FORSCHUNG

ORPHEUS THE ORPHEUS PROJECT: BUILDING THE WORKFLOW FOR END-TO-END OBJECT-BASED AUDIO BROADCASTING

MICHAEL WEITNAUER AND ANDREAS SILZLE

Driven by the technological challenges of the media landscape, as well as new user expectations and demands, the so-called 'object-based' audio technology has been developed. Object-based media is a revolutionary approach for creating and deploying interactive, personalised, scalable and immersive content.

ORPHEUS is a H2020-funded EU project involving ten European major players – broadcasters, manufacturers and research institutions – sharing the common vision to make the best use of this technology for both broadcasters and end users. During a 30-month run-time, the consortium develops, implements and validates an object-based end-to-end media chain for audio content in a real-life broadcast environment. ORPHEUS is driven by two pilots to demonstrate both linear and non-linear audio experiences using an IP-based custom-built broadcast chain.

This article presents an overview of the basic architecture, explains the challenges in developing this system and outlines the innovative tools that were created for recording, mixing, monitoring, storing, distributing, playing-out and rendering of object-based audio.

Introduction

The media landscape has been subject to significant changes over the past years. The advent of mobile devices with broadband internet access and on-demand services has led to new media consumption patterns and consumer expectations.

However, broadcast infrastructure and production workflows have not significantly evolved over the past decades. Therefore, especially all broadcasters face increasing efforts to support growing numbers of formats and target platforms. Driven by these problems, the so-called object-based audio technology has been developed.

Object-based audio is a ground-breaking new concept, transforming the way media is created, produced and provisioned. It is a highly effective approach for delivering programmes as individual elements to the audience for both future broadcast emission and IP delivery.

The ORPHEUS partners (Figure 1) are building an end-to-end object-based audio solution that addresses current issues and offers useful personalization options, e.g. speech level control, to the end user. The project's main objective is to create a reference architecture, sketching the possibilities and advantages of producing and delivering object-based audio in a chain of interoperable building blocks, based on open standard technologies and formats.

The basic architecture has already proven its applicability in the first pilot phase. The ORPHEUS project is presently elaborating the second phase of the pilot, reaching the final straight of the 30-month project.

Object-based Audio

Object-based audio is a revolutionary approach for the creation and deployment of interactive, scalable, immersive and cross media applications for any type of media content. This is possible by representing the audio content as a set of individual assets together with metadata describing their relationships and associations. Examples for this metadata information are e.g. 3D positions, gain or a semantic label such as "Dialogue". This allows media objects to be assembled in novel ways to enhance user experiences. The most important features are:

- Personalization of content representation (e.g. speech level customization or multi-lingual features)
- Immersive experience for any kind of content
- The delivery of audio content in a format- and system-agnostic manner.

Further insight into curation, responsiveness and user-experience of object-based broadcasting is available in a paper by Shotton et al. [1].

ORPHEUS' Objectives

In detail, the ORPHEUS objectives are to:

1. Develop, implement and validate a complete end-to-end object-based broadcasting chain under real-world conditions;
2. Apply, contribute and disseminate open standards of

If not indicated otherwise, all images: the authors

Figure 1. ORPHEUS consortium

Experience our object-based vision with the ORPHEUS App on your iOS device: https://orpheus-audio.eu/ios-app"

object-based formats and tools within the architecture;

3. Examine the adaptability of existing broadcast technology to object-based production and develop a concept for the transition of existing infrastructure, systems and software/tools to a regular operational service of object-based audio;
4. Provide a reference architecture and guidelines on how to implement an end-to-end broadcasting chain for object-based audio.

However, even though ORPHEUS is focussing on audio-only content, it is required that the workflow, the lessons learnt and the guidelines will be applicable to TV or Video broadcasting infrastructures, too.

Requirements for Pilots 1 and 2

The ORPHEUS project first designed a broad specification, which then was iteratively refined through implementation and experimentation. First, requirements from the point of view of both content creators and listeners have been identified. To this purpose, ORPHEUS carried out an in-depth study of current radio production [2]. This included proposals for new workflows incorporating object-based technologies, from using RFID to identify radio studio participants (e.g. which person is talking), to incorporating production documentation, social media contributions, and 3D panning for spatial audio.

