PRESS PAPER

MPEG-H AUDIO

THE NEXT-GENERATION AUDIO SYSTEM
FOR IMMERSIVE AND PERSONALIZED SOUND
IN BROADCAST AND STREAMING

Fraunhofer IIS has developed this system, based on the MPEG-H 3D Audio standard [1], to offer an enhanced sound experience for broadcast and new media services such as UHDTV, immersive music services, 4K video streaming or Virtual Reality.
## Key Benefits

### Immersive Audio
- Be in the middle of the action: The MPEG-H Audio system delivers enveloping immersive sound from all around the listener
- Be closer to your idol: Immersive music services powered by MPEG-H bring fans and artists closer together. Listeners will be immersed in sound from every direction as intended by the content creator

### Personalized Sound and Improved Accessibility
- Select your preferred audio presentation from several preconfigured versions in one stream
- Seamlessly switching between languages
- Adapt the audio mix to your needs and preferences
- Enjoy an enriched sports experience – multiple commentaries, home and away team announcers, venue sound like being at the event
- Benefit from enhanced speech intelligibility with Dialog+ and advanced audio description

### Ease of Use Proven on-Air 24/7 and in Trials
- The MPEG-H Audio system is proven to work with today’s streaming and broadcast workflows. It has a complete, free authoring tool chain and it does not require video frame alignment or special transmission paths in production.
- Immersive sound can be played back over TV sets, loudspeakers, headphones, or MPEG-H equipped soundbars.

### Universal Delivery
- Enjoy MPEG-H everywhere over multiple platforms – on the go – in your living room – in your home theatre – in the car
- One production, one stream to all devices – always the best possible sound experience

### Trademark Program
- To indicate to consumers that products implementing the MPEG-H Audio system will interoperate correctly, Fraunhofer operates the MPEG-H Audio system trademark program

### International Standard
- The system is based on the MPEG-H 3D Audio standard from ISO/IEC MPEG, the international standards group responsible for many globally dominant media standards such as MP3, AAC, MPEG-1, MPEG-2, MPEG-4, AVC/H.264 and HEVC/H.265.
- The MPEG-H Audio system is included in the ATSC, DVB, TTA and SBTVD TV standards.
- It is the audio codec selected for ATSC 3.0 broadcasting in South Korea and is the first next-generation audio system used on-the-air since broadcasts began in South Korea on May 31, 2017.
IMMERSIVE AUDIO

The television industry and related standards bodies around the world are preparing for the delivery of UHD video through new standards such as ATSC 3.0, ISDB-Tb, and DVB-UHDTV. All of these standards include or are considering new next-generation audio (NGA) systems to provide additional features or performance beyond those offered today. MPEG-H Audio opens an entire next level of sound going beyond stereo and surround. With sound coming from above, a third dimension is added to the audio experience. The impression of height gives the listeners a much more realistic and natural sound experience. Based on the playback system, the soundscape can be extended with sounds from below like footsteps down on the floor.

MPEG-H Audio makes it possible to truly immerse and envelop each user in sound from all directions. Watching a live sports event in immersive sound allows viewers to hear the crowd cheering around them, they can experience the stadium announcer from above and hear the fans of both teams on different sides of the stadium – all that within the comfort of their living room. Listening to a live music concert streaming one can hear the audience from all around and hear the music from the front stage.

Immersive MPEG-H Audio content can be enjoyed by consumers with AVRs, smart speakers, immersive soundbars, in cars and on mobile devices.

MPEG-H Audio is the only next-generation audio system that can deliver immersive sound using three types of audio elements or any combination of those:

Channel-based: each transmission channel is associated with a precisely defined and fixed location of the loudspeakers relative to the listener. An immersive sound experience is usually created by adding two to four height loudspeakers to the traditional 5.1 or 7.1 surround configurations. An advanced example would be a 22.2 speaker configuration as specified for Japanese TV broadcasting.
**Object-based:** each individual audio object may be positioned in three dimensions independently of loudspeaker positions. The main difference between objects and channels is that the spatial position of an audio object can vary over time. The associated metadata enables the decoder to render the object to the final loudspeaker setup at the receiver side.

