ORPHEUS AUDIO PROJECT:
PILOTING AN END-TO-END OBJECT-BASED
AUDIO BROADCASTING CHAIN

Andreas Silzle[1], Michael Weitnauer[2], Olivier Warusfel[3], Werner Bleisteiner[4], Tilman Herberger[5], Nicolas Epain[6], Benjamin Duval[7], Niels Bogaards[8], Chris Baume[9], Uwe Herzog[10]

[1] Fraunhofer IIS, Erlangen, Germany, andreas.silzle@iis.fraunhofer.de
[2] Institut für Rundfunktechnik, München, Germany, weitnauer@irt.de
[3] ircam, Paris, France, olivier.warusfel@ircam.fr
[4] Bayerischer Rundfunk, München, Germany, werner.bleisteiner@br.de
[5] MAGIX, Dresden, Germany, therberger@magix.net
[6] b<>com, Rennes, France, nicolas.epain@b-com.com
[7] Trinnov Audio, Paris, France, benjamin.duval@trinnov.com
[8] Elephantcandy, Amsterdam, The Netherlands, niels@elephantcandy.com
[9] BBC R&D, London, United Kingdom, chris.baume@bbc.co.uk
[10] Eurescom, Heidelberg, Germany, herzog@eurescom.eu

ABSTRACT

Object-based media is a revolutionary approach for creating and deploying interactive, personalised, scalable and immersive content. It allows media objects to be assembled in novel ways to create new and enhanced user experiences. Object-based media is flexible and responsive to user needs as well as environmental and platform-specific factors.

ORPHEUS is a H2020-funded EU project involving ten European major players – broadcasters, manufacturers and research institutions. During a 30-month project, we develop, implement and validate an object-based end-to-end media chain for audio content. We are running two pilots to demonstrate both linear and non-linear audio experiences using a custom-built broadcast chain. The first pilot was a live radio broadcast with enhanced functionalities, including immersive sound, foreground/background control, language selection, and in-depth programme metadata.

This paper presents initial results of the first pilot, 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. To encourage the support of the broadcast industry in adopting this new technology, ORPHEUS is working towards the publication of a reference architecture and general guidelines for successful implementation of object-based audio in a real-life broadcast environment.

INTRODUCTION

The media landscape has been subject to significant changes over the past years. The advent of mobile devices with mobile 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, all broadcasters face increasing efforts to support growing numbers of formats and target platforms.

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 in the fully IP-connected media eco-system of the future.

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. It enables:

a) Multi-dimensional and multi-lingual features;

b) Novel interactive user experiences and personalised audio content, and;

c) The delivery of audio content in a format-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

The ORPHEUS objectives are to:

1. Develop, implement and validate a complete end-to-end object-based broadcasting chain, including the research required to implement missing tools for recording, mixing and monitoring, storing, play-out, distributing, as well as receiving of object-based audio and metadata;

2. 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;

3. Demonstrate a new, prodigious user experience through the creation of workflows and applications for the use of object-based audio as an emerging future broadcast technology;

4. Based on the findings of objectives 1 to 3, create reference architecture and guidelines on how to implement an end-to-end broadcasting chain for object-based audio.

Requirements and Pilots 1 and 2

The ORPHEUS project is structured to firstly design a broad specification, then iteratively refine that through implementation and experimentation. First, requirements from the point of view of users (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 participants, to incorporating production documentation, social media contributions, and 3D panning for spatial audio.

On the reception side, 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 for background noise level or room equalisation.

In two pilots, the developed technology is to proof the ORPHEUS vision.

Pilot 1, recently completed, 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, yet to come, 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."

PILOT PHASE 1

Programme Planning and Preparing

From the beginning, radio has always been a live medium, and still today, live events such as sports and concert transmissions drive radio's continued popularity. Yet, with the development of on-demand listening and implicit demand to arrange and structure content, as well as to optimize the listening experience in its technical and aesthetic quality, refined models of programme and scheduling schemes have been applied.

