

Objektbasiertes Audio

Seminararbeit

im Studiengang
Audiovisuelle Medien

vorgelegt von

Alexander Weller

Matr.-Nr.: 41569

am 25. Februar 2021

an der Hochschule der Medien Stuttgart

Erstprüfer/in:

Zweitprüfer/in:

Inhaltsverzeichnis

Inhaltsverzeichnis	2
Abkürzungsverzeichnis	3
1 Einleitung	4
2 Kanal-, Szenen- und Objektbasiertes Audio.....	5
2.1 Kanalbasiertes Audio.....	5
2.2 Szenenbasiertes Audio.....	6
2.3 Objektbasiertes Audio	8
2.4 Hybride Ansätze	9
3 Audiocodecs.....	10
3.1 MPEG-H.....	10
3.2 AC-4	12
4 Nutzungsmöglichkeiten.....	13
4.1 Orpheus.....	13
4.2 BBC	13
4.3 S3A.....	14
5 Beispielhafter objektbasierter Workflow mit der EAR Production Suite... 	15
5.1 Die EAR Production Suite.....	15
5.2 Definition des Beispielprojekts	15
5.3 Arbeitsschritte.....	15
5.3.1 Tracksetup	15
5.3.2 Routingsetup.....	16
5.3.3 Plugin Setup.....	17
5.3.4 Automation und Interaktivität	20
5.3.5 Export	21
6 Fazit	22
Quellenverzeichnis	23
Abbildungsverzeichnis.....	25

Abkürzungsverzeichnis

HdM	Hochschule der Medien
NGA	Next Generation Audio
ADM	Audio Definition Model
EAR	EBU ADM Renderer
EPS	EAR Production Suite
OBA	Object Based Audio
CBA	Channel Based Audio
SBA	Scene Based Audio
HOA	Higher Order Ambisonics
USAC	Unified Speech and Audio Coding
DAW	Digital Audio Workstation
VST	Virtual Studio Technology
SAOC	Spatial Audio Object Coding

1 Einleitung

Audioproduktionen und die Art, wie Medien mit Audioinhalt konsumiert werden, hat bereits mehrere Iterationen durchlaufen. In den Anfängen wurde monophon produziert und konsumiert. Um die Immersion zu erhöhen folgten Stereo und Surroundsetups wie 5.1. Durch die zusätzlichen Kanäle und Lautsprecher ermöglichten diese eine bessere Ortung der Schallquellen und somit eine höhere Immersion für die Konsument*innen.¹

Limitierungen gibt es bei diesen kanalbasierten Formaten vor allem dadurch, dass für ein bestimmtes Lautsprechersetup z.B. 5.1 abgemischt wird und sich die Mischung nicht an abweichende Setups anpassen kann, da lediglich Lautsprecherkanäle in der Mischung vorhanden sind und keinerlei Information zu den einzelnen Klangobjekten und deren Position im 3D Raum enthalten sind. Auch eine Anpassung der Mischung z.B. eine Lautstärkeänderung des Dialogs im Vergleich zu Hintergrundgeräuschen ist nicht möglich, da diese Informationen bei kanalbasiertem Audio nicht voneinander trennbar sind.²

Durch das „Next Generation Audio“, zu dem objektbasiertes Audio gehört, sollen diese Limitierungen überwunden werden.³

¹ Vgl. Roginska, A., & Geluso, P. (2017). S. 40.

² Vgl. Herre, J., Hilpert, J., Kuntz, A., & Plogsties, J. (2015). S. 2.

³ Vgl. Olivieri, F., Peters, N., & Sen, D. (2019). S. 12.

2 Kanal-, Szenen- und Objektbasiertes Audio

Das NGA ermöglicht die Produktion und Vermischung von drei Audioformaten: kanalbasiertes Audio (CBA), szenenbasiertes Audio (SBA) und objektbasiertes Audio (OBA).

In den folgenden Unterkapiteln werden diese kurz anhand eines Beispiels aus der „EBU Technical Review: Scene-Based Audio and Higher Order Ambisonics: A technology overview and application to Next-Generation Audio, VR and 360 ° Video“⁴ definiert.

2.1 Kanalbasiertes Audio

Bei kanalbasiertem Audio wird explizit für bestimmte Lautsprecher setups gemischt. Jeder Lautsprecher bekommt einen Kanal zugeteilt.⁵

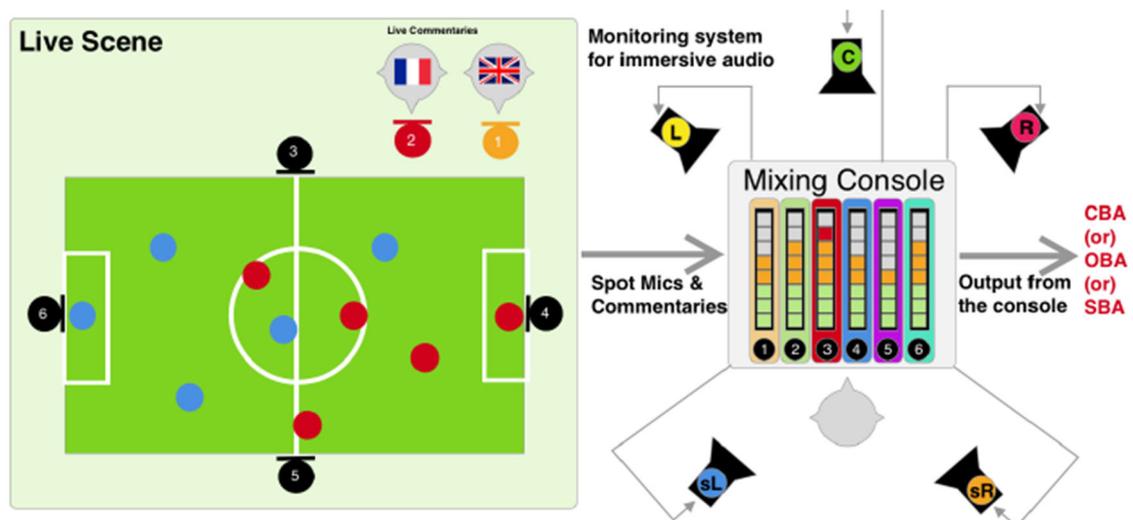


Abbildung 1: Beispielszenario der EBU⁶

Im Beispiel der EBU Technical Review wird ein Fußballspiel von einer Rundfunkanstalt kanalbasiert übertragen. Es wird jeweils eine 5.0 und eine Stereo Mischung mit englischem und französischem Kommentator produziert. Diese müssen alle separat abgemischt und übertragen werden.⁷

⁴ Vgl. Olivieri, F., Peters, N., & Sen, D. (2019).

⁵ Vgl. ebd., S. 12.

⁶ Vgl. ebd., S. 7.

⁷ Vgl. ebd. S.7.

Die beiden Beispielkonsumenten 1 und 2 zeigen dabei die Limitierungen des kanalbasierten Audios auf:

Konsument 1 hat zwar ein 5.0 Lautsprechersetup, hat diese allerdings nicht normgerecht positioniert. Die übertragene Mischung wird daher nicht korrekt wiedergegeben.⁸

Konsument 2 hat ein korrekt aufgebautes Stereosetup und empfängt auch ein Stereosignal, der Rundfunkanbieter hat allerdings das Spiel nur auf Französisch übertragen, um die zu übertragende Datenmenge zu reduzieren. Da Konsument 2 allerdings Englisch spricht, empfängt er das falsche Signal und kann dies nicht korrigieren.⁹