**Scene-based (or Ambisonics):** a sound scene is represented by a set of spherical harmonic coefficients that have no direct relationship to channels or objects, but instead describe the sound field. Scene-based audio is natively carried in MPEG-H Audio as Higher Order Ambisonics (HOA).

### PERSONALIZED SOUND AND IMPROVED ACCESSIBILITY

With its unique personalization features, MPEG-H Audio offers consumers great flexibility to actively engage with the content and adapt it to their own preferences. The easiest way to interact with the content is to select one of several different audio presentations, predefined by the content creator. Those are complete audio mixes with a descriptive label attached to them, for example »Standard TV mix«, »Dialog Enhanced Audio« or »Live Venue Sound«. Furthermore, simple adjustments are possible, such as increasing the dialogue volume in relation to other audio elements if enabled by the content creator. Advanced listeners can dive deep into the audio representation where audio elements of the mix may be selected and adjusted in their level and/or position.

A user interface displaying all personalization options is available on all MPEG-H Audio enabled devices or apps so that viewers can personalize their content using, for example, the remote control of the TV or the touch screen on a mobile device. With its safe and innovative system design, the MPEG-H Audio user interface will automatically adapt to the content creator’s intentions and only display the interactivity options currently available.
Dialog+
Dialog intelligibility is perhaps the most important aspect for broadcasters when offering a program to their audience. The MPEG-H Audio system enables the listeners to adjust the dialog level to their preference under the full control of the broadcaster or content creator.

Several studies and experiments have shown that even for good mixes most viewers value the possibility of adjusting the dialog level relative to the rest of the audio mix. People with hearing impairments benefit from a higher level of the dialog. Similarly, a higher level for dialog can help viewers consuming the content in a different language than their mother tongue to better understand the dialog.

Another very important factor influencing the intelligibility is the environment where the content is consumed. With mobile devices it can be hard to understand the dialog in noisy environments (e.g., at a train station, in the plane, etc.). In these situations, the Dialog+ feature gives the audience the option to boost the dialog.

MPEG-H Audio enables this by transmitting each dialog element as an independent audio object. The viewer can therefore increase the dialog level for better intelligibility, or simply adjust it to create their own customized mix.

Fraunhofer’s Dialog+ algorithm also enables dialog level adjustment for legacy content, where the commentary or dialog are not available as individual audio elements. Automatic audio mixing may be an important step towards making audio descriptions affordable, even for low budget productions, and is a big time-saver for productions that have to go on-air quickly. At the same time, the advanced MPEG-H Audio metadata gives broadcasters the possibility to carefully control each feature they offer to their viewers.

By offering personalization features, broadcasters can easily comply with accessibility regulations.
EASE OF USE PROVEN ON-AIR 24/7 AND IN TRIALS

The MPEG-H Audio System was designed to enable the most advanced features in existing and future broadcast and streaming systems. Commercially on-air 24/7 in South Korea since 2017 [2] and proven in numerous broadcast tests around the world, broadcast specialists have experienced the easy integration into their existing SDI-based workflows.

During the Eurovision song contest 2018 and 2019 the European Broadcasting Union EBU and Fraunhofer IIS have successfully conducted trials for live production of immersive and interactive content using the MPEG-H TV Audio System [3].

In May 2018 France Télévisions conducted the first live broadcast over DVB terrestrial and satellite platforms with MPEG-H in Europe during the French Tennis Tournament in Paris. During the tournament, France Télévisions has enabled various personalization options as well as multiple languages together with the immersive sound for the live coverage of the court Philippe Chatrier. For the 2019 edition, France Télévisions has integrated the production tools from Jünger and Fraunhofer IIS in the OB van and conducted the first MPEG-H remote production [4].