For the receiver, ORPHEUS has defined an architecture wherein compressed audio streams and metadata are decoded, then user interaction features are provided by altering metadata in a personalisation stage. A renderer receives the audio streams and metadata and automatically adapts to the playback setup (headphones, 2D or 3D loudspeakers). Environmental adaptation is proposed with inputs from sensors for head-tracking of binaural rendering, and for compensation of background noise level or room equalisation.

Pilot 1, recently completed [13], used live content creation and distribution, from microphones in the studio all the way through to listeners using mobile devices, in-home hi-fi, or web browsers.

Pilot 2, presently under development, will integrate responsive features for on-demand listening, adding variable-length programming, e.g. requesting a version of a programme to have a desired duration: „I have 15 minutes to listen to a drama," or extending one: „I'd like this topic to be expanded."

Architecture Overview

A pilot implementation architecture has been developed (Figure 2), which will be further refined during the project time, based on experiences from the project pilots and lead to a reference architecture for an end-to-end object-based audio production workflow. The latest version of the architecture for Pilot 1 covers the established workflow stages of broadcasting: recording, pre-production & mixing, radio studio, distribution and reception.

The contribution aspect of the architecture for Pilot 1 was designed to suit for both legacy channel-based content (e. g. Stereo or 5.1) as well as pre-produced and live produced object-based content. Yet, the distribution focuses mainly on the Internet, as the broadcast emission of object-based content via legacy channels like DVB or DAB.

While some of the necessary components are partly available in proprietary implementations, ORPHEUS is aiming at closing existing gaps with own developments. The focus of the entire workflow was to provide solutions by applying open and international standards such as the Au-