At present, with the deployment of radio content not only live 'on air' but also 'on-demand', this has lead to complex tasks of appending necessary metadata for preparing and provisioning content for multiple online platforms.

ORPHEUS pilot 1 started with radio's origins, staging and transmitting live a radio drama – including all metadata for optimal adaptability to later on-demand exploitation.

Following that, pre-recorded object-based content (i.e. music, documentaries etc.) were added, to create a run-down like in a typical 'magazine' programme.

Architecture Overview

At present a pilot implementation architecture has been developed (Figure 1). It 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 contains five macroblocks: recording, pre-production & mixing, radio studio, distribution and reception.
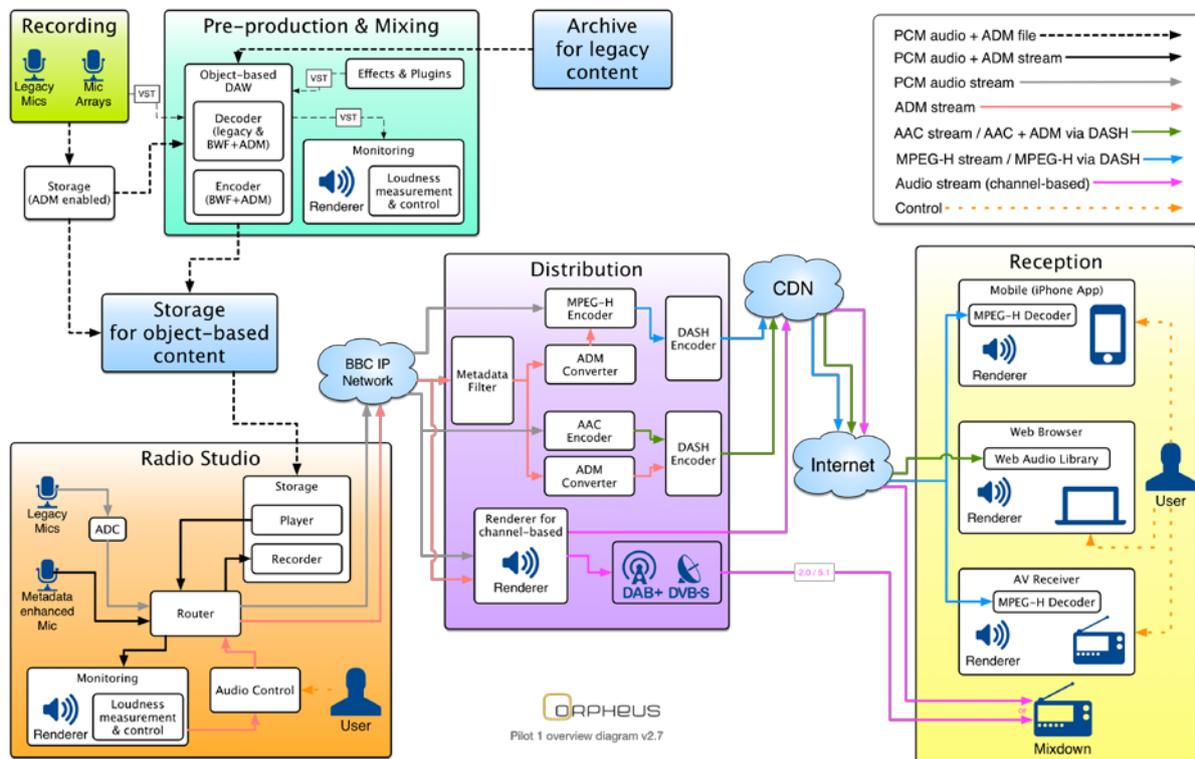

Figure 1: Pilot implementation architecture

The architecture for Pilot 1 was designed to suite for both legacy channel-based content as well as pre-produced and live produced object-based content. Yet, the distribution was focused mainly on the Internet, as the broadcast emission of object-based content via legacy channels like DVB or DAB do not currently support the specific object-based features

Recording

The recording macroblock provides hardware to capture audio signals and related metadata for a genuine object-based production. Capturing of sound may employ: a) proximity microphones, which will be used for sound objects; or b) two different compact microphone arrays, which will be used to produce a multichannel or Higher-Order Ambisonics (HOA) bed for the ambient part.