During the 2018 Football World Cup in Russia, Fraunhofer IIS together with the Norwegian broadcaster NRK conducted a streaming trial using MPEG-H Audio to demonstrate the interactive features in combination with 5.1 surround sound [5].

In August 2018, the EBU, together with EBU members BBC, France Télévisions, IRT, RAI and ZDF, as well as a range of technology partners, conducted one of the most complex broadcast trials in Europe so far during the 2018 European Athletics Championships (EAC) in Berlin [6]. The trial demonstrated the world’s first live production and distribution of Ultra High Definition (UHD) content, with High Frame Rate (HFR), High Dynamic Range (HDR), and Next Generation Audio (NGA).

In October 2019, Globo, the largest media group in Brazil, successfully used MPEG-H Audio for a live broadcast over the enhanced ISDB-Tb system during one of the world’s biggest music festivals, Rock in Rio [7][13].
UNIVERSAL DELIVERY

Crafted for the latest ways to consume media, the MPEG-H Audio system brings the highest quality and user experience everywhere and on all devices from mobile phones to TV sets. The built-in renderer adapts the content to the playback capabilities of the device. There is no need to choose from complicated selections to find the content version which best works for the device.

There is only one kind of MPEG-H stream delivered through all streaming or broadcast platforms to all kinds of receiver and playback devices, from headphones to sound bars and discrete loudspeaker systems. The device will take the MPEG-H stream and automatically adapt the audio mix to the current device capabilities.

MPEG-H Audio provides not just an audio codec, but a complete integrated audio solution for delivering the best possible audio experience, independent of the reproduction system. It includes advanced Loudness and Dynamic Range Control (DRC). The loudness normalization feature ensures consistent loudness across programs and channels, for different presets and playback configurations, based on loudness information embedded in the MPEG-H Audio stream.
TRADEMARK PROGRAM

The MPEG-H Trademark is an indication to consumers that a product design has passed third-party testing to verify MPEG-H functions as it should, so they may purchase it with the confidence that it will work with other MPEG-H devices and deliver programming with advanced audio features.

Fraunhofer operates the MPEG-H Audio system trademark program, but does not test products. Authorized test centers selected by Fraunhofer will perform product testing according to the program specifications. Manufacturers are free to choose any test center.

The program has been designed to thoroughly test the user interfaces, input and output connections, audio controls and processing, and basic MPEG-H decoding functions of a product. This ensures connection issues with other trademarked components are minimized.

The program test suites verify that products include the features and profiles that broadcasters use in their programming, preventing consumer disappointment when their device does not support viewing programs with additional audio elements.

A full list of licensed original design manufacturers, licensed manufacturers and available devices can be found online at www.mpeg-h.com [16].

INTERNATIONAL STANDARD

MPEG-H 3D Audio or ISO/IEC 23008-3 is an open, international standard from MPEG (Moving Pictures Experts Group), the international standards body responsible for many globally dominant media standards such as mp3, AAC, MPEG-1, MPEG-2, MPEG-4, AVC/H.264 and HEVC/H.265. Already adopted by ATSC, DVB, TTA (Korean TV) and SBTVD (Brazilian TV) TV standards, the MPEG-H Audio system is the sole audio codec selected for ATSC 3.0 broadcasting in South Korea and the first NGA system used on-the-air since broadcasts began in South Korea on May 31, 2017.

The performance of MPEG-H Audio was carefully evaluated by MPEG and documented in two MPEG Verification Test Reports [14][15]. With more than one million subjective ratings from eight independent test labs (Fraunhofer IIS, Sony, NHK, Gaudio, Nokia, Orange, Qualcomm, Dolby and ETRI), the MPEG-H 3D Audio standard is the best tested next-generation audio codec and has demonstrated excellent quality for all use cases evaluated.

MPEG-H Audio has been widely adopted and included in various standards worldwide. Currently, MPEG-H Audio is the only Next Generation Audio system standardized for broadcast, streaming, hybrid and VR applications:
ATSC: The Advanced Television Systems Committee has successfully included MPEG-H Audio in its ATSC 3.0 suite of standards as ATSC Standard A/342-3 [2]. The corresponding transport layer signaling is specified in ATSC A/331 [8].