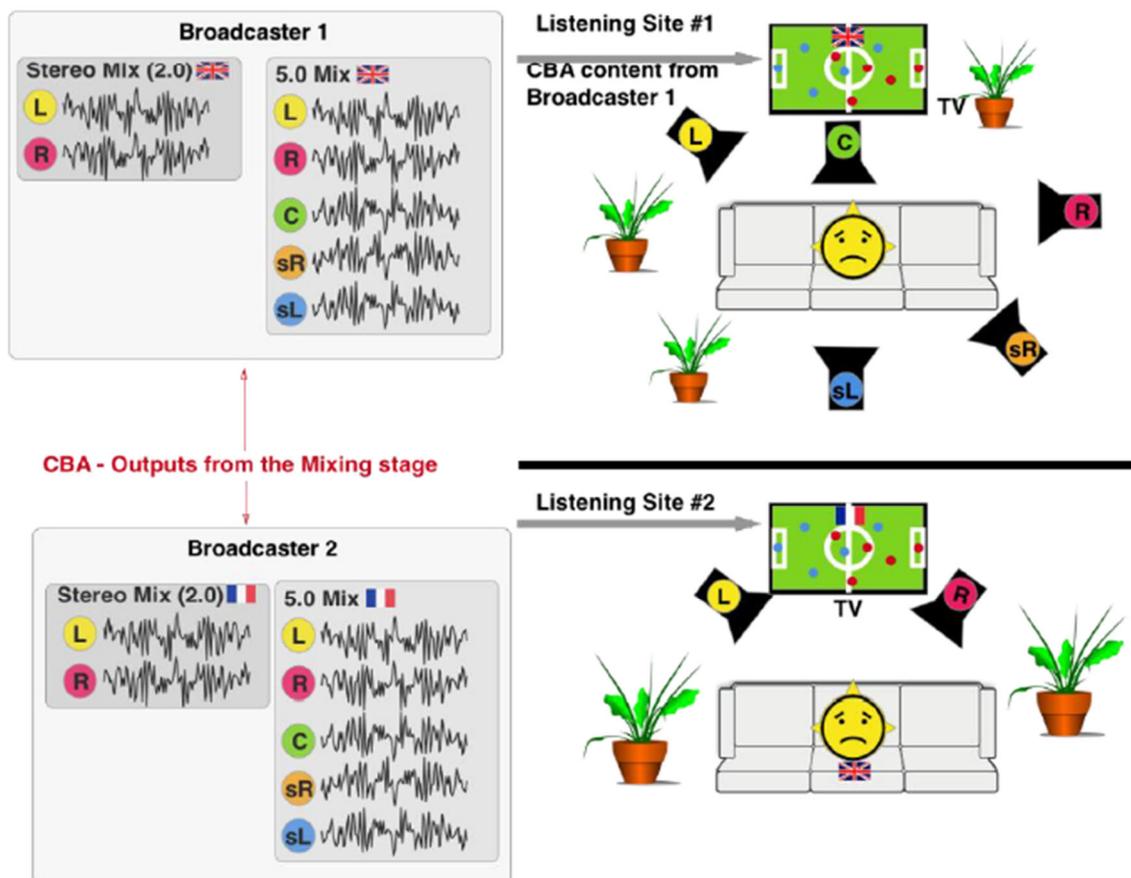


Abbildung 2: Sowohl Konsument 1 als auch Konsument 2 bekommen eine falsche Mischung.¹⁰

2.2 Szenenbasiertes Audio

Basierend auf Higher Order Ambisonics (HOA) ermöglicht szenenbasiertes Audio die Trennung der Mischung von einem spezifischen Lautsprecher-setup bei der Wiedergabe. Das Fußballspiel wird nicht als separate Kanäle übertragen, sondern als 3D Klangfeld.

⁸ Vgl. Olivieri, F., Peters, N., & Sen, D. (2019). S. 7.

⁹ Vgl. ebd., S. 8.

¹⁰ Vgl. ebd.

Bei Konsument 1 wird die Mischung korrekt wiedergegeben trotz normabweichendem Lautsprecher-setup, da beim Dekodieren das Audio an das spezifische Setup angepasst werden kann.¹¹

Bei Konsument 2 wird die Mischung zwar korrekt wiedergegeben, der Kommentator ist aber weiterhin fixer Bestandteil der Mischung und kann nicht ausgetauscht werden und spricht dadurch die falsche Sprache.

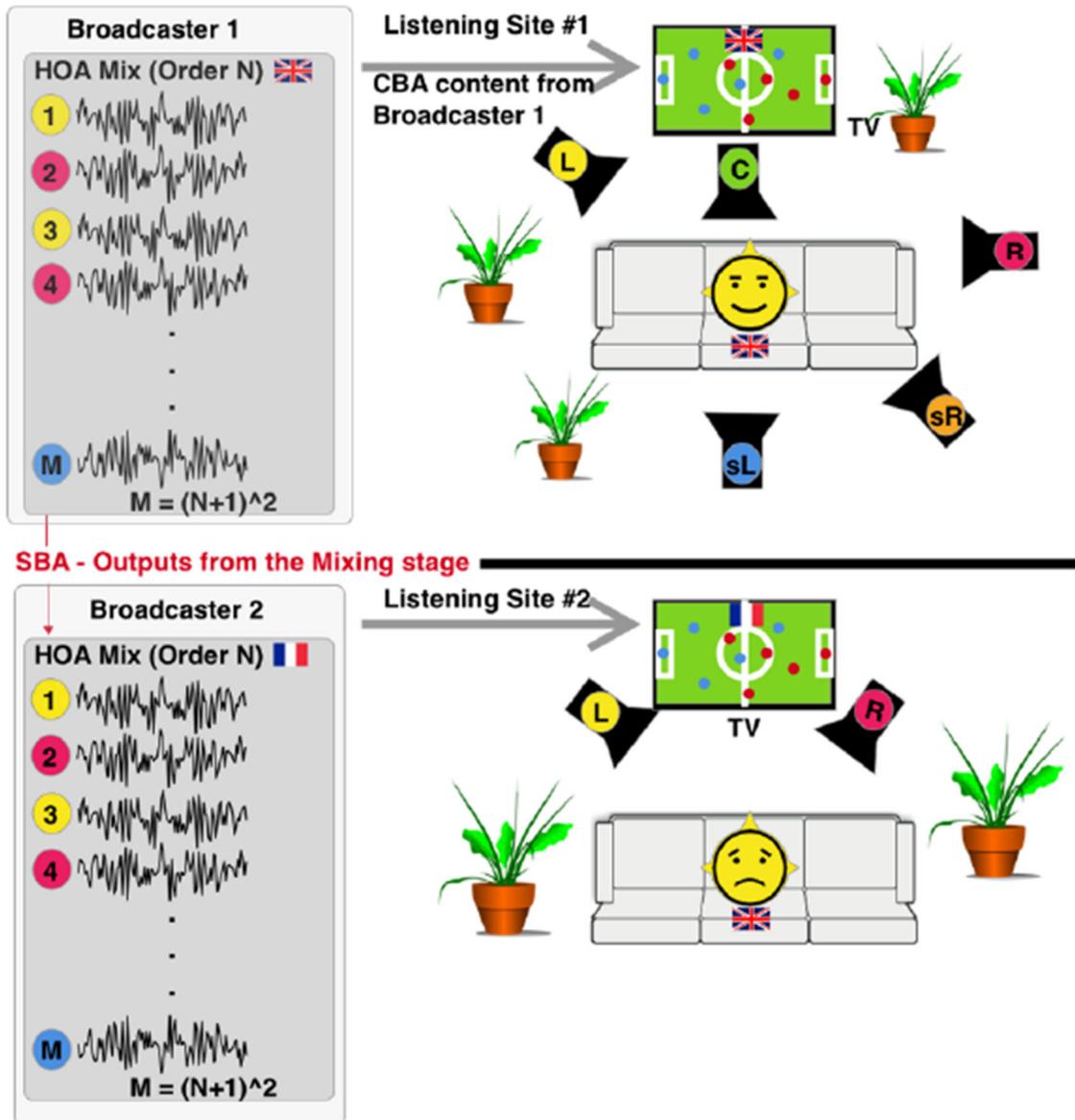


Abbildung 3: Konsument 1 bekommt eine richtige Mischung. Konsument 2 bekommt eine falsche Mischung.¹²

¹¹ Vgl. Olivieri, F., Peters, N., & Sen, D. (2019). S. 12.

¹² Vgl. ebd., S. 11.

2.3 Objektbasiertes Audio

Objektbasiertes Audio ermöglicht es, Klangobjekte auch bei der Übertragung und Wiedergabe voneinander zu trennen. Anstatt dass wie beim kanalbasierten Audio eine fertige Mischung übertragen wird, werden mehrere Objekte mit Metadaten gesendet.¹³

Dadurch ist es Konsument 2 möglich, zwischen mehreren Kommentator-Objekten auszuwählen und die korrekte Sprache zu empfangen.

Konsument 1 ist es allerdings nicht möglich, die Mischung dahingehend anzupassen, dass sie auf seinem Lautsprechersetup korrekt wiedergegeben wird.