Ideally, object-based audio content is recorded in the ADM (Audio Definition Model) format directly in the DAW. 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.

The following main ADM components in MAGIX object based DAW Sequoia have been configured for the first pilot phase in summer 2017:

Pre-Production and Mixing

The purpose of the pre-production and mixing block in Figure 1 is to deliver tools for editing existing object-based content or creating such content from legacy audio material or other sources. The core of this block is the object-based DAW Sequoia by MAGIX.

A typical application for a DAW is the appliance of effects or plugins onto audio tracks. Usually effects and plugins can be applied to audio content within the DAW.

ORPHEUS partner b<>com has developed several plugins for analysing and manipulating HOA recordings for object-based audio workflows.

A special use case is applying reverb to the content. 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. ORPHEUS partner IRCAM has investigated and published a range of techniques that can be used to overcome these challenges, and have proposed ways to integrate this with the ADM format [6].

ICRAM also created a suite of applications (Mac and PC) with advanced ADM functionalities: ADM Player, ADM Recorder, ADM Renderer and ADM Extractor (Figure 2).
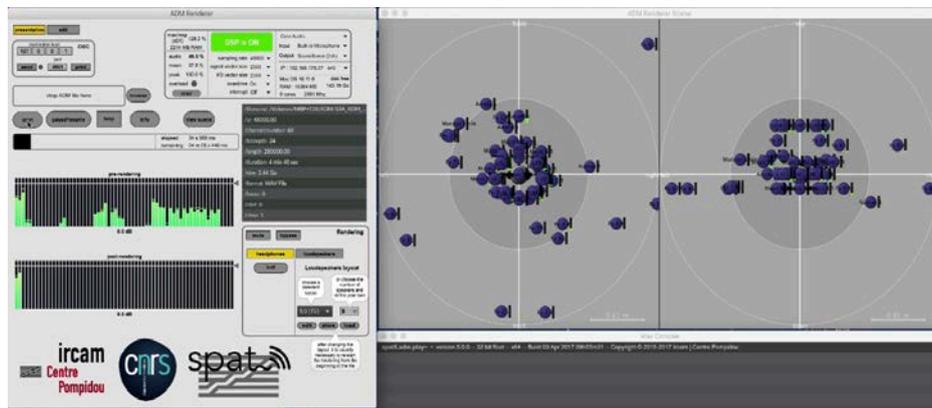

Figure 2: IRCAM ADM tools

Monitoring

For previewing the object-based content and simulating the user experience, it is important to allow monitoring with different speaker setups, including binaural monitoring. For this purpose, Fraunhofer IIS' MPEG-H renderer was implemented in MAGIX DAW Sequoia.

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

The ORPHEUS radio studio is located in BBC Broadcasting House (Figure 3). The studio was re-fitted with 11 loudspeakers (7 head height, 4 above head height). The mixing console was replaced with a large touch-screen and a PC. There are two microphones, connected by AES67.

Figure 3: Orpheus studio at BBC Broadcasting House

The remaining technical equipment is centred on BBC R&D's "IP Studio". This is a production platform using commodity computer hardware and IP networks [7]. It serves as a framework for developing standards in several international forums, such as the Advanced Media Workflow Association (AMWA)[1] [8]. For the ORPHEUS project, the IP Studio software runs on an off-the-shelf, rack-mounted server.

Distribution

Hardware or software solutions provide means for decoding and rendering audio streams received over the Internet. Two main delivery formats are implemented: MPEG-H streams (including audio essence and metadata), and legacy encoded audio objects (AAC) with separate ADM metadata. For more details about the features of the MPEG-H format see [9, 10].