DVB: DVB has also selected and included MPEG-H Audio in the specification ETSI TS 101 154 v2.3.1 [9] defining the usage of audio and video codecs for DVB systems. The proper signaling for MPEG 2 TS DVB systems was specified in ETSI EN 300 468 [18].

TTA (South Korea): The Telecommunications Technology Association (TTA) has selected MPEG-H Audio as the sole audio system for ATSC 3.0 in South Korea, as specified in TTAK-KO-07.0127 [10].

SBTVD: The SBTVD (Sistema Brasileiro de Televisão Digital / Brazilian Digital Television System) Forum has adopted MPEG-H Audio inclusion into the ABNT specification for TV 2.5 in Brazil, ABNT NBR 15602-2 [11]. The additional signaling for transport layer is specified in ABNT NBR 15603 [19] and the MPEG-H Audio receiver specification is provided in ABNT NBR 15604 [20].

SCTE: SCTE/ISBE has included MPEG-H Audio in its coding and carriage suite of standards for cable television, ANSI/SCTE 242-3 [21] and ANSI/SCTE 243-3 [22].

China: MPEG-H 3D Audio has been chosen as the 3D Audio standard for the upcoming Chinese UHDTV services and it is in the final stage of the standardization process [23].


CTA: CTA has included MPEG-H Audio into their specification: CTA-CEB32.5 Recommended Practice for ATSC 3.0 Television Sets, Audio [26].

3GPP: 3GPP has selected MPEG-H Audio as the only audio format for 360° video streaming services over 5G within Release-15 of the specifications, TS 26.118 3GPP Virtual reality profiles for streaming applications [12].

VRIF: The VR Industry Forum (VRIF) has selected MPEG-H Audio as the only audio codec for VR streaming applications and specifies MPEG-H Audio into their VR Industry Forum Guidelines [27].

MPEG OMAF: MPEG has adopted MPEG-H Audio as the only immersive audio system for Omnidirectional media applications (VR/AR/360-degree video) [28].

All major OTT Specifications include MPEG-H Audio:

- MPEG CMAF: ISO/IEC 23000-19:2020, Information technology — Multimedia application format (MPEG-A) — Part 19: Common media application format (CMAF) for segmented media [29];
- CTA WAVE: CTA: CTA-5001, Web Application Video Ecosystem – Content Specification [30];
- DTV Play: The SBTVD Forum has added MPEG-H Audio support into the Profile D set of specification for DTV Play, ABNT NBR 15606-1:2018 [31].
- HbbTV: HbbTV 2.0.2 Specification (ETSI TS 102 796): Hybrid Broadcast Broadband TV [32];
- DASH-IF: DASH-IF Guidelines for Implementation: DASH-IF Interoperability Point for ATSC 3.0 [33];
- DVB DASH: MPEG-DASH Profile for Transport of ISO BMFF Based DVB Services over IP Based Networks [34].
REFERENCES


[12] ETSI TS 126 118 V15.0.0 (2018-10), 5G; 3GPP Virtual reality profiles for streaming applications (3GPP TS 26.118 version 15.0.0 Release 15): https://www.etsi.org/deliver/etsi_Ts/126100_126199/126118/15.00.00_60/ts_126118v150000p.pdf
[13] Shows do Rock in Rio transmitidos com tecnologia desenvolvida pelo Fraunhofer IIS:

[14] N16584, MPEG-H 3D Audio Verification Test Report:
http://mpeg.chiariglione.org/standards/mpeg-h/3d-audio/mpeg-h-3d-audio-verification-test-report


[16] MPEG-H Audio Trademark Program:


