**The Basis for your MPEG-H
Production Tool Solution**

The MPEG-H Production Library

MPEG-H Audio makes it possible for listeners to interact with individual audio elements. They can choose from various presets, languages, and commentators, change the position of audio objects in the room, and even benefit from accessibility features such as enhanced dialogue and audio descriptions.

Producers of MPEG-H content define the range and scope of all personalization options in a process called "authoring". During this process, they create the metadata required for user interactivity and automated reproduction optimization on individual devices. The result of the authoring process is a so-called "MPEG-H Scene". During production, it is the central representation of MPEG-H NGA authorings and enables monitoring, export, and import.

The MPEG-H Production Library is a toolbox which enables the creation of unique products. It comprises everything that is required for the authoring and monitoring of metadata. This includes tools for measuring and monitoring the loudness of MPEG-H content, which helps meet loudness regulations. The MPEG-H Production Library offers the export and import of MPEG-H Scenes in various MPEG-H Master formats including open formats like ADM. Exports are ready for emission and serve as input for MPEG-H encoders. Its documentation includes code snippets, use case examples, and a Command Line Interface (CLI) demo implementation example.

Specifications

Supported Platforms and Library Characteristics
- macOS: x86 and Apple silicon,
  Windows: x86, Linux: x86 and ARM
- static library without stl
- object-oriented C++ 98

OSS Dependencies
Dependent on the feature set used, the MPEG-H Production Library has dependencies to:
- libxml2
- zlib
- libxslt
- asdcplib
- ONNX Runtime

Authoring & Monitoring

Components
Components are the smallest adressable units of an MPEG-H Scene. They consist of audio tracks and associated metadata. Components that do not change their position during playback are usually channel-based components or static objects. This applies to both, multichannel audio (e.g. 5.1+4H bed) and mono signals such as mono commentary. Components that are intended to change their position over time or can be moved by the consumer are referred to as dynamic objects (e.g. birds, audio description). The MPEG-H Production Library enables the definition of both kinds of objects and also ensures that all necessary characteristics are assigned.

Switch Groups

Components that are not permitted to be played simultaneously, but have either/or characteristics, can be placed in a switch group. Such a switch group could, for instance, be used to define several language versions of a commentary track, each of which can be selected by the consumer or consumer electronics device during playback.

Presets

Presets combine the previously defined components and switch groups into predefined audio scenes. If multiple presets are included, the consumer can choose between different representations of the audio content. The MPEG-H Production Library helps create different mixes for certain requirements (e.g. ambience only, home and away team).

Monitoring Layouts

The MPEG-H Production Library provides a preview of the presets and interactivity settings and makes it possible to listen to an MPEG-H Scene on various output speaker layouts. During monitoring, the producer can switch between playback formats. This enables a preview of the decoder output. The monitoring layouts range from legacy layouts (Mono, Stereo, 5.1 etc.) to immersive layouts (5.1+4H, 7.1+4H, 22.2 etc.) and binaural rendering for headphones.

I/O

ADM

The Audio Definition Model (ADM) according to ITU-R BS.2076 defines an open metadata format for production, exchange, and archiving of NGA content in file-based workflows. Its comprehensive metadata syntax allows describing many types of audio content including channel-, object-, and scene-based representations for immersive and interactive audio experiences.

S-ADM

A serial representation of the Audio Definition Model (S-ADM) is specified in ITU-R BS.2125 and defines a segmentation of the original ADM for use in linear workflows such as real-time production for broadcasting and streaming applications.

MPEG-H BWF/ADM and S-ADM

An MPEG-H Broadcast Wave Format with embedded ADM metadata (BWF/ADM) file is a multichannel wave-file which contains all audio and metadata of the MPEG-H Scene. The exported BWF/ADM file is compliant with the MPEG-H ADM Profile. MPEG-H profiled S-ADM is defined in a corresponding manner.