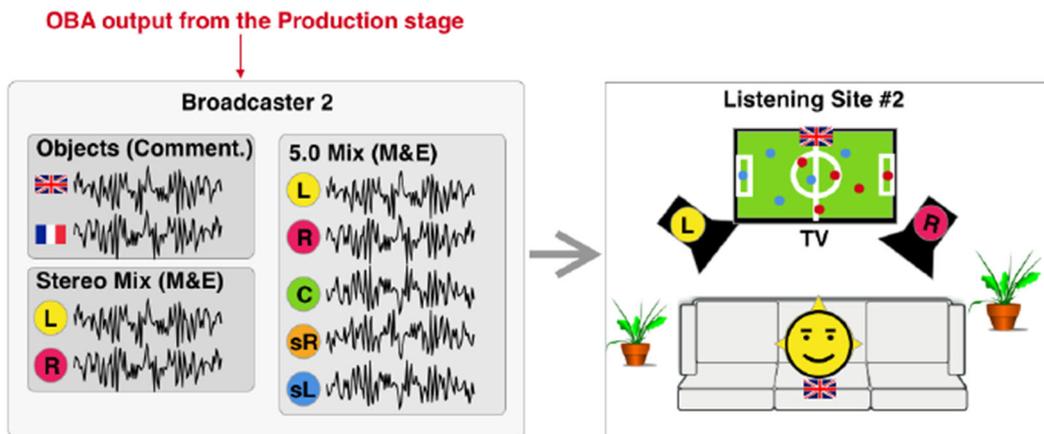


Abbildung 4: Mit OBA bekommt auch Konsument 2 die richtige Mischung.¹⁴

Ein weiterer Vorteil für die Rundfunkanstalt ist, dass sie Bandbreite sparen würden, wenn sie mehrere Sprachen übertragen wollen. Abbildung 6 zeigt ein Beispiel von Dolby AC-4 bei dem bei der Übertragung von einem Programm mit drei Sprachen die Datenrate um 54% gesenkt werden kann.¹⁵

Conventional Approach		Presentation Approach	
Stream	Data Rate/kbps	Stream	Data Rate/kbps
English 5.1	144	English Commentary 1.0	40
Spanish 5.1	144	Spanish Commentary 1.0	40
Chinese 5.1	144	Chinese Commentary 1.0	40
Commentary-free 5.1	144	Commentary-free 5.1	144
TOTAL	576	TOTAL	264
		Savings	54%

Abbildung 5: Beispielhafte Berechnung der ersparten Bandbreite bei Nutzung von OBA mit AC-4.¹⁶

¹³ Vgl. Olivieri, F., Peters, N., & Sen, D. (2019). S. 9.

¹⁴ Vgl. ebd., S. 10.

¹⁵ Vgl. Dolby. (2015). S. 13.

¹⁶ Vgl. ebd.

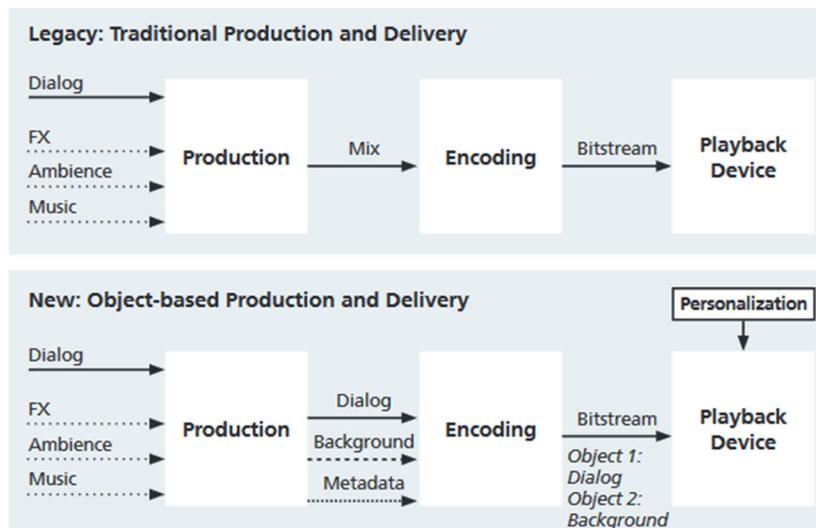


Abbildung 6: Vergleich der Produktion von CBA und OBA.¹⁷

2.4 Hybride Ansätze

Da die verschiedenen Formate CBA, SBA und OBA jeweils Vor- und Nachteile mit sich bringen, wird bei einer NGA Produktion empfohlen, eine Kombination der Varianten zu benutzen.¹⁸ Dadurch ist es dank SBA möglich, unabhängig von den Lautsprecher-Setups zu produzieren und die Bandbreite zu reduzieren. OBA ermöglicht es, bestimmte Anteile der Mischung zu personalisieren indem z.B. die Kommentatoren Sprache auswählbar ist oder der Dialog im Verhältnis zu Hintergrundgeräuschen lauter gemacht werden kann.¹⁹

¹⁷ Vgl. Simon, C., Torcoli, M., & Paulus, J. (2019). S. 5.

¹⁸ Vgl. Olivieri, F., Peters, N., & Sen, D. (2019). S. 12.

¹⁹ Vgl. ebd., S. 20.

3 Audiocodecs

In den letzten Jahren wurden mehrere Audiocodecs entwickelt, welche objektbasiertes Audio unterstützen. In den folgenden Unterkapitel werden zwei dieser Codecs präsentiert.

3.1 MPEG-H

Der MPEG-H Codec basiert einerseits auf dem Audio Definition Model für die Art wie die Daten und Metadaten abgespeichert werden, als Codec-Basis dient allerdings MPEG Unified Speech and Audio Coding (USAC) der um 3D Komponenten erweitert wurde.²⁰

Das Eingangssignal durchläuft einen USAC-3D Decoder und wird dann in die einzelnen Komponenten der Kanäle, Objekte, Metadaten, HOA und SAOC aufgeteilt und durchläuft die folgenden Schritte²¹:

- Die Kanaldaten werden in einem Format Converter an das vorhandene Lautsprecher Layout zum Beispiel durch intelligentes Downmixing angepasst.
- Objekte werden ebenfalls angepasst an das Lautsprecher Layout wiedergegeben. Zusätzlich werden die Metadaten miteinbezogen, um zum Beispiel Panning zu ermöglichen.
- Objekte, die mit dem erweiterten Spatial Audio Object Coding kodiert wurden, werden im SAOC-3D Decoder mithilfe der Metadaten wieder dekodiert.
- Higher Order Ambisonics Koeffizienten werden im HOA Renderer verarbeitet.

Dieser Mix wird nun kanalbasiert entweder an einen Binaural Renderer für eine Kopfhörerermischung oder an ein Lautsprechersystem weitergeleitet.²²

Ein Diagramm des Signalflusses befindet sich in Abbildung 7.

²⁰ Vgl. Herre, J., Hilpert, J., Kuntz, A., & Plogsties, J. (2015). S. 4.

²¹ Vgl. ebd.

²² Vgl. ebd., S. 3-4

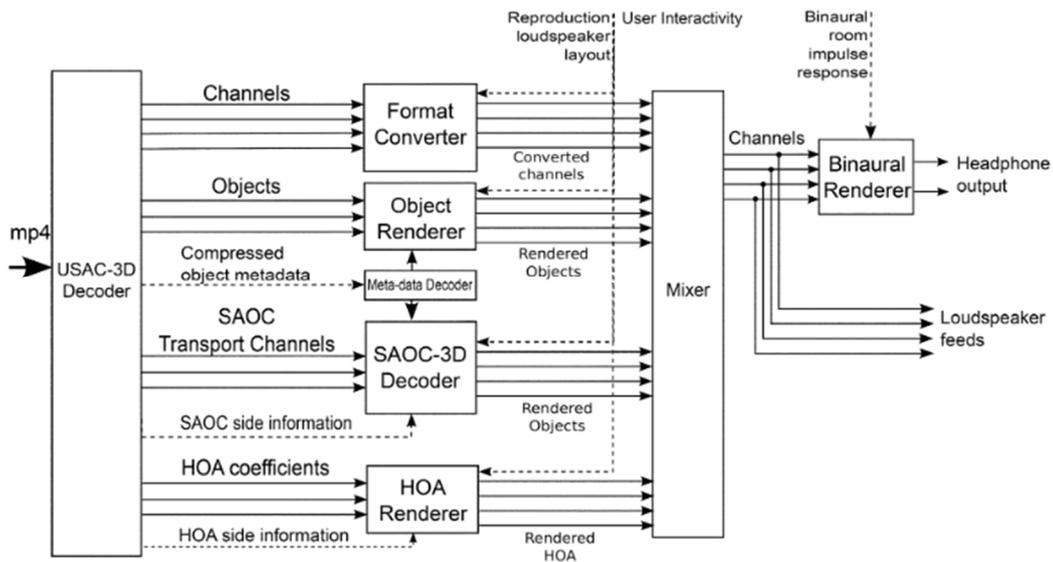


Abbildung 7: Blockdiagramm des MPEG-H 3D Decoder ²³

In Abbildung 8 durchläuft das Signal auch noch zusätzlich eine Dynamic Range Control und eine Loudness Normalisierung. Dies erlaubt Konsument*innen noch zusätzlich die Dynamik und Lautheit zu steuern. ²⁴