According to the available bandwidth of the network and processing power of the end-user device the receiver shall select/ask for 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". The streams can also be delivered at various bitrates, depending on the end-user's situation. Additionally, streams shall be selected based on user preference (such as language, available listening time, etc.). The play-out module will support a large diversity of standardised 2D or 3D, e.g. 2.0, 5.1, 7.1+4, and binaural.

Reception and User Experience

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

- mobile device: iOS app (Elephantcandy)
- on the PC: HTML5 WebAudio API compatible Browser (BBC R&D)
- living room: AV receiver (Trinnov)

In all of 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

In addition, various audio rendering functionalities can be stored as customized profiles, representing different environment and listening conditions.

---

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

The screen shots of the iOS prototype app in Figure 4 exemplify the user-friendly implementation of these features following dedicated user requirement analysis [11] and usability evaluations.

Figure 4: Screenshots of the Pilot 1 prototype iOS app with three windows: a) element navigation b), profile selection (pre-sets of parameters), c) reproduction layout selection.

PILOT PHASE 1: A LIVE OBJECT-BASED RADIO DRAMA

The features that demonstrate the capabilities of the object-based approach were applied in BBC's live object-based radio drama, produced in the ORPHEUS studio, see Figure 5.

Multiple streams
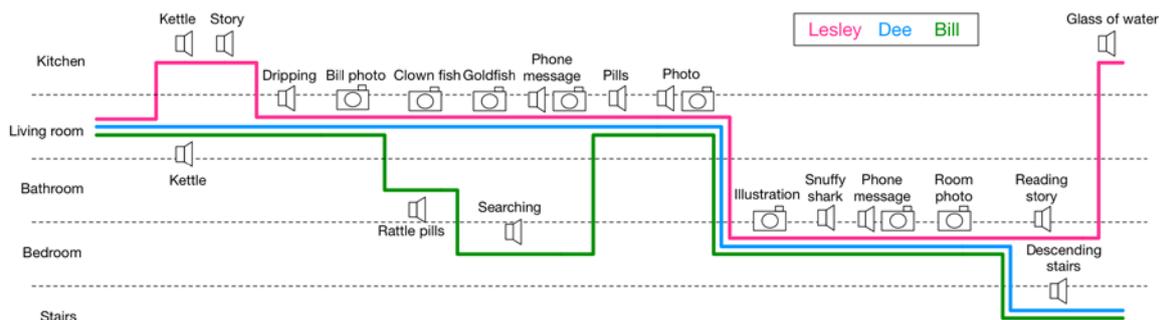
Figure 5: Draft diagram of character movements in the drama. Sound effects are marked with a speaker symbol, and images with a camera symbol

The three characters in the story were moving around the apartment into various rooms at different times. This introduced multiple parallel audio streams where action is happening simultaneously. The users had the option to choose which character they follow. Their experience of the story was therefore vary depending on who they follow, which gives a different perspective on what really happened.

Supplementary photos

Photos and images appeared at different points in the drama. These were be used to enhance the story by giving the audience a glimpse into the fictional world of the story, and to supplement the audio (e.g. displaying a text message). The images were prepared in advance and triggered in the studio at the right moments.
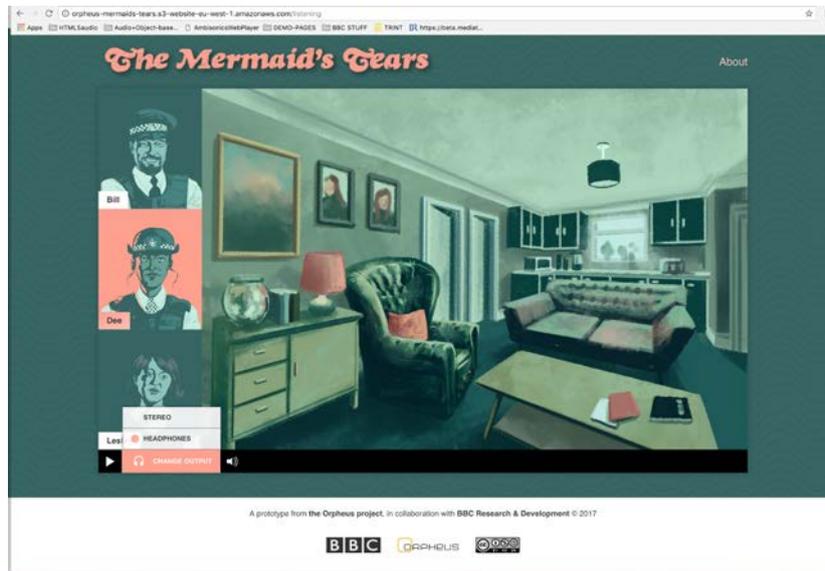
Figure 6: Web representation of BBC's radio drama in pilot 1

