MPEG-H AUDIO

MPEG-H Audio Production Workflow

The Myth of the 128 Objects

In this article, we'll compare the MPEG-H Audio System to another established Next Generation Audio option.

Version 1.0
Daniela Rieger

Date 2024-10-01

MPEG-H Audio Production Workflow – The Myth of the 128 Objects

A common question that often arises in talks with broadcasters and producers is why MPEG-H Audio only supports 16 audio channels. While it is indeed correct that an MPEG-H Audio delivery file consists of 16 channels (typically, channels 1-15 contain audio signals, channel 16 is the "Control Track" and reserved for metadata), this does not mean that MPEG-H Audio production is limited to those 16 elements. Originally, this channel count derives from the 16 audio channels typically used in broadcast workflows based on SDI.

Yes, the MPEG-H Authoring Plugin allows the creation of 15 audio components only. This may seem limiting at first glance, but it only indicates that, at the end of the production workflow, 15 dedicated objects or channels with labels, object panning, and user personalization can be used. During the mixing process, this simply means that the multitude of input tracks – including, if we assume a documentary mix, the music score, the individual foleys, ambiences, and other elements – are routed to buses, which then go into the MPEG-H Authoring Tool.

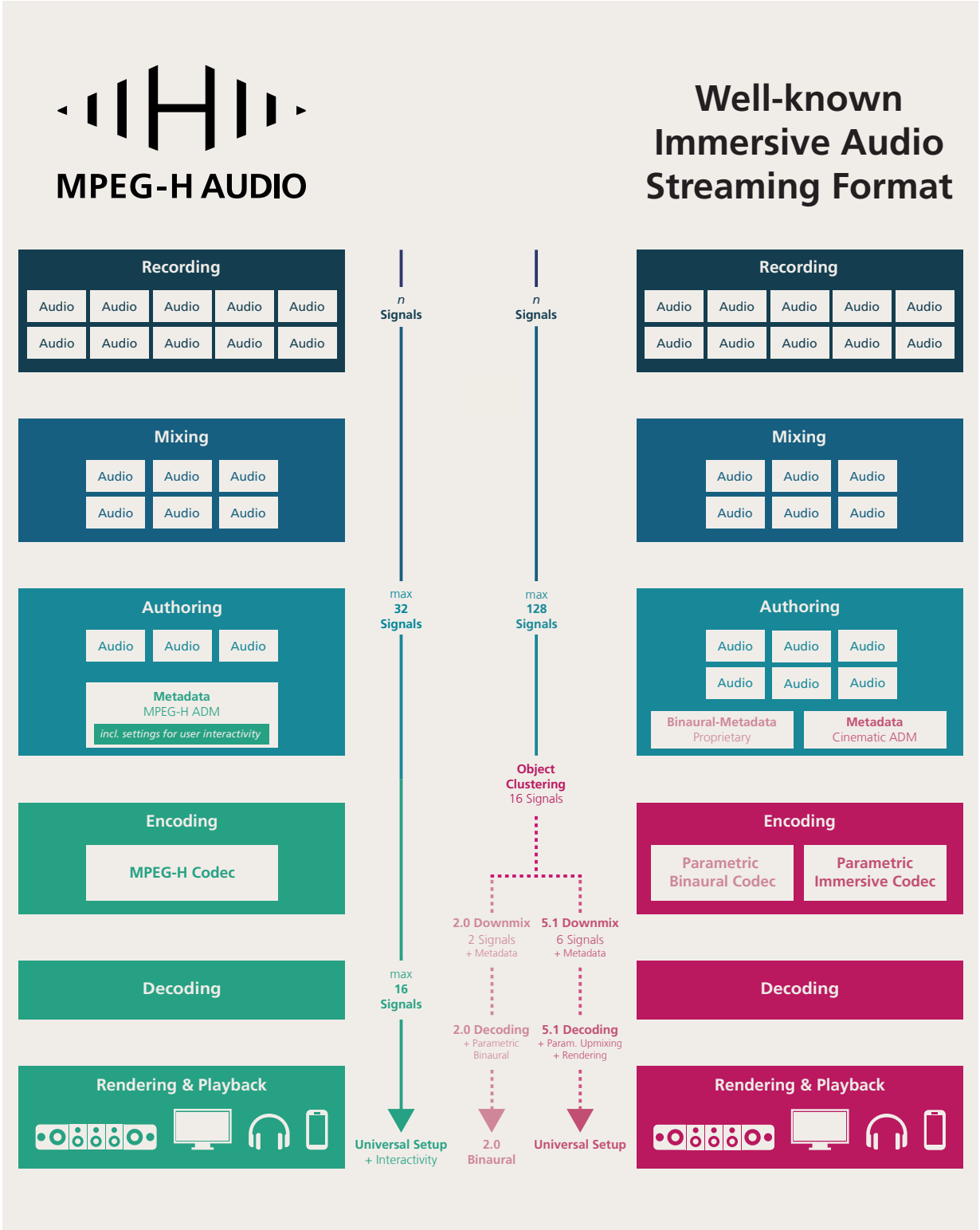
It is not necessary for every single sound element that is part of the mix to be a dedicated object in the MPEG-H Authoring stage. Working with pre-mixes and summing up and grouping elements that are similar has always been an essential step in audio mixing and does also apply to NGA productions, where, for example, audio elements are sorted into a bed and separate audio objects. In this respect, workflows have not changed much, and it is still good practice to create an ambience or foley bus, route all individual sounds to this track, and then route this bus to the MPEG-H Authoring Tool as an ambience or foley component.