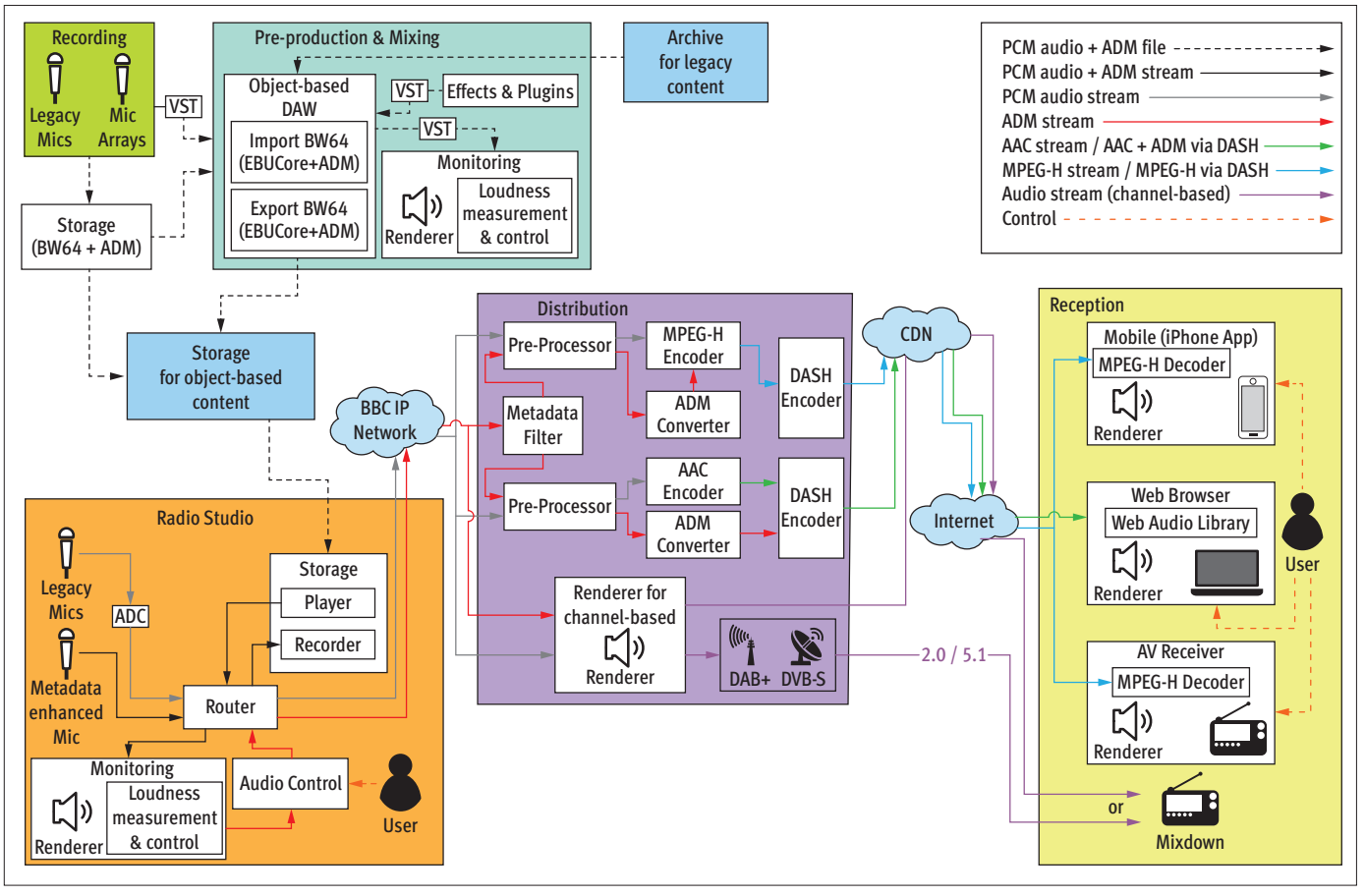

Figure 2. ORPHEUS reference architecture

dio Definition Model (ADM) [4] for the production stages. Moreover, the formats and tools for all production blocks are to allow for best possible interoperability, not being tied to specific commercial systems. The usage of uncompressed audio signals, either as BW64 [3] file exchange or uncompressed AES67 streams [14], throughout all production stages has been implemented to allow for the best possible exchange with other system components. Another argument for uncompressed audio and ADM is the archiving aspect of broadcast workflow, in order to preserve content in best possible quality for future reuse.

Recording

The recording stage provides hardware to capture audio signals and related metadata for a genuine object-based production. Capturing of sound may employ spot microphones, which will be used for sound objects; or different microphone arrays, which will be used to produce a multichannel or Higher-Order Ambisonics (HOA) bed for the ambient part. Other well-known microphone setups can also be used for a channel-bed or even stereophonic objects.

Ideally, object-based audio content is recorded directly in a DAW (Digital Audio Workstation) using the ADM format. Yet, this may also be achieved independently from the DAW. In this case the recorded content is transferred in form of multichannel BW64 [3] files with ADM [4, 5] metadata. Depending on the established broadcast architecture, these files may be stored in a temporary ADM-enabled storage or directly in the storage for object-based content.

Pre-Production and Mixing

The purpose of the pre-production and mixing stage in Figure 2 is to deliver tools for editing existing object-based content or creating such content from recordings or legacy audio material. Within ORPHEUS, the core of this block is the object-based DAW Sequoia by MAGIX. A typical application for a DAW is to apply effects or plugins onto audio tracks. During authoring, the DAW is also used for the positioning and other metadata editing of objects, as shown in Figure 3.

To deliver objects along with HOA beds, the ORPHEUS partner b<>com has developed several plugins for analysing and manipulating HOA recordings for object-based audio workflows.

Another important use case is the usage of reverberation with audio objects. Currently, reverb could only be included as a fixed part of the individual objects or as an additional audio object, which partially limits the possibilities of ADM. IRCAM has investigated and published a range of techniques that can be used to overcome these challenges, and have proposed ways to integrate this within the ADM format [6].

A suite of applications (Mac and PC) with advanced ADM functionalities was also created by IRCAM: ADM Player, ADM Recorder, ADM Renderer and ADM Extractor (Figure 4), [7]. These tools are now freely available for everyone on the IRCAM web page [15].

Monitoring

To preview the object-based content and simulate the user experience, it is important to monitor the content with different speaker setups, including binaural monitoring. As the work on a standardized production renderer within ITU-R is still ongoing, one of the candidates for this standard (Fraunhofer IIS' MPEG-H renderer) was implemented in Sequoia and the radio studio environment for this purpose.