[18] ETSI EN 300 468 V1.16.1 (2019-08), Digital Video Broadcasting (DVB); Specification for Service Information (SI) in DVB systems:
https://www.etsi.org/deliver/etsi_en/300400_300499/300468/01.16.01_60/en_300468v011601p.pdf


https://webstore.ansi.org/standards/scte/ansiscte2422017-1710634


[23] MPEG-H selected for chinese 3D audio standard:

https://www.itu.int/dms_pubrec/itu-r/rec/bs/R-REC-BS.1196-7-201901-S!!PDF-E.pdf

[25] Ultra HD Forum Guidelines:

[26] CTA-CEB32.5 Recommended Practice for ATSC 3.0 Television Sets, Audio, June 2017:
https://shop.cta.tech/products/recommended-practice-for-atsc-3-television-sets-audio

https://www.iso.org/standard/73310.html

[29] ISO/IEC 23000-19:2020, Information technology — Multimedia application format (MPEG-A) — Part 19: Common media application format (CMAF) for segmented media:
https://www.iso.org/standard/79106.html


[32] HbbTV: HbbTV 2.0.2 Specification (ETSI TS 102 796): Hybrid Broadcast Broadband TV:
https://www.etsi.org/deliver/etsi_ts/102700_102799/102796/01.05.01_60/ts_102796v010501p.pdf

[33] DASH-IF: Guidelines for Implementation: DASH-IF Interoperability Point for ATSC 3.0:
https://dashif.org/docs/DASH-IF-IOP-for-ATSC3-0-v1.1.pdf

[34] Digital Video Broadcasting (DVB): MPEG-DASH Profile for Transport of ISO BMFF Based DVB Services over IP Based Networks:
https://www.etsi.org/deliver/etsi_ts/103200_103299/103285/01.03.01_60/ts_103285v010301p.pdf
ABOUT FRAUNHOFER IIS

Fraunhofer IIS, based in Erlangen, Germany, is the largest institute within Fraunhofer-Gesellschaft, Europe’s leading application-oriented research organization.

For over 30 years, the institute’s Audio and Media Technologies division has been shaping the globally deployed standards and technologies in the fields of audio coding and moving picture production. Fraunhofer IIS systems and tools help create, transmit and provide excellent audio and video content as well as enable high-quality real-time communication. Today, almost all computers, mobile phones and consumer electronic devices are equipped with technologies from Erlangen and are used by billions of people around the world every day.

It all started with the creation of mp3, then evolved with the co-development of AAC and HE-AAC. Now the fourth generation of best-in-class audio technologies – MPEG-H Audio, EVS, LC3/LC3plus and xHE-AAC – elevates the media experience to new heights. In terms of audio signal processing, Symphoria and the Sonamic product family provide enveloping and enhanced sound in cars, while the upHear product family dramatically improves 3D audio playback or recording quality of professional and consumer devices. Fraunhofer technologies also power digital radio: first and foremost in the form of the ContentServer, combining audio encoding, multimedia data management and multiplexing. In the field of moving picture technologies, establishing the Digital Cinema Initiative test plan boosted the creation of professional tools for digital film and media production, such as easyDCP, Realception and JPEG XS.

The interdisciplinary team transforms science into best-in-class applications with new functionalities for end users as well as optimum efficiency, reducing transmission costs while increasing reliability. Always taking into account the demands of the market, Fraunhofer IIS develops technology that makes memorable moments.

We hear the bigger picture.

Copyright © May 2021 Fraunhofer-Gesellschaft

For more information:
audio-info@iis.fraunhofer.de
www.mpeg-h.com.de

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. INFORMATION IN THIS DOCUMENT IS OWNED AND COPYRIGHTED BY THE FRAUNHOFER-GESELLSCHAFT AND MAY BE CHANGED AND/OR UPDATED AT ANY TIME WITHOUT FURTHER NOTICE. PERMISSION IS HEREBY NOT GRANTED FOR RESALE OR COMMERCIAL USE OF THIS SERVICE, IN WHOLE OR IN PART, NOR BY ITSELF OR INCORPORATED IN ANOTHER PRODUCT.