In addition to this radio drama, the ORPHEUS pilot phase 1 consists of four pre-produced examples of object-based audio content, represented in Elephantcandy's iOS ORPHEUS app (Figure 7):

- Heute im Stadion (BR-German): football match report with immersive audio, commentary enhancement and transcript

- Afternoon show (BBC-English): magazine excerpt with commentary enhancement and transcript

- Herbst (FH IIS-German/English): soundscape with immersive audio, commentary enhancement, transcript and individually adjustable audio object

- The Turning Forest (BBC-English/German): soundscape with immersive audio, commentary enhancement and transcript
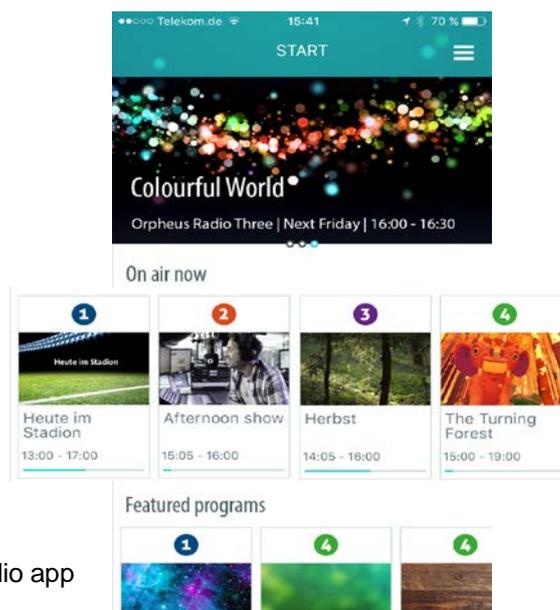

Figure 7: iOS ORPHEUS radio app

SUMMARY

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

- CREATION: Various object-based content productions completed by BR and BBC
- PRODUCTION: a variety of capturing, analysing, editing and mixing tools for object-based audio by partners FHG IIS, b<>com, IRCAM, as well as a fully ADM-compatible SEQUOIA DAW by partner MAGIX

- PLAYOUT: prototype of fully IP-connected radio studio with implemented object-based features by partner BBC R&D implementing technology by FHG IIS and MAGIX
- DISTRIBUTION: content server with encoding capabilities for object-based audio ADM plus PCM/AAC audio streams as well as MPEG-H by FH IIS, BBC R&D, IRT
- RECEPTION: implementation of object-based audio decoding and reproduction in browser, iOS app and CE AV-receiver by Elephantcandy and Trinnov, IRT and BBC R&D

The ORPHEUS project demonstrates 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, scalable and genuine IP-based. 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.

More details about the project can be found in the ORPHEUS Tonmeistertagung 2016 paper [12]. ORPHEUS publishes the project deliverables and the reference architecture guidelines on how to implement object-based audio chains on their website at http://orpheus-audio.eu/public-deliverables/.

ACKNOWLEDGMENTS

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

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

[9] 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.

[10] 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.

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

[12] Silzle, A., R. Sazdov, M. Weitnauer, et al. The EU Project ORPHEUS: Object-based Broadcasting – for Next Generation Audio Experiences. 29th Tonmeistertagung - VDT International Convention. 2016. Cologne, Germany.