

Audio coding specifications for ISDB-T3

23rd September 2026

Toru Chinen[†]

Adrian Murtaza^{† †}

Satoshi Teshigawara^{†††}

[†] Principal Researcher, Spatial Acoustic Technology Department, Sony Corporation.

Chair, Audio Coding Technology WG, ARIB

^{††} Chief Technology and Standardization Manager, Fraunhofer IIS

^{†††} Senior Manager, Lead, Devices & AV Ecosystem, Licensing Sales, Dolby Japan K.K

ARIB STD-B32 Part 2 Audio Coding Systems

The following audio systems are specified in this standard.

- MPEG-2 AAC Digital Broadcasting
- MPEG-4 AAC Advanced Broadcasting
- MPEG-4 ALS Lossless Audio Transmission
- MPEG-H 3D Audio Next-Generation Services for Advanced Broadcasting
- AC-4 Next-Generation Services for Advanced Broadcasting

MPEG-H 3D Audio

MPEG-H 3D Audio is an object-based audio system based on the **open international standard** ISO/IEC 23008-3.

It provides high-quality and efficient delivery of **immersive and personalized sound experiences**.

The MPEG-H Audio system offers advanced personalization and accessibility features, while preserving the **full control over interactivity options** in the hands of the content creator.

With support for seamless configuration changes, ad-insertion and metadata authoring during live production, MPEG-H Audio is providing the **most advanced user experience for live** broadcast and streaming applications.



MPEG-H 3D Audio

MPEG-H 3D Audio is the only standardized next generation audio codec with native support for **22.2 multichannel audio format, in Level 4 – as adopted in ARIB standard STD-B32 part 2 v4.1.**

The MPEG-H Audio system performance was carefully evaluated by MPEG, at **nine prestigious independent test labs** around the world:

- Fraunhofer IIS, Sony, NHK, Gaudio, Nokia, Orange, Qualcomm, Dolby and ETRI

The MPEG-H 3D Audio standard is **the most thoroughly tested** NGA codec and has been evaluated as the most advanced next generation audio system by **independent test laboratories in Brazil** (TV 3.0 Test Report).

LEVELS AND THEIR CORRESPONDING RESTRICTION FOR BASELINE AND LOW COMPLEXITY PROFILES

LEVEL	1	2	3	4	5
Max. sampling rate [kHz]	48	48	48	48	96
Max. number of core channels in compressed data stream	10	18	32	56	56
Max. number of decoder-processed core channels	5	9	16 ^(*)	28	28
Max. number of channels in referenceLayout	5	9	16 ^(*)	24	24

(*) The Baseline profile supports in Level 3 up to 24 objects if the additional complexity restrictions given in ISO/IEC 23008-3, subclause 4.8.2.5.2 [6]) are applied on the encoding process.

Services Enabled by MPEG-H (South Korea and Japan)

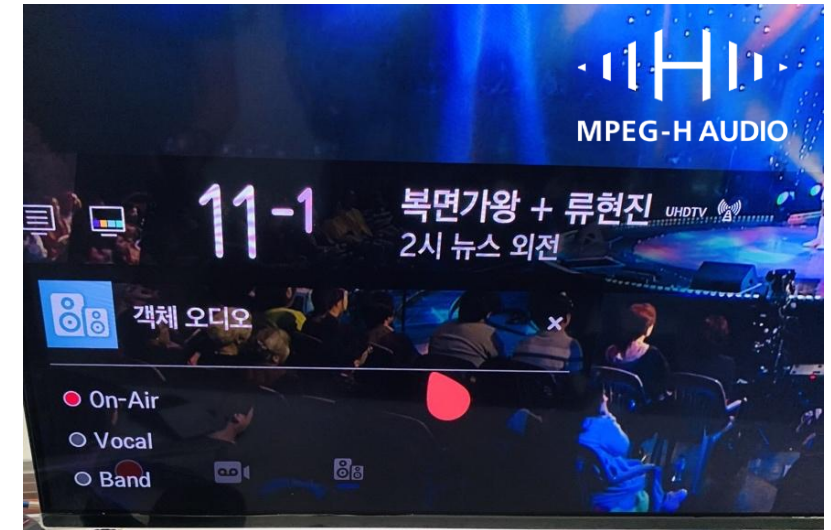
MPEG-H Audio has been selected as the **only audio system for the UHDTV terrestrial broadcast system in South Korea.**

- First regular terrestrial UHDTV service worldwide using ATSC 3.0 with MPEG-H Audio started in May 2017.
- Regular UHDTV service with broadcasts from: KBS, SBS and MBC.

