörspiele und Feature in Sequoia produziert werden oder ob ProTools auch insofern weiterentwickelt werden wird, dass es in der gleichen Art und Weise objektbasiert eingesetzt werden kann, wie Sequoia. Ob die Abhängigkeit von Dolby gelöst und die Verwendung offener Standards unterstützt werden wird, bleibt abzuwarten.

Für die Hörer wird es ebenfalls eine Umstellung geben, schließlich müssen sie über die neue Technologie und deren neue Möglichkeiten aufgeklärt werden. Doch durch einfache und intuitiv gestaltete Bedienoberflächen von App und Browser, fördert dies die Anwenderfreundlichkeit und neue Funktionen können leicht genutzt werden. Der Hörer hat es somit leichter, den erheblichen Mehrwert zu erkennen. Durch die einfache Verbreitung über das Internet sind die Inhalte fast Jedem leicht zugänglich. Ebenso fallen durch die Installation der App auf dem Smartphone oder die Nutzung des Browsers keine zusätzlichen Kosten an.

Der objektbasierte Ansatz löst auch ein Problem, mit dem Rundfunkanbieter häufig konfrontiert werden: die Beschwerden vieler Hörer, die Musik sei zu laut und die Sprache zu leise. Objektbasiert haben die Hörer nun die Möglichkeit, das Mischungsverhältnis zwischen Musik und Sprache selbst einzustellen und an ihr persönliches Empfinden anzupassen.

Objektbasiertes Produzieren erscheint somit als ein innovativer, konsequenter und universeller Schritt in die Zukunft der Medienproduktion, in der auf die vielfältigen Anforderungen der Konsumenten durch individuelle Anpassungsmöglichkeiten eingegangen wird.

Innerhalb von Orpheus geht es zunächst nur darum zu untersuchen, was objektbasiert alles denkbar und technisch möglich ist. Wie viele von diesen Ideen und Möglichkeiten künftig dann tatsächlich in der Realität umgesetzt werden, wird sich zeigen.

Danksagung

An dieser Stelle möchte ich mich herzlich bei meinen beiden Betreuern bedanken, Herrn Prof. Oliver Curdt und Herrn Werner Bleisteiner, die mich beim Erstellen dieser Arbeit unterstützt haben und mir immer mit Rat & Tat und einem offenen Ohr zur Seite standen. Herrn Prof. Curdt möchte ich überdies herzlich für sein Engagement im Tonbereich während meines gesamten Studiums danken.

Für ihre Unterstützung beim Erstellen dieser Arbeit geht ein großer Dank an Katrin Roth, Eva Demmelhuber, Jens Kehne und Tobias Ricken.

Ein besonderer Dank gilt meinen Eltern, die mich während meines Studiums immer unterstützt haben und mir auch dabei halfen, meinen großen Traum vom Auslandsjahr in Finnland zu verwirklichen.
Auch meinen Großeltern möchte ich für ihre Unterstützung während meines Studiums danken.

Ein weiterer Dank geht an meine Arbeitskollegen beim BR, insbesondere an Elisabeth Weininger, Helmut Kormann, Winfried Messmer, Fabian Zweck und Helge Schwarz.

Herzlich danken möchte ich meiner lieben Münchner WG, Sarah & Matthias Brockhaus und Thomas Maierhofer, die mich mein ganzes Studium über unterstützt haben und mich bei sich beherbergten, wenn ich in München war. Ebenso auch ein riesiges Dankeschön an Ursula & Martin Leitzinger. Ihr ward mir stets eine große Hilfe.

Und zuletzt gilt mein herzlichster Dank noch Sophia Paarmann, Birgit Kleinschrot & Martin Paarmann für ihre Hilfe und Unterstützung in den vergangenen Monaten.

Abbildungsverzeichnis

2.1 Das 5.1 Format im Kino - aus Dolby Laboratories (2014) S. 5 10
2.2 Kanalbasierte Produktion für eine Stereomischung - von Homepage IRT (2016) 13
2.3 Objektbasierte Produktion - von Homepage IRT (2016) 13

3.1 Grafische Übersicht der Projektstruktur - aus *D2.2 Interim Reference Architecture Specification And Integration* (2017) S. 8 19

4.1 Grafische Übersicht über die Makroblöcke der Pilotarchitektur - aus *D2.3 Interim Pilot Progress Report* (2017) S. 7 27

6.1 Zylindrisches Mikrofonarray - aus *D3.3 Object-based Capture* (2017) S. 10 35
6.2 Kugelmikrofonarray *EigenMike* - von Homepage mh acoustics (2017) . . 35
6.3 HoaScope Plugin - aus *D3.3 Object-based Capture* (2017) S. 19 36
6.4 Sound Field Editor Plugin - aus *D3.3 Object-based Capture* (2017) S. 21 . 36

7.1 EDDIE-HOG Anbindung an das Studio 9 des BRs - Interne Dokumentation von W. Bleisteiner (BR) 40
7.2 Prototyp-Experimental-System EDDIE - Interne Dokumentation von W. Bleisteiner (BR) 41
7.3 ProTools Session der mehrsprachigen Kurzfassung des Features - Interne Dokumentation von W. Bleisteiner (BR) 44

9.1 Bewegung der Charaktere im Hörspiel - aus *D2.3 Interim Pilot Progress Report* (2017) S. 10 49

11.1 Prototyp iOS App - aus *D5.2 Implementation And Documentation Of Intermediate Version Of Object-based Renderers And User Interfaces* (2017) S. 18 57

Quellenverzeichnis

b<>com (2017). Homepage.
URL: *https://b-com.com/en/institute/about-bcom (Online-Abfrage: 28. Mai 2017)*

Bleisteiner, W. (2017). Auskünfte zum aktuellen Projektstand in persönlichen Gesprächen und Telefonaten zwischen März und August 2017.