In addition to pure audio monitoring, other parts of the user experience can be simulated within the DAW, e.g. transport control, dialog foreground/background balance and exchanging language specific objects.

Radio Studio

As the ORPHEUS project is focussing on audio-only content, a radio workflow was the natural choice. Radio production is usually made live with several pre-produced elements and the moderator of a radio show is typically using a radio studio where the programme is composed and distributed.
The ORPHEUS radio studio is situated in the BBC Broadcasting House (Figure 5). The studio was re-fitted with 11 loudspeakers (7 head height, 4 elevated). The mixing console was replaced by a large touch-screen and a PC. There are two microphones, connected via AES67.

The adjacent technical equipment is centred on BBC R&D's "IP Studio". This is a production platform using commodity computer hardware and IP networks [8]. It serves as a framework for developing standards in several international forums, such as the Advanced Media Workflow Association (AMWA) [9]. For the ORPHEUS project, the IP Studio software runs on an off-the-shelf, rack-mounted server.

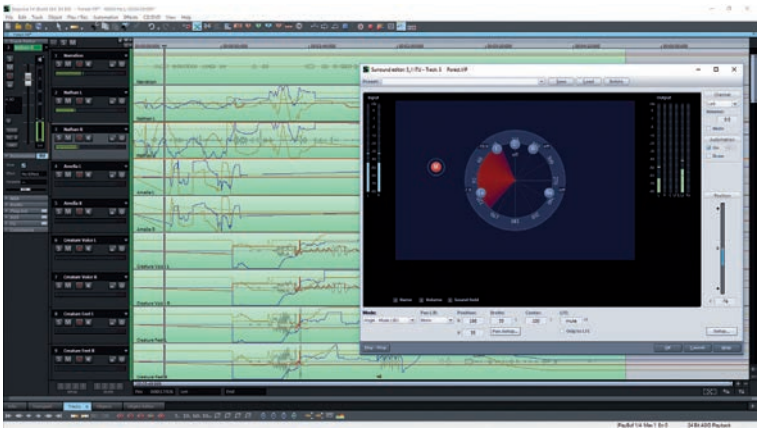

Figure 3. Sequoia displaying ADM tracks and object panner window

Distribution

Due to the diversity of platforms, which need to be served from a broadcaster (e.g. broadcast channel, own Video on Demand platform, 3rd party platforms, smartphone apps, etc.), it is very unlikely that a broadcaster will distribute its content with one NGA (Next Generation Audio) system only. Therefore, it was important for ORPHEUS to design a workflow that is basically independent from the play-out format. This demand has also been reflected in the ORPHEUS architecture by providing two delivery formats of the content: MPEG-H streams (including audio essence and metadata) for mobile devices and the AV Receiver, as well as legacy encoded audio objects (AAC) with separate ADM metadata for web browser representation. Both delivery formats are transported by MPEG-DASH streams [17] . For more details about the features of the MPEG-H format see [10, 11].

To offer the best possible flexibility for object-based productions, it is essential to provide an automatic approach to limit the number of objects throughout the workflow stages. Different production steps may require different numbers of objects, e.g. authoring, archive and distribution. A manual reduction conducted by a sound engineer appears as optimum, but not necessarily from an economical point of view. Hence, IRT is investigating and developing a solution that can reduce objects automatically.

According to limitations of the chosen play-out codec, the available bandwidth of the network or processing power

MICHAEL WEITNAUER

is an R&D Engineer and Project Manager at IRT and coordinating IRT's activities in ORPHEUS.

ANDREAS SILZLE

is a senior scientist at Fraunhofer IIS in Erlangen. He is the technical coordinator of the EU project ORPHEUS.

of the end-user device, the receiver will be offered an appropriate streaming format, i.e. where some audio objects are kept independent while others are pre-combined or pre-rendered into a monophonic or multi-channel "proxy object" or beds. The streams can also be delivered at various bitrates, depending on the end-user's listening situation. Additionally, streams may be selected based upon user preferences (such as language, available listening time, etc.). The play-out module will support a large diversity of standardised 2D or 3D formats, e.g. 2.0, 5.1, 7.1+4, and binaural.