NHK has conducted several end-to-end tests of the **Advanced ISDB-T broadcast system with MPEG-H Audio since 2023.**

Successful end-to-end verification held in **four cities (Fukuoka, Osaka, Tokyo, and Nagoya)** and proved that Advanced ISDB-T can simultaneously provide UHD and HD transmissions:

- UHD: VVC Video (9.0 Mbps) + MPEG-H Audio 22.2 (704kbps)
- HD: VVC Video (0.7 Mbps) + MPEG-H Audio stereo (96kbps)



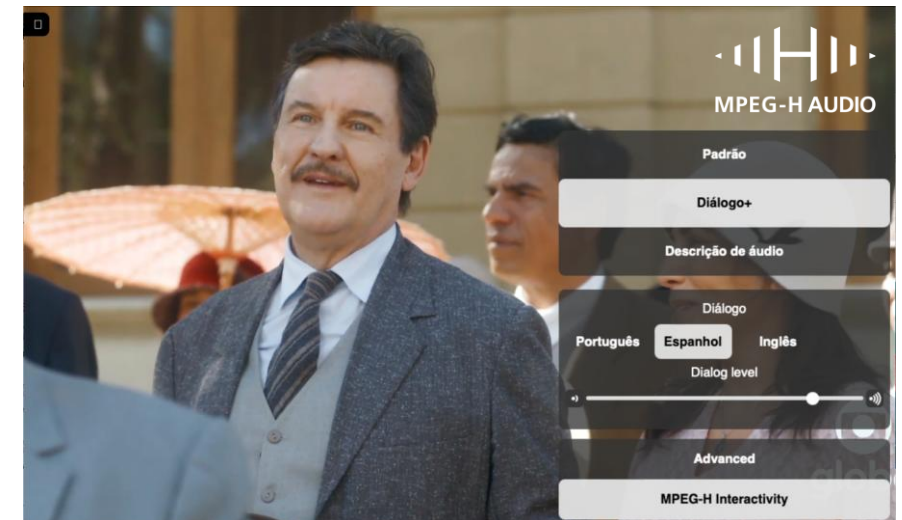
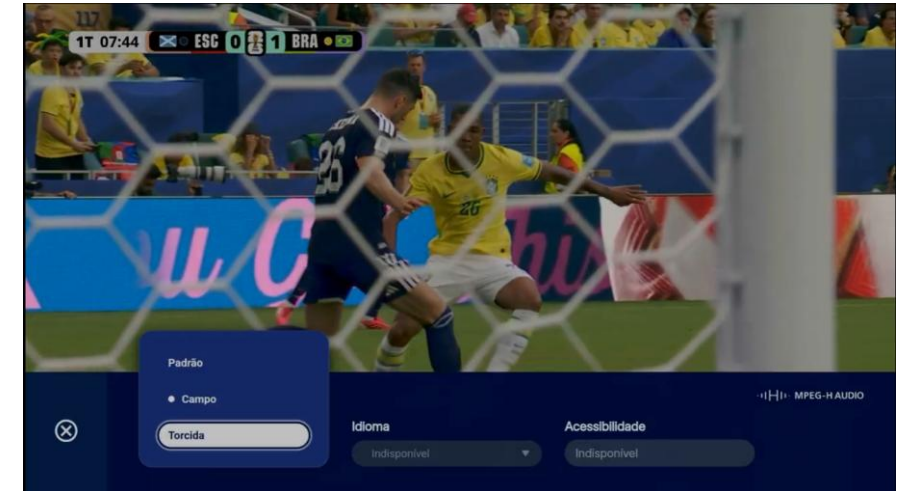
Services Enabled by MPEG-H (Brazil)

MPEG-H Audio has been selected after a detailed technical evaluation as the **sole audio codec for the TV 3.0 broadcast system in Brazil**.

- MPEG-H Audio is used 24/7 broadcast over ATSC 3.0 since June 2026 by Globo and public broadcasters such as EPC.
- Personalization and Accessibility options enabled for sports, and soap opera.

MPEG-H Audio is used also for broadcast over ISDB-T in Brazil (TV 2.5) by several major broadcasters:

- Globo (Rio de Janeiro, Sao Paulo, Recife),
- TV Cultura,
- Rede Amazonica and Amazon Sat.