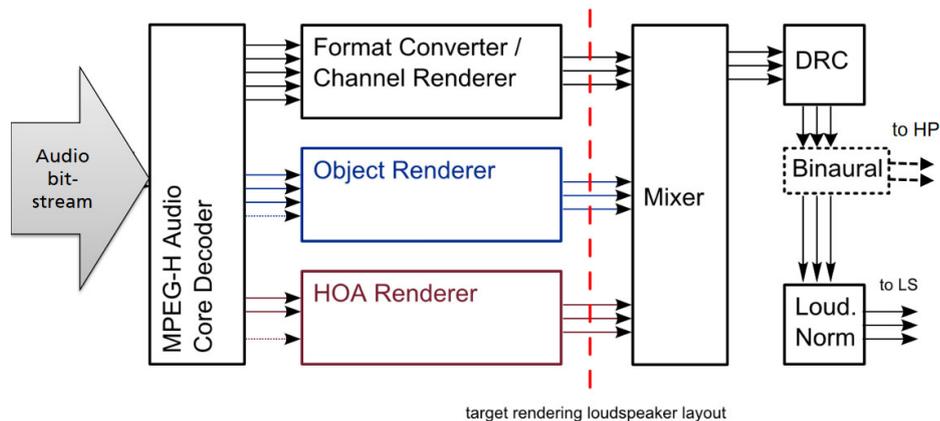


Abbildung 8: Vereinfachtes Blockdiagramm des MPEG-H Decoder. ²⁵

²³ Vgl. Herre, J., Hilpert, J., Kuntz, A., & Plogsties, J. (2015). S. 4.

²⁴ Vgl. Plogsties, J. (2015, Juli 15). S. 12-13.

²⁵ Vgl. ebd., S. 6.

3.2 AC-4

Der von Dolby entwickelte und von dem europäischen Institut für Telekommunikationsnormen standardisierte AC-4 Codec kann sowohl kanalbasiertes als auch objektbasiertes Audio verarbeiten.²⁶ Zu beachten ist allerdings, dass zur Erstellung und Wiedergabe jeweils proprietäre Dolby-Technologie wie z.B. ein Dolby Atmos Renderer benutzt werden muss.²⁷

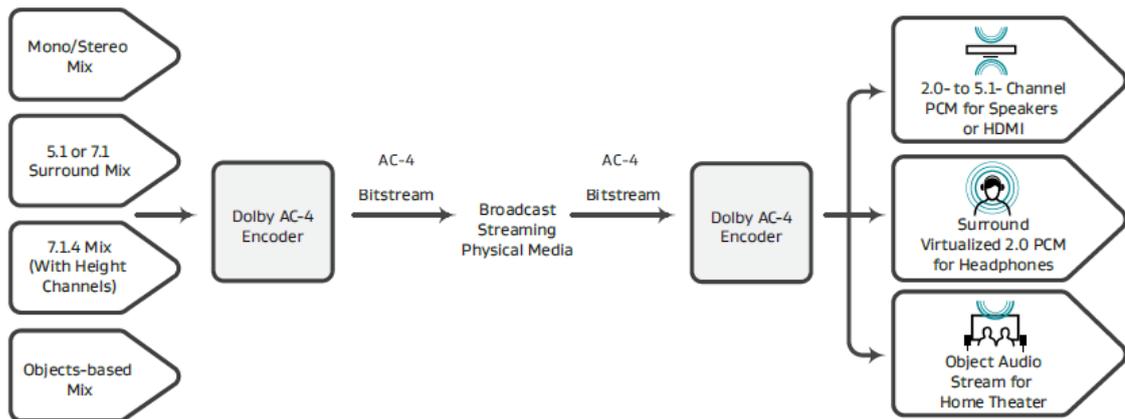


Abbildung 9: Nutzungsmöglichkeiten für den Dolby AC-4 Codec.²⁸

²⁶ Vgl. Dolby. (2015). S. 3.

²⁷ Vgl. ebd., S. 29.

²⁸ Vgl. ebd., S. 3.

4 Nutzungsmöglichkeiten

Durch die standardisierten Codecs wie MPEG-H und AC-4 ist es erstmals möglich, objektbasiertes Audio zu produzieren und zu konsumieren. Die folgenden Forschungsarbeiten haben sich mit diesem Thema beschäftigt und zeigen auf, welche Möglichkeiten objektbasierte Medien bieten.

4.1 Orpheus

Orpheus ist ein Forschungsprojekt des Fraunhofer IIS (DE – coordinator), Eurescom (DE), BBC (UK), IRT (DE), Elephantcandy (NL), Trinnov (FR), bcom (FR), IRCAM (FR), BR (DE) und Magix (DE) das von 2015 bis 2018 aktiv war.²⁹

Der Fokus lag dabei darauf, das Potential von objektbasierten Medien durch Prototypen aufzuzeigen.

Es wurde zum Beispiel eine iOS-App entwickelt, auf der ein Portfolio an object-based Projekten ausprobiert werden konnte. Ein Beispiel war ein Prototyp für Variable Length Wiedergabe. Bei dieser kann der/die Konsument*in während der Wiedergabe die Länge der Audiodatei verändern, indem er/sie die Wunschlänge angibt und der Player Abschnitte der Datei entfernt oder wieder einfügt, bis diese Länge erreicht ist.³⁰

Es wurde auch ein EBU ADM Renderer (EAR) entwickelt, welcher das komplette Audio Definition Model (ADM) unterstützt und als Grundlage für die EAR Production Suite dient.³¹

4.2 BBC

Das BBC R&D Department hat auch unabhängig von dem Orpheus Projekt objektbasierte Medien im Rundfunk erforscht. Auf ihrer Website beschreiben sie unter anderem, wie ein Tag für BBC-Konsument*innen im Jahr 2022 aussehen könnte. Dabei beschreiben sie, dass der Wettersprecher wenn notwendig automatisch durch Gebärdensprache ersetzt wird, Wettervorhersagen an den Terminkalender angepasst werden, automatische personalisierte Zusammenfassungen verpasster Folgen einer Serie erstellt werden, Podcasts sich an die Dauer der Pendelstrecke anpassen und die Musikauswahl des Radios sich an Hörer*innen anpasst.³²

²⁹ Vgl. EBU. (2018). S. 3.

³⁰ Vgl. Orpheus Results. (o. J.).

³¹ Vgl. ebd.

³² Vgl. A day in the lives of BBC audiences in 2022. (o. J.).