MPF

An MPEG-H Production Format (MPF) file is a multichannel wave-file which contains all audio and metadata of the MPEG-H Scene. The metadata is stored in the control track, which is one of the audio tracks of the multichannel wave-file and contains a time-code-like signal that is robust against sample rate conversions or level changes.

MPEG-H Production Template XML

An MPEG-H Production Template XML is a representation that consists of the static attributes of an MPEG-H Scene. The XML can be used as a scene template for live authoring devices.

Supported Export Formats

- MPEG-H BWF/ADM
- MPF
- MPEG-H compliant S-ADM (ITU-R BS.2125) byte blobs
- MPEG-H Production Template XML

Supported Import Formats

- MPEG-H Masters
  - MPEG-H MPF
  - MPEG-H BWF/ADM
- BWF/ADM according to Dolby Atmos Master ADM Profile Specification
- IMF/MXF/IAB (SMPTE ST 2067-201)
- DCP/MXF/IAB (SMPTE ST 429-18)
- MPEG-H compliant S-ADM (ITU-R BS.2125) byte blobs
- MPEG-H Production Template XML

Real-Time Transport

The MPEG-H Production Library exports the metadata in real time, tightly coupled with the audio signals and synchronized with the video signal on any of the connections that are commonly used in linear productions, such as SDI, MADI, and IP.

To ensure the integrity of metadata in an SDI-environment in any production step, the metadata is delivered in the control track. The control track is a time-code-like audio signal and can be treated as a regular audio channel.

This ensures the synchronization of metadata with the associated audio and video signals. The control track is robust enough to survive A/D and D/A conversions, level changes, sample rate conversions, and frame-wise editing. It does not force audio equipment to be put into data mode or non-audio mode in order to pass through.

Authoring Features

Loudness Measurement

The MPEG-H Production Library includes loudness measurement for file-based offline workflows. The loudness measurement according to ITU-R BS.1770-4 for presets and components is performed automatically during export or can be triggered manually.
The MPEG-H Production Library provides programm-based as well as anchor-based loudness measurement functionalities.

Loudness Metering

The MPEG-H Production Library includes methods for live loudness metering tailored to real-time workflows. The loudness metering provides short-term loudness, momentary loudness, and integrated loudness according to ITU-R BS.1770-4.

Dynamic Gain Sequences

Dynamic gain sequences allow content creators to create and transmit volume automation to control component levels in the renderer of the consumer's playback device. It can be used to provide different dynamic mixes of components, depending on the playback scenario selected by the consumer or CE device. The MPEG-H Production Library writes the gain sequence into the dynamic metadata of all export formats.

Adaptive Background Attenuation (ABA)

ABA is a method which ensures that background sounds are lowered where they would interfere with speech intelligibility. The reduction works adaptively and only when dialog is present. In passages with no active speech, the background sounds remain unchanged. The MPEG-H Production Library provides offline and real-time ABA functionality.

Dialog+

Dialog+ is a dialogue enhancement technology based on a deep learning model. It separates a full stereo mix into a dialogue and a non-dialogue component. Afterwards, the relative level of these components can be adjusted by attenuating the background. The background attenuation works statically, dynamically, or through a combination of both.

Downmix

The MPEG-H Production Library supports generic downmixing to all supported target layouts. It also includes customizable downmix options enabling content-specific downmixing that is configurable for each layout. Through this, the producer has full control over the mix.

Video Frame Alignment

The MPEG-H Production Library makes possible the alignment of the dynamic metadata to video frame rates, for instance 50 fps and 59.94 fps. This enables video-frame accurate editing.

Contact

Fraunhofer Institute for Integrated Circuits IIS

Management of the institute
Prof. Albert Heuberger
(executive)
Prof. Bernhard Grill
Prof. Alexander Martin

Am Wolfsmantel 33
91058 Erlangen, Germany
Phone +49 9131 776-0
info@iis.fraunhofer.de

Mandy Garcia
Phone +49 9131 776-6178
mandy.garcia@iis.fraunhofer.de

www.mpeg-h.com
www.iis.fraunhofer.de