Reception and User Experience

The dedicated user devices for object-based media consumption within ORPHEUS are fully IP connected. 'Legacy broadcast' like DAB+ can only be supplied with broadcaster-hosted rendering and this limits the possibilities for user interaction. For the ORPHEUS pilots, three reception scenarios and respective solutions were identified and developed:

- mobile device: iOS app (Elephantcandy)
- web browser: WebAudio API renderer (BBC R&D and IRT)
- living room: AV receiver (Trinnov)

In these use cases, the specific key features of object-based audio broadcasting are implemented:

- navigation within programme elements
- transcripts with replay modus
- foreground/background levels adjustment
- dynamic range control
- rendering to various reproduction layouts

The most mature implementation of user features is provided by the iOS app, where – in addition – various audio rendering functionalities can be stored as customized profiles, representing different environment and listening conditions. Furthermore, it is even possible to activate those profiles automatically by i.e. a GPS position.

The screen shots of the iOS prototype app in Figure 6 exemplify the user-friendly implementation of these features, following dedicated user requirement analysis [12] and usability evaluations.

Summary

The following achievements have been made during the course of the project so far:

- CREATION: Various object-based content productions completed and published
- PRODUCTION: a variety of capturing, analysing, editing and mixing tools for object-based audio as well as an ADM-compatible Sequoia DAW
- PLAY-OUT: prototype of fully IP-connected radio studio with implemented object-based features
- DISTRIBUTION: content server with encoding capabilities for object-based audio ADM plus AAC audio streams as well as MPEG-H
- RECEPTION: implementation of object-based audio decoding and reproduction in browser, iOS app and CE AV-receiver

The ORPHEUS project is demonstrating the advantages of object-based media as an innovative, universal and consequent approach for media production. It also positions this technology as an essential component for emerging cross-media demands, being integrative and scalable. Moreover, it is capable of supporting the transition from linear to non-linear, and both on-air and on-demand listening, using broadcast and IP technology.

The consortium is heavily involved in the standardization work of the relevant organizations (ITU-R, DVB, MPEG & EBU) to contribute to open standards and to make sure that all necessary formats and protocols will work successfully together.

ORPHEUS publishes project deliverables, the reference architecture and guidelines on how to implement object-based audio chains on their website at http://orpheus-audio.eu/public-deliverables/.