Für die Wiedergabe mit variabler Länge wurde, ebenso wie bei dem Orpheus Projekt, ein Prototyp namens „Responsive Radio“ entwickelt und veröffentlicht. ³³

4.3 S3A

BBC R&D hat gemeinsam mit den Universities of Surrey, Salford und Southampton ein fünfjähriges Forschungsprojekt durchgeführt, bei dem es um das Potential von 3D Sound für immersive Projekte ging.³⁴

Im Bereich object-based spatial audio entstand dabei unter anderem eine Forschungsarbeit bei dem objekt-basiertes Audio genutzt wurde, um den Konsument*innen persönliche Mischungen zu erlauben, mit einem Fokus auf Verständlichkeit. Mittels eines Reglers kann man fließend zwischen Immersion und Verständnis wählen. Je nach Position des Reglers werden dann mehr oder weniger Hintergrundgeräusche eingefügt, welche zwar das auditive Gesamtbild unterstützen, aber nicht zwingend zur Narrative beitragen und so auch die Verständlichkeit der Sprache mindern können.³⁵

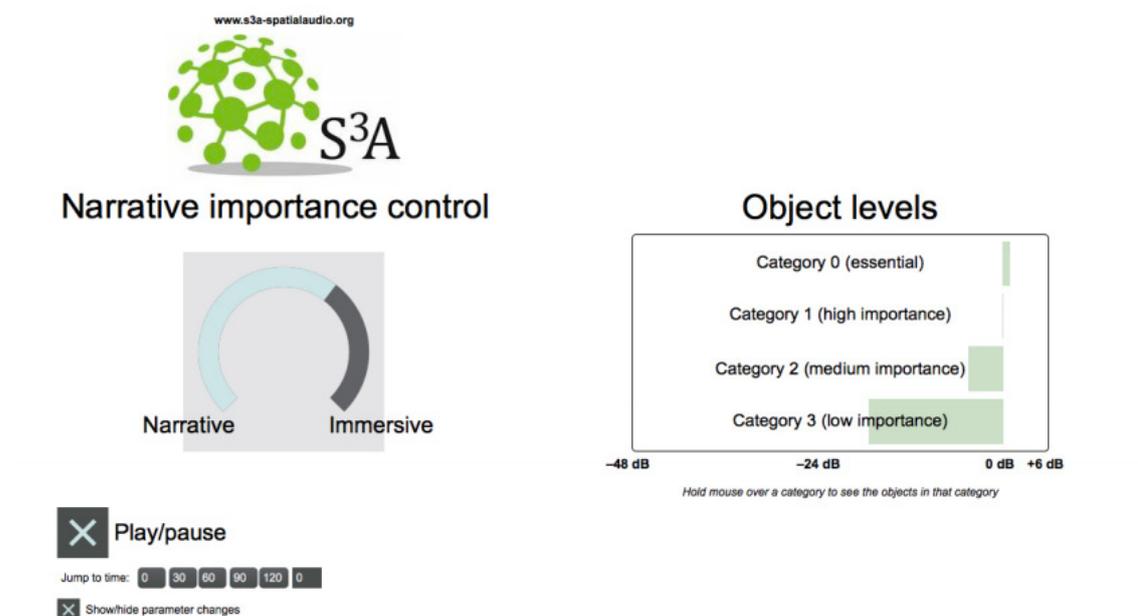


Abbildung 10: Userinterface für die Narrative importance control.³⁶

³³ Vgl. Brooks, M. (2015, März 13).

³⁴ Vgl. About S3A. (o. J.).

³⁵ Vgl. Shirley, B., Ward, L., & Chourdakis, E. T. (2019). S. 1.

³⁶ Vgl. Thornton, M. (2019).

5 Beispielhafter objektbasierter Workflow mit der EAR Production Suite

5.1 Die EAR Production Suite

Die EAR Production Suite ist eine Sammlung von VST Plugins, um NGA-Content in einer DAW zu produzieren. Die Suite erlaubt es den Nutzer*innen objekt- und kanalbasierte Tracks zu definieren und diese auch im BW64 Format mit ADM Metadaten zu speichern.³⁷

Ausgelegt ist die Plugin-Sammlung für die DAW Reaper.³⁸

5.2 Definition des Beispielprojekts

Um den Workflow mit der EAR Production Suite aufzuzeigen wird ein exemplarisches Projekt definiert. Enthalten in dem Projekt sollen sein:

- Eine Sprecherin die einen Text vorliest (Mono, Center)
- Eine Atmosphäre (Stereo)
- Ein Auto das vorbeifährt (Mono)

Die Sprecherin und die Atmosphäre werden als kanalbasierte Objekte gehandhabt, während das Auto als Objekt an den Konsument*innen vorbeifahren wird.

5.3 Arbeitsschritte

5.3.1 Tracksetup

Das Basisprojekt besteht aus drei Tracks mit den jeweiligen Audiodateien. Dieser Workflow beschreibt nicht Mixing oder Mastering, sondern es wird von finalen Audiospuren bzw. -dateien ausgegangen, welche nun objektbasiert veröffentlicht werden sollen. Lediglich Einstellungen der Ear Production Suite Plugins und deren Automation werden beschrieben.

³⁷ Vgl. Firth, M., Bailey, R., & Pike, C. (2020, Oktober). S. 3.

³⁸ EAR Production Suite. (o. J.).

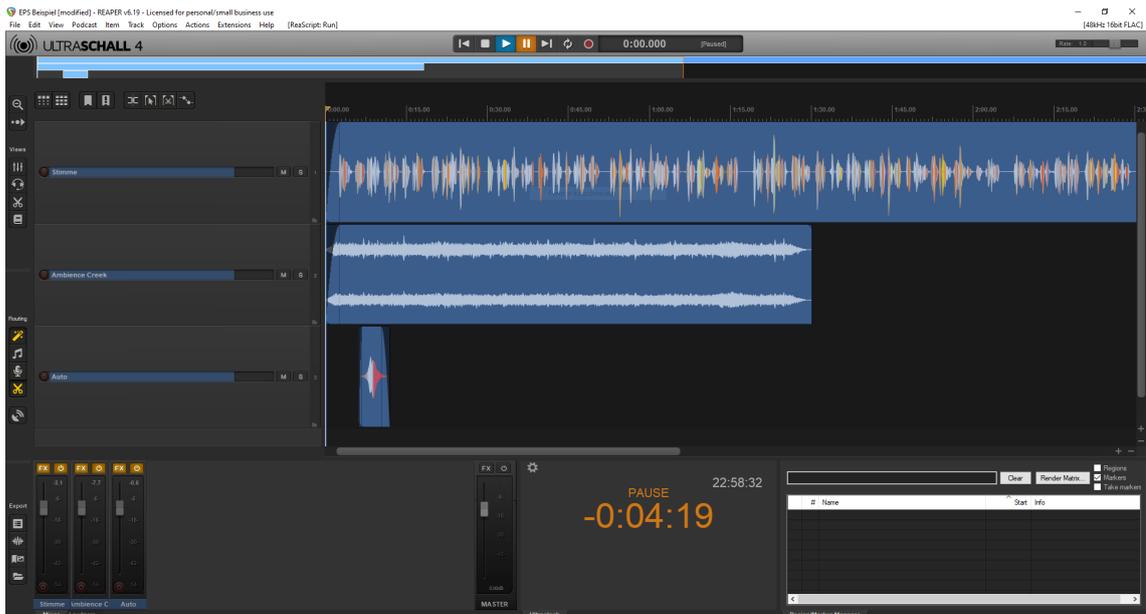


Abbildung 11: Beispielprojekt mit drei Tracks.

Zusätzlich zu den drei Tracks werden noch zusätzliche Tracks für die EAR Scene und EAR Monitoring Plugins benötigt.

In diesem Fall fügen wir einen Track für die Scene ein und jeweils einen Track für Stereo Monitoring und einen Track für 5.1 Monitoring.

Die Scene- und Monitoring-Tracks benötigen eine Kanalbreite von 64.

5.3.2 Routingsetup

Bei den drei Audiotracks und dem Scene-Track muss eingestellt werden, dass sie kein Audio an den Mastertrack senden.

Die Audiotracks müssen jeweils an den Scene-Track senden und der Scene-Track an die Monitoring-Tracks.

Die Breite der Sends der Audiotracks müssen an das Audio angepasst sein. Für ein Mono- und Objekt-Signal benötigt der Send eine Breite von 1, für Stereo 2 usw. Jeder Track muss sein Signal aber an einen separaten Kanal der Scene schicken.