AC-4

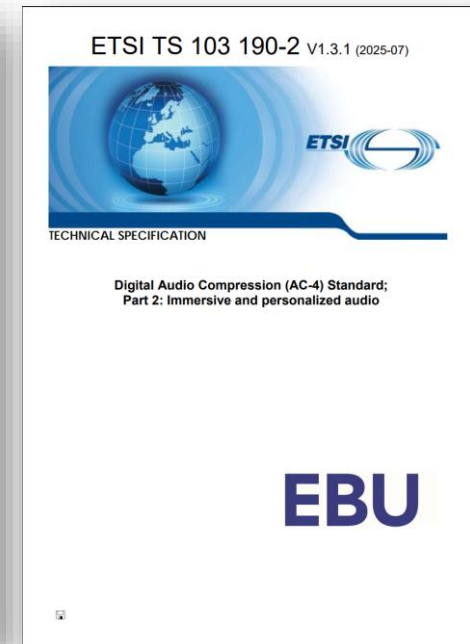
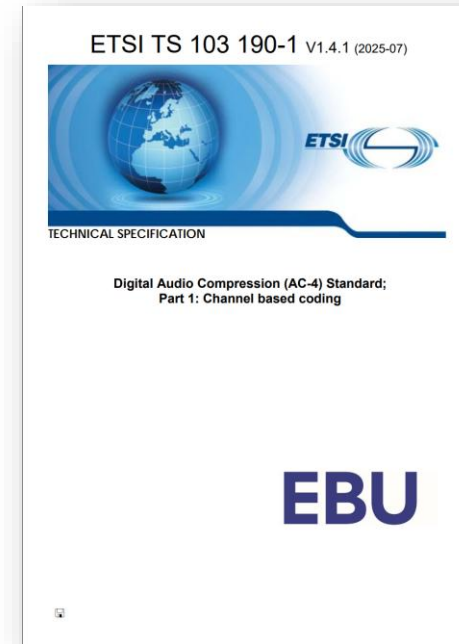
STD-B32 part 2 v4.1 references the following **international standards**

- ETSI TS 103 190-1 V1.4.1 (2025-07)
 - Digital Audio Compression (AC-4) Standard;
Part 1: Channel based coding
- ETSI TS 103 190-2 V1.3.1 (2025-07)
 - Digital Audio Compression (AC-4) Standard;
Part 2: Immersive and personalized audio

AC-4

- Is ground-breaking **Next Generation Audio**
- Introduces **intelligent loudness management**, and **A/V frame alignment**
- Provides **state of the art audio compression efficiency**

AC-4 has been approved for Japan's next-gen DTT broadcast standardization under **MIC**, through encoding-to-decoding tests contracted to **A-PAB** and supported by **NHK**, **NEC** and **Dolby**, including multiplexing with VVC video into MMT transport.



AC-4 in ARIB STD-B32 part 2 v4.1

Within STD-B32 part 2 v4.1, AC-4 supports the following **Object-Based Audio** input audio formats

- Mono
 - Stereo
 - Multichannel Stereo
 - 3.0 (3/0), 4.0 (2/2), 5.0 (3/2), 5.1 (3/2.1), 7.1 (5/2.1, 3/2/2.1), 5.1.2 (2/0/0-3/0/2-0.1), 5.1.4 (2/0/2-3/0/2-0.1), 7.1.4 (2/0/2-3/2/2-0.1)
 - Object
 - object
- ✂ Other formats, i.e., 3.0 (2/1), 4.0 (3/1), 6.1 (3/3.1) and 22.2 (3/3/3-5/2/3-3/0/0.2), can be encoded with AC-4 by treating certain channel(s) as object(s) and/or using tools such as Spatial Coding and A-JOC (Advanced Joint Object Coding).

Multichannel stereo audio mode notation used in the parentheses above:

- A channel configuration is represented as:
“Upper layer (front/side/rear) - Middle layer (front/side/rear) - Lower layer (front/side/rear).LFE”
where the numbers indicate the number of channels assigned to each position.
- For audio modes consisting only of a middle layer, the notation is simplified to: “Middle layer (front/side/rear).LFE”
- Furthermore, for middle-layer-only multichannel stereo audio modes without side channels, the notation is further simplified to: “Middle layer (front/rear).LFE”

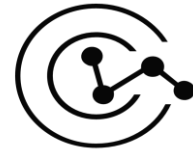
Services Enabled by AC-4

AC-4 brings NGA to ISDB-based Broadcast and OTT services

- **Dialogue Enhancement** - viewers to tailor the dialogue level to suit their individual preferences
- **Personalization** – a single bit stream can offer viewers a selection of multiple audio presentations
- **Immersive audio** – Dolby Atmos immersive audio at low data rates

Dialogue Enhancement

AC-4 gives the audience the ability to **adjust the level of dialogue** according to their personal preference. This includes support for both legacy (pre-mixed audio) and next-generation (music and effects plus dialogue) program sources.



Presentation Selection

New production techniques delivered by AC-4 satisfy **audience personalization** needs by delivering the individual audio program elements with metadata (crowd mics, alternative language track, announcer mics, music, and sound effects) all the way to the playback device, where those elements can be mixed on the preference of the viewer.



Immersive audio

AC-4 natively supports **immersive audio** to move sound around the audience in three-dimensional space.

 **Dolby** Atmos



Summary

ARIB STD-B32 Part 2 specifies:

- MPEG-2 AAC
- MPEG-4 AAC
- MPEG-4 ALS
- MPEG-H 3D Audio
- AC-4

Among them, MPEG-H and AC-4 provide key capabilities for next-generation advanced broadcasting services:

- Immersive Audio
- Object Audio
- Greater Flexibility
(e.g., Personalization, Dialogue Enhancement)

MPEG-H and AC-4 are expected to play an important role in future advanced broadcasting services.

Thank you for your attention