D2.2 Interim Reference Architecture Specification And Integration (2017). Public Deliverable. Orpheus Konsortium (Weitnauer, M. - IRT).
URL: *https://orpheus-audio.eu/public-deliverables/ (Online-Abfrage: 10. Juni 2017)*

D2.3 Interim Pilot Progress Report (2017). Public Deliverable. Orpheus Konsortium (Baume, C. - BBC).
URL: *https://orpheus-audio.eu/public-deliverables/ (Online-Abfrage: 10. Juni 2017)*

D3.2 Implementation And Documentation Of Reverberation For Object-based Broadcast (2016). Public Deliverable. Orpheus Konsortium (Noisternig, M. - Ircam).
URL: *https://orpheus-audio.eu/public-deliverables/ (Online-Abfrage: 8. Juli 2017)*

D3.3 Object-based Capture (2017). Public Deliverable. Orpheus Konsortium (Epain, N. - b<>com).
URL: *https://orpheus-audio.eu/public-deliverables/ (Online-Abfrage: 15. Juli2017)*

D3.4 Implementation And Documentation Of A Live Object-based Production Environment (2016). Public Deliverable. Orpheus Konsortium (Baume, C. - BBC).
URL: *https://orpheus-audio.eu/public-deliverables/ (Online-Abfrage: 29. Mai 2017)*

D3.5 Specification And Implementation Of Reference Audio Processing For The Use In Content Creation And Consumption Based On Novel Broadcast Auality Standards (2017). Public Deliverable. Orpheus Konsortium (Mason, A. - BBC).
URL: *https://orpheus-audio.eu/public-deliverables/ (Online-Abfrage: 14. Juni 2017)*

D4.2 Interim Report Object-based Radio (2017). Public Deliverable. Orpheus Konsortium (Silzle, A. - Fraunhofer IIS).
URL: *https://orpheus-audio.eu/public-deliverables/ (Online-Abfrage: 17. Mai 2017)*

D5.2 Implementation And Documentation Of Intermediate Version Of Object-based Renderers And User Interfaces (2017). Public Deliverable. Orpheus Konsortium (Duval, B. - Trinnov; Bogaards, N. - Elephantcandy).
URL: *https://orpheus-audio.eu/public-deliverables/ (Online-Abfrage: 24. Mai 2017)*

D5.3 Document On Methodology Of Evaluation Of User Experience (2017). Public Deliverable. Orpheus Konsortium (Bleisteiner, W. - BR).
URL: *https://orpheus-audio.eu/public-deliverables/ (Online-Abfrage: 5. Juli 2017)*

D6.1 Dissemination Plan And Innovation Roadmap (2016). Public Deliverable. Orpheus Konsortium (Gupta, M. - Eurescom).
URL: *https://orpheus-audio.eu/public-deliverables/ (Online-Abfrage: 13. Mai 2017)*

Dolby Laboratories (2014). Dolby Atmos Next-Generation Audio for Cinema, White Paper.
URL: *https://www.dolby.com/us/en/technologies/dolby-atmos/dolby-atmos-next-generation-audio-for-cinema-white-paper.pdf (Online-Abfrage: 25. Mai 2017)*

Eurescom (2017). Homepage.
URL: *https://www.eurescom.eu/home.html (Online-Abfrage: 28. Mai 2017)*

Fraunhofer IIS (2017). Homepage.
URL: *https://www.iis.fraunhofer.de/de/ff/amm/for/forschaudiomulti.html (Online-Abfrage: 28. Mai 2017)*

IRT (2016). Homepage.
URL: *https://www.irt.de/fileadmin/media/downloads/veranstaltungen/2016/TMT_VR/IRT-Thementag_VR_Michael_Weitnauer.pdf (Online-Abfrage: 13. Juli 2017)*

Magix (2017). Homepage.
URL: *http://www.magix.com/de/unternehmen/ (Online-Abfrage: 28. Mai 2017)*

mh acoustics (2017). Homepage.
URL: *https://mhacoustics.com/products#eigenmike1 (Online-Abfrage: 14. Juli 2017)*

Orpheus (2017). „Object-Based Audio“ Seminar der EBU, Homepage.
URL: *https://orpheus-audio.eu/ebu-oba-2017/ (Online-Abfrage: 11. Juli 2017)*

Trinnov (2017). Altitude 32 mit integriertem Einmesssystem, Homepage.
URL: *http://www.trinnov.com/products/home-theater/altitude32/introduction-altitude32/?lang=en_us (Online-Abfrage: 9. August 2017)*

Wikipedia - Ambisonics (2017). Homepage.
URL: *https://en.wikipedia.org/wiki/Ambisonics#Higher-order_Ambisonics (Online-Abfrage: 9. August 2017)*

Anhang

Auf dem beigefügten Datenträger befinden sich:

- die vorliegende Arbeit als PDF-Datei
- die Langfassung des Features „Donnerschacht und Frösche in Schnapsgläsern - Die Kunst des Geräuschemachens“ (in Binaural als mp3-Datei)
- die deutschsprachige Kurzfassung des Features (in Binaural als mp4-Datei)
- die englischsprachige Kurzfassung des Features (in Binaural als mp4-Datei)
- die französischsprachige Kurzfassung des Features (in Binaural als mp4-Datei)