In dem Beispielprojekt sendet:

- Stimme Mono 1 auf EPS-Scene 1
- Ambience Creek Stereo 1/2 auf EPS-Scene 2/3
- Auto Mono 1 auf EPS-Scene 4
- EPS-Scene 1/64 auf Stereo-Monitoring und 5.1-Monitoring jeweils 1/64
- Stereo-Monitoring und 5.1-Monitoring an den Mastertrack

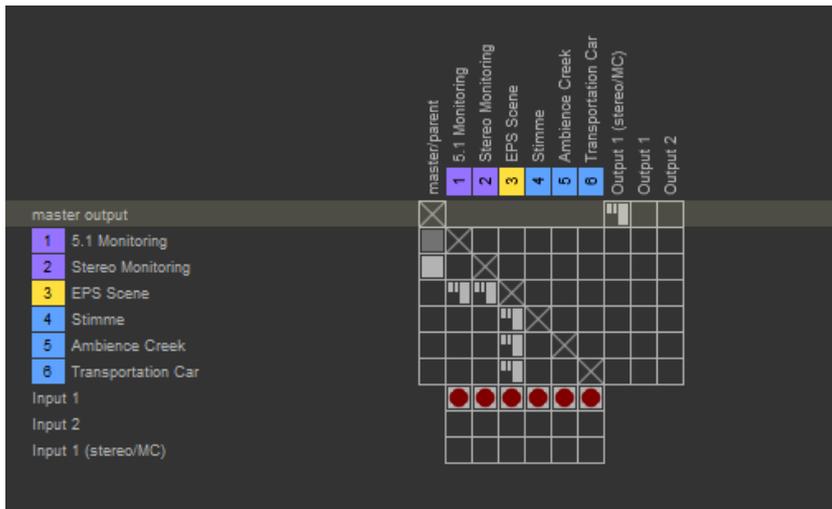


Abbildung 12: Routing des Beispielprojekts.

5.3.3 Plugin Setup

In den EPS-Scene Track wird das „EAR Scene Plugin“ eingefügt. Standardmäßig ist dieses im „Auto Mode“, wenn es eingefügt wird. Im „Auto Mode“ fügt es alle Elemente automatisch in das Audioprogramm ein. Diese können im Nachhinein wieder entfernt oder in andere Programme verschoben werden.

In den Track für die Stimme und die Ambience wird das Plugin „EAR Direct Speakers“ eingefügt. Dieses Plugin behandelt den Track als kanalbasiertes Audio. Es muss eingestellt werden, welches Format (Mono, Stereo, 5.1, ...) der Track benutzt und an welchen Kanal in der Szene das Audio geschickt wird.

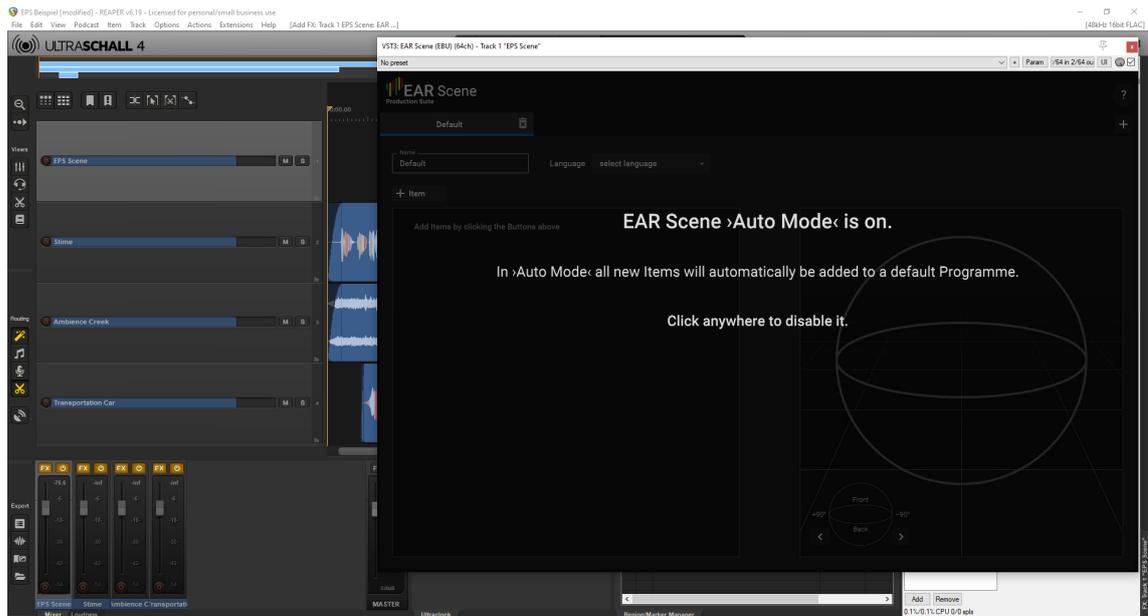


Abbildung 13: Das EAR Scene Plugin im „Auto Mode“.

In den Track des objektbasierten Audios wird das Plugin „EAR Object“ eingefügt. Es muss auch wieder angegeben werden, auf welchem Kanal die Szene das Objektaudio empfangen wird.

Im Gegensatz zu dem Direct Speakers Plugin können im Object Plugin auch die Parameter des Objekts angepasst werden. In Abbildung 14 können die Benutzeroberflächen verglichen werden.

Unter anderem kann der/die Benutzer*in die Lautstärke, Distanz und Diffusion des Objekts anpassen und auch automatisieren.

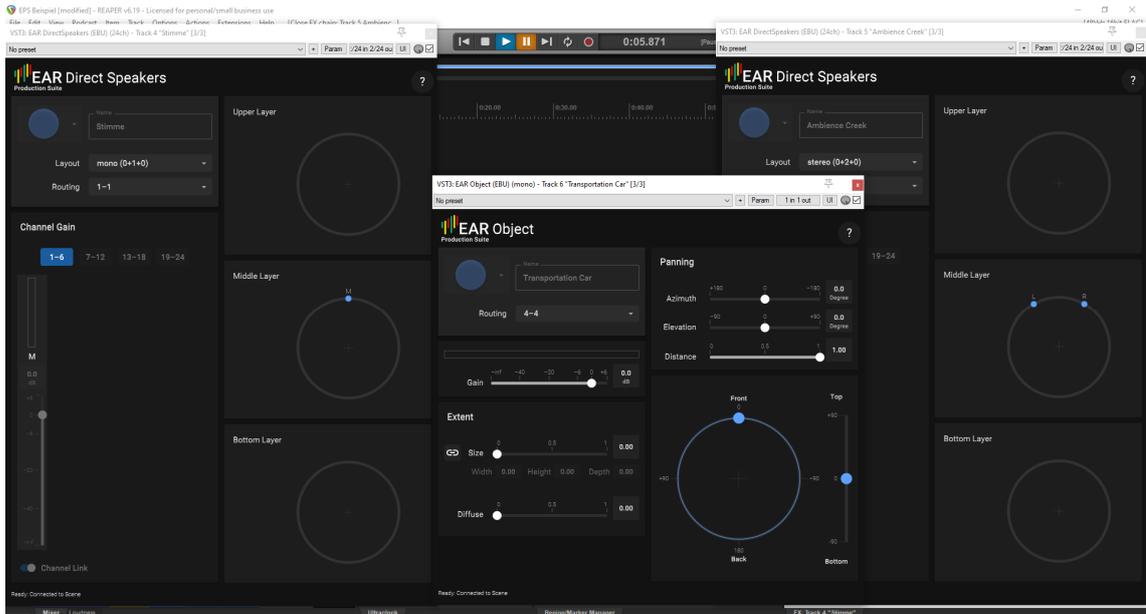


Abbildung 14: Die „Direct Speakers“-Plugins und das „Object“-Plugin für das Beispielprojekt.

Als letztes müssen noch die Plugins in die Monitoring Tracks eingefügt werden. In den 5.1 Monitoring Track wird das „EAR Monitoring 0+5+0“ eingefügt und in den Stereo Monitoring Track das „EAR Monitoring 0+2+0“.

Nach diesem Schritt kann das objektbasierte Audio sowohl über Stereo als auch 5.1 gegengehört und angepasst werden.