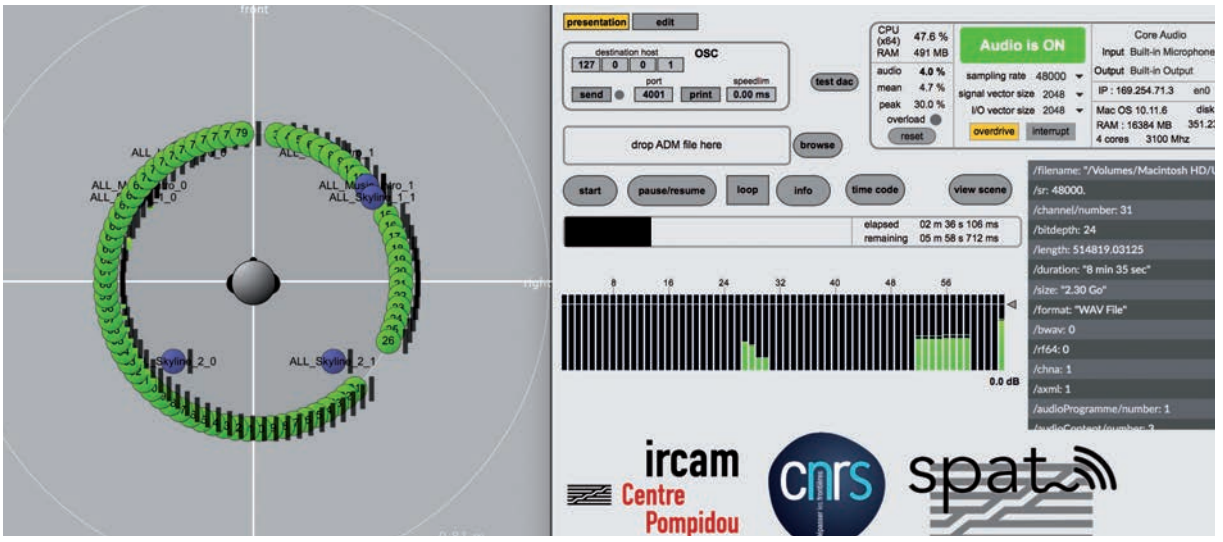

Figure 4. IRCAM ADM tools

Figure 5. ORPHEUS studio at BBC Broadcasting House

Acknowledgment

This work was partly funded from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 687645 (ORPHEUS project). The paper only reflects the authors' views. The Commission is not responsible for any use that may be made of the information it contains.

References

[1] Shotton, M., F. Melchior, M. Evans, et al. Object-based Broadcasting – Curation, Responsiveness and User Experience. IBC. 2014. Amsterdam, The Netherland.

[2] D3.1: Requirements, designs and workflows of an object-based production environment; Deliverable of EU project ORPHEUS. 2016. http://orpheus-audio.eu/wp-content/uploads/2016/04/Orpheus-D3.1_Requirements-designs-and-workflows-of-an-object-based-production-environment.pdf.

[3] ITU-R BS.2088-0, Long-form File Format for the International Exchange of Audio Programme Materials with Metadata. 2015, Intern. Telecom Union, Geneva, Switzerland. https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.2088-0-201510-I!!PDF-E.pdf.

[4] ITU-R BS.2076-0, Audio Definition Model. 2015, Intern. Telecom Union, Geneva, Switzerland. https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.2076-0-201506-I!!PDF-E.pdf.

[5] Füg, S., D. Marston, and S. Norcross. The Audio Definition Model – A Flexible Standardised Representation for Next Generation Audio Content in Broadcasting and Beyond. 141st AES Convention. 2016. Los Angeles, USA.

[6] D3.2: Implementation and documentation of reverberation for object-based audio broadcasting; Deliverable of EU project ORPHEUS. 2016. http://orpheus-audio.eu/wp-content/uploads/2016/12/orpheus-d3-2_implementation-and-documentation-of-reverberation-for-object-based-broadcast.pdf.

[7] Geier, M., T. Carpentier, M. Noisternig, et al. Software Tools for Object-based Audio Production Using the Audio Definition Model (ITU-R Recommendation BS.2076). ICSA. 2017. Graz, Austria.

[8] Brightwell, P.J., J.D. Rosser, R.N.J. Wadge, et al. The IP Studio. IBC. 2013. Amsterdam, The Netherlands.

[9] Networked Media Reference Architecture v1.0. 2015, Joint Taskforce for Networked Media (JT-NM). http://jt-nm.org/RA-1.0/.

[10] Herre, J., J. Hilpert, A. Kuntz, et al., MPEG-H 3D Audio – The New Standard for Coding of Immersive Spatial Audio. IEEE Journal of Selected Topics in Signal Processing – Special Issue on Spatial Audio, 2015. 9(5): p. 770-779.

[11] Bleidt, R.L., D. Sen, A. Niedermeier, et al., Development of the MPEG-H TV Audio System for ATSC 3.0. IEEE Transactions on Broadcasting, 2017. 63(1): p. 202..236, DOI: https://doi.org/10.1109/TBC.2017.2661258.

[12] D5.1: Document on user requirements; Deliverable of EU project ORPHEUS. 2016. http://orpheus-audio.eu/public-deliverables/.

[13] http://www.bbc.co.uk/rd/blog/2017-09-mermaids-tears-object-based-audio

[14] AES67-2015: "AES standard for audio applications of networks – High-performance streaming audio-over-IP interoperability". Audio Engineering Society. 2015-09-21

[15] http://forumnet.ircam.fr/product/spat-en/admix-en/

[16] http://www.amwa.tv/

[17] ISO/IEC 23009-1, Information technology – Dynamic adaptive streaming over HTTP (DASH) – Part 1: Media presentation description and segment format

Figure 6. Screenshots of the Pilot 1 prototype iOS app with three windows: element navigation, transcripts and profile selection (pre-sets of parameters).