The relevant question to ask during production is: Does this sound element need to be a dedicated object? If an audio element is supposed to be controlled separately by the user or the user device, it should be an object. All other audio elements (without interactivity) can be mixed into beds.

- **Object – Example 1:** Elements that are meant to be manipulated by the user in a gradual process (volume increase and decrease)
- **Object – Example 2:** Elements that are part of an "either/or" selection (the English, German, and Spanish commentary of a sports mix which should only be played one at a time). These three elements must be individual objects/channels and have to be combined into a so-called "switch group". This switch group could be called "Commentary" and allow the user to select only one of the three commentators at a time during playback

As can be seen in the flowchart, with MPEG-H Audio, the clustering of the many input elements simply happens during the mixing stage under full control and in a process that is fully transparent to the producer, which ensures easy monitoring. When producing in another well-known immersive audio format (shown in the right column of the flowchart), objects are also clustered – the difference being that this occurs later in the chain and without the possibility for the producer to control or monitor the process.
The following flowchart illustrates what an MPEG-H Audio production from recording to playback looks like in comparison to a well-known immersive audio format.

MPEG-H AUDIO


MPEG-H AUDIO
Well-known Immersive Audio Streaming Format
Recording
Audio
Mixing
Audio
Authoring
Audio
Metadata
MPEG-H ADM
incl. settings for user interactivity
Encoding
MPEG-H Codec
Decoding
Rendering & Playback
n Signals
max 32 Signals
max 16 Signals
Universal Setup
+ Interactivity
n Signals
max 128 Signals
Object Clustering
16 Signals
2.0 Downmix
2 Signals
+ Metadata
5.1 Downmix
6 Signals
+ Metadata
2.0 Decoding
+ Parametric Binaural
5.1 Decoding
+ Param. Upmixing
+ Rendering
2.0 Binaural
Universal Setup
Recording
Audio
Mixing
Audio
Authoring
Audio
Binaural-Metadata
Proprietary
Metadata
Cinematic ADM
Encoding
Parametric Binaural Codec
Parametric Immersive Codec
Decoding
Rendering & Playback

Disclaimer

INFORMATION IN THIS DOCUMENT IS PROVIDED 'AS IS' AND WITHOUT ANY EXPRESS OR IMPLIED WARRANTIES, INCLUDING, WITHOUT LIMITATION, THE IMPLIED WARRANTIES OF MERCHANTABILITY AND FITNESS FOR A PARTICULAR PURPOSE.

All rights reserved. No part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from the publisher.

Product and corporate names may be trademarks or registered trademarks of other companies. They are used for explanation only, with no intent to infringe.

Copyright (c) 2024 Fraunhofer-Gesellschaft zur Foerderung der angewandten Forschung e.V. as legal entity for its Fraunhofer Institute for Integrated Circuits IIS ('Fraunhofer').

MPEG-H AUDIO