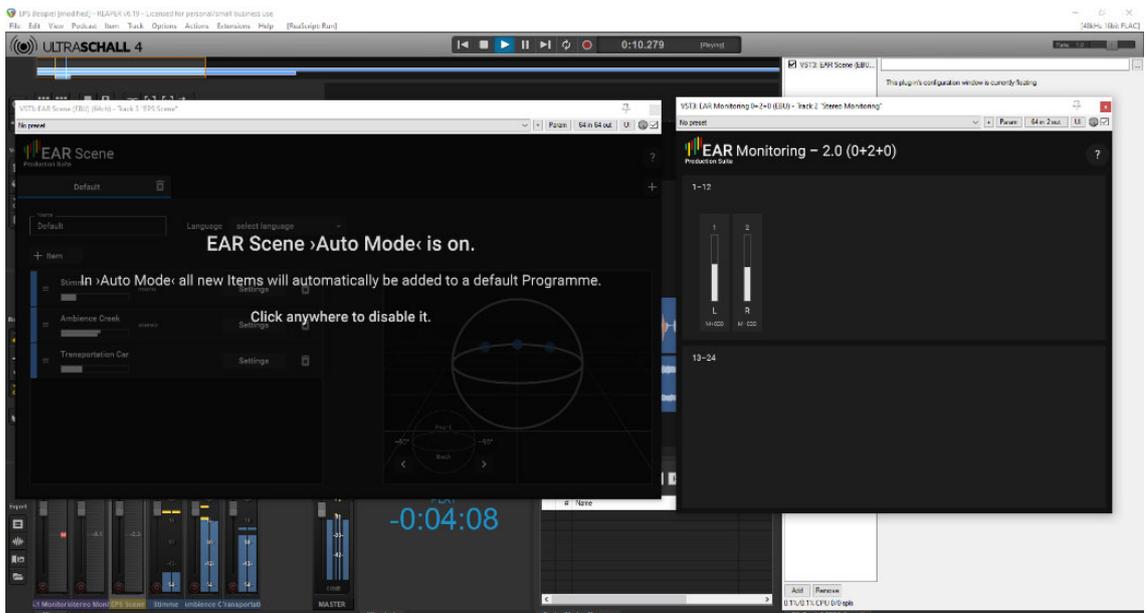


Abbildung 15: Das Signal der Scene kommt beim Stereo-Monitoring-Plugin an.

5.3.4 Automation und Interaktivität

In diesem beispielhaften Workflow soll das Auto an den Hörer*innen vorbeifahren. Um dies zu erreichen muss der Azimuth-Wert des Objekts automatisiert werden. Das „EAR Object“-Plugin stellt diesen Parameter Reaper zur Automation zur Verfügung. Über eine Automationskurve kann dieser einfach angepasst werden.

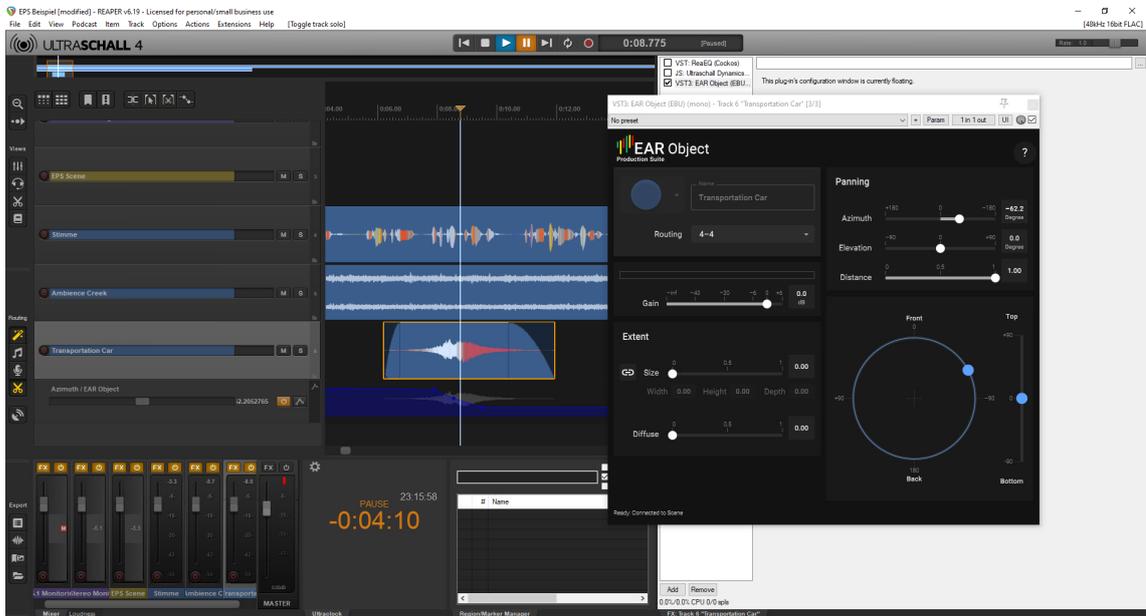


Abbildung 16: Die angepasste Automationskurve des Azimuth Parameters.

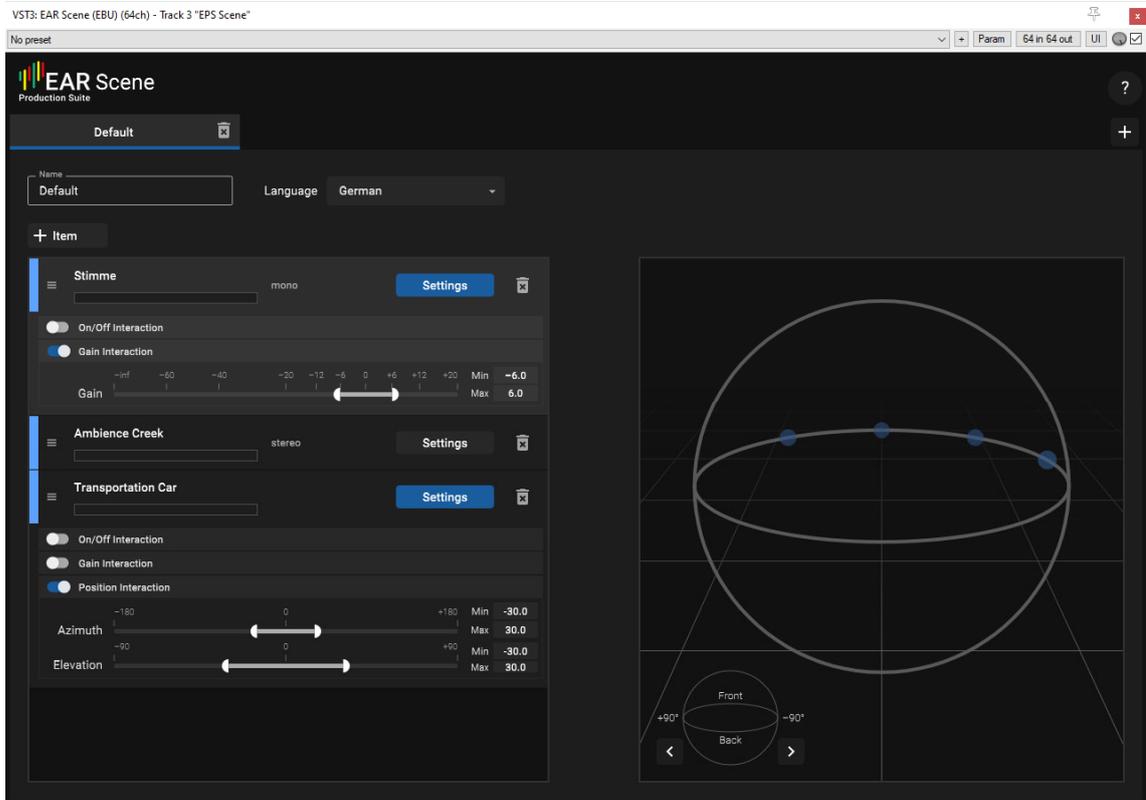
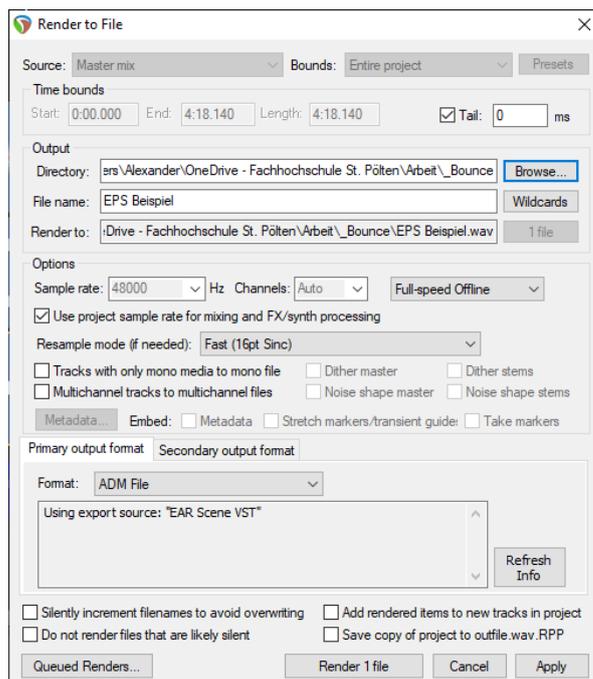


Abbildung 17: In dem Scene-Plugin können Einstellungen bezüglich der Interaktivität des Programms angepasst werden.

Zusätzlich kann bei dem Scene-Plugin eingestellt werden, inwiefern das Programm interaktiv für Konsument*innen sein soll. So kann eingestellt werden, ob von einzelnen Tracks die Lautstärke, unabhängig von anderen Kanälen, veränderbar sein soll. Tracks, die objektbasiertes Audio enthalten, können auch noch weiter angepasst werden und zum Beispiel von den Konsument*innen selber im Raum verschoben werden. Diese Einstellungen können, je nach Bedarf, sowohl aktiviert, deaktiviert als auch limitiert werden, so dass zum Beispiel die Lautstärke nur um +/- 6 dB verändert werden kann. Ein Beispiel für diese Einstellungen sind in Abbildung 17 zu sehen.

5.3.5 Export

Das Projekt kann entweder kanalbasiert über die Monitoringtracks exportiert werden, wodurch die Interaktivitätsmöglichkeiten allerdings verloren gehen, oder im BW64-Format (.wav) mit ADM-Metadaten. In dem BW64-Format bleiben die objektbasierten Vorteile bestehen, es muss allerdings auch ein ADM-Renderer benutzt werden, um die Dateien abzuspielen.



Um das Projekt objektbasiert zu bouncen muss im Reaper-Renderer als Dateiformat „ADM File“ angegeben werden.

Abbildung 18: Export des Projekts als ADM Datei in Reaper.

6 Fazit

Objektbasiertes Audio bietet viele Vorteile im Vergleich zu kanal- und szenenbasiertem Audio. Es kann unabhängig von der Wiedergabesituation produziert werden und einzelne Aspekte der Produktion können von den Konsument*innen den Bedürfnissen entsprechend während der Wiedergabe angepasst werden. Zusätzlich bietet es auch auf der Produktionsseite neue Möglichkeiten. Objekte können einfach im 3D Raum bewegt und akkurat wiedergegeben werden, bei der Übertragung kann die notwendige Datenrate gesenkt werden und auch Lösungen für Binauralwiedergabe sind bei OBA Codecs oft enthalten.

Zum aktuellen Zeitpunkt ist die Technologie allerdings noch im Entwicklungsstadium. Es gibt nur wenige Geräte die z.B. MPEG-H decodieren können und auch die Produktionsstools für objektbasiertes Audio sind noch nicht ausgereift. Tools wie die EAR Production Suite ermöglichen es zwar schon ADM Dateien zu exportieren und importieren, dabei kommt es aber oftmals zu Fehlern und Abstürzen. Auch ist der Begriff des objektbasierten Audios im Vergleich zu Next Generation Audio noch sehr limitiert implementiert. Prototypen wie sie z.B. die BBC entwickelt mit Variable Length Playback sind zu diesem Zeitpunkt mit OBA nicht möglich, da sich die Objekte bei OBA bis jetzt immer auf Tracks und nicht Clips beziehen.

Es wäre weitere Forschung und Entwicklung notwendig, welche einerseits die bestehenden Lösungen auf ein produktionsreifes Niveau anheben und andererseits Projekte, die aufzeigen was mit einem wirklichen objektbasiertem Audio, welches sich nicht nur auf Tracks sondern auch auf Clips beziehen kann, möglich wäre.

Quellenverzeichnis

A day in the lives of BBC audiences in 2022. (o. J.). Abgerufen 23. Februar 2021, von BBC website: <https://www.bbc.co.uk/rd/object-based-media#audiences>

About S3A. (o. J.). Abgerufen 23. Februar 2021, von S3A website: <https://www.s3a-spatialaudio.org/about-s3a>

Brooks, M. (2015, März 13). Future Content Experiences: The First Steps For Object-Based Broadcasting. Abgerufen 23. Februar 2021, von BBC website: <https://www.bbc.co.uk/rd/blog/2015-03-future-content-experiences-the-first-steps-for-object-based-broadcasting>

Dolby. (2015). *Dolby AC-4: Audio Delivery for Next-Generation Entertainment Services*.

EAR Production Suite. (o. J.). Abgerufen 23. Februar 2021, von EAR Production Suite website: <https://ear-production-suite.ebu.io/>

EBU. (2018). *TR042: EXAMPLE OF AN END-TO-END OBA BROADCAST ARCHITECTURE AND WORKFLOW*. SOURCE: ORPHEUS PROJECT. Geneva.

Firth, M., Bailey, R., & Pike, C. (2020, Oktober). *The EAR Production Suite*. BBC R&D. Abgerufen von <http://downloads.bbc.co.uk/rd/workstreams/next-generation-audio/bbc-rd-ear-production-suite-oct20-mf.pdf>

Herre, J., Hilpert, J., Kuntz, A., & Plogsties, J. (2015). *MPEG-H Audio—The New Standard for Universal Spatial/3D Audio Coding*. Erlangen, Germany: International Audio Laboratories Erlangen, A Joint Institution of Universität Erlangen-Nürnberg and Fraunhofer IIS. Abgerufen von International Audio Laboratories Erlangen, A Joint Institution of

Universität Erlangen-Nürnberg and Fraunhofer IIS website: <http://www.aes.org/e-lib/browse.cfm?elib=17556>

Olivieri, F., Peters, N., & Sen, D. (2019). *EBU Technical Review: Scene Based Audio and Higher Order Ambisonics: A technology overview and application to Next-Generation Audio, VR and 360° Video*. San Diego, California, USA: EBU.

Orpheus Results. (o. J.). Abgerufen 23. Februar 2021, von Orpheus website: <https://www.orpheus-audio.eu/results/>

Plogsties, J. (2015, Juli 15). *MPEG-H Audio System for Broadcasting (ITU-R Workshop "Topics on the Future of Audio in Broadcasting")*. Fraunhofer IIS.

Roginska, A., & Geluso, P. (2017). *Immersive Sound The Art and Science of Binaural and Multi-Channel Audio*. Routledge.

Shirley, B., Ward, L., & Chourdakis, E. T. (2019). *Personalization of Object-based Audio for Accessibility using Narrative Importance*. Manchester, UK: TVX 2019.

Simon, C., Torcoli, M., & Paulus, J. (2019). *MPEG-H AUDIO FOR IMPROVING ACCESSIBILITY IN BROADCASTING AND STREAMING*. Fraunhofer Institute for Integrated Circuits IIS.

Thornton, M. (2019). Object Based Audio Can Do So Much More Than Just Dolby Atmos? We Explore. Abgerufen von Pro Tools Expert website: <https://www.pro-tools-expert.com/production-expert-1/2019/3/15/is-there-more-to-object-based-audio-than-just-dolby-atmos-we-explore>

Abbildungsverzeichnis

Abbildung 1: Beispielszenario der EBU	5
Abbildung 2: Sowohl Konsument 1 als auch Konsument 2 bekommen eine falsche Mischung.	6
Abbildung 3: Konsument 1 bekommt eine richtige Mischung. Konsument 2 bekommt eine falsche Mischung. ¹²	7
Abbildung 4: Mit OBA bekommt auch Konsument 2 die richtige Mischung.	8
Abbildung 5: Beispielhafte Berechnung der ersparten Bandbreite bei Nutzung von OBA mit AC-4. ¹⁶	8
Abbildung 6: Vergleich der Produktion von CBA und OBA.....	9
Abbildung 7: Blockdiagramm des MPEG-H 3D Decoder	11
Abbildung 8: Vereinfachtes Blockdiagramm des MPEG-H Decoder.....	11
Abbildung 9: Nutzungsmöglichkeiten für den Dolby AC-4 Codec.	12
Abbildung 10: Userinterface für die Narrative importance control.	14
Abbildung 11: Beispielprojekt mit drei Tracks.	16
Abbildung 12: Routing des Beispielprojekts.....	17
Abbildung 13: Das EAR Scene Plugin im „Auto Mode“	18
Abbildung 14: Die „Direct Speakers“-Plugins und das „Object“-Plugin für das Beispielprojekt.	19
Abbildung 15: Das Signal der Scene kommt beim Stereo-Monitoring-Plugin an.	19
Abbildung 16: Die angepasste Automationskurve des Azimuth Parameters.	20
Abbildung 17: In dem Scene-Plugin können Einstellungen bezüglich der Interaktivität des Programms angepasst werden.	20
Abbildung 18: Export des Projekts als ADM Datei in Reaper.....	21