DiBEG
Digital Broadcasting Experts Group

SET eXPerience 2021

# Response to TV 3.0 Project CfP Phase-1 and 2 Audio Coding

DiBEG/Japan

# Proposal on Audio Coding by DiBEG

DiBEG
Digital Broadcasting Experts Group

1. General

We, DiBEG, or Digital Broadcasting Experts Group, of Japan, proposed Ph-1 and Ph-2 of the audio coding in response to the TV 3.0 CfP (Call for Proposal) issued by the SBTVD Forum. We describe the outline of our technical proposal of Ph-1 and Ph-2 for the audio coding.

2. Outline of audio coding

Our proposed audio coding is the latest audio coding with MPEG-H 3D audio which we are considering for the next-generation digital terrestrial TV in Japan (Advanced ISDB-T). We submitted the proposal about MPEG-H 3DA technology jointly with Fraunhofer, Ateme and ATSC. This document describes the outline of our proposed MPEG-H 3DA features.

# Ph-1 Proposal on Audio Coding

DiBEG
Digital Broadcasting Experts Group

We submitted the statement based on the CfP requirements.
The following is the outline of the statement for MPEG-H 3D audio.

AC1.1: channel based; 2.0, 5.1, 5.1+4H

Our proposed MPEG-H 3DA applies 22.2 multichannel sound and 7.1 (5.1+2H) multi-channel sound.

AC1.2: Object-based

Our proposed MPEG-H 3DA is complied Object-based.

AC1.3: Scene-based (HOA)

Our proposed MPEG-H 3DA complies with Scene-based (LC profile).

AC2.1: Switch components (audio objects/alternative full mix substreams)

Our proposed MPEG-H 3DA will be applied to Multi-lingual services and multi-comment services.

# Ph-1 Proposal on Audio Coding

DiBEG
Digital Broadcasting Experts Group

The following is the outline of the statement for MPEG-H 3D audio.

## AC2.2 & 2.3: Adjust object loudness and position

Our proposed MPEG-H 3DA renderer have capability to adjust audio object loudness.

## AC2.4: Enable interactivity when using external sound reproduction devices

This function is considered to be applied to Advanced ISDB-T.

## AC3.1 to 3.3: Audio description delivery in the same stream as the main audio, an alternative full mix and an additional audio object with metadata

Our proposed MPEG-H 3DA renderer enables to control sound object such as description.

# Ph-1 Proposal on Audio Coding

DiBEG
Digital Broadcasting Experts Group

The following is the outline of the statement for MPEG-H 3D audio.

## AC4.1: Emergency warning information audio description

This emergency description is included in Advanced ISDB-T. It is considered to be implemented in the Advanced ISDB-T.

## AC5.1: Flexible loudspeaker configuration render

Our proposed MPEG-H 3DA renderer have capability to reproduce a set of audio signals for several loudspeaker layouts.

## AC5.2: Binaural render

Our proposed MPEG-H 3DA renderer supports binaural processing.

## AC6.1: Consistent loudness across programs

This function is considered to be applied to Advanced ISDB-T by utilizing MPEG-H 3DA renderer and meta data.

# Ph-1 Proposal on Audio Coding

DiBEG
Digital Broadcasting Experts Group

The following is the outline of the statement for MPEG-H 3D audio.

## AC6.2: Consistent loudness after user interaction

Our proposed ADM or MPEG-H 3DA renderer have capability to adjust audio object loudness. These services are required to study.

## AC7.1: Seamless playback during configuration changes

Our proposed MPEG-H 3DA renderer have capability to reproduce several loudspeaker arrangements.

## AC7.2: Seamless playback during user interaction

Our proposed MPEG-H 3DA renderer have capability to reproduce sound object with or without several audio elements.

## AC7.3: Seamless playback during changes in production

Our proposed MPEG-H 3DA renderer have capability to reproduce sound with or without several sound objects.

# Ph-1 Proposal on Audio Coding

DiBEG
Digital Broadcasting Experts Group

The following is the outline of the statement for MPEG-H 3D audio.

## AC7.4: Seamless and sample-accurate stream splicing

Our proposed MPEG-H 3DA enables to improve the seamless and sample-accurate stream splicing or ad-insertion at any time instance, even if some of the streams come from different distribution platforms.

## AC8.1: Bit rate; kbps@MOS4/MUSHRA>80 or equivalent objective metric- 2.0, 5.1, 5.1+4H, 5.1+4H+2mono objects+2stereo object

Our proposed MPEG-H 3DA is effective for bit rate reduction, targeting 50% lower than AAC.

## AC9.1 & 9.2: Real-time encoding and

Our proposed MPEG-H 3DA adopts real-time encoding.

## AC10.1: A/V sync; frame-accurate

This function considers to apply to Advanced ISDB-T technology.

# Ph-1 Proposal on Audio Coding

DiBEG
Digital Broadcasting Experts Group

The following is the outline of the statement for MPEG-H 3D audio.

## AC11.1: VR/AR/XR/3DoF/6DoF support

The VR series are considered to be applied to Advanced ISDB-T.

## AC12.1: Interoperability with different distribution platforms

Our proposed MPEG-H 3DA will have suitable interoperability.

## AC13.1: Scalability

This function is considered to be applied to Advanced ISDB-T.

## AC14.1: Extensibility

This function is considered to be applied to Advanced ISDB-T.

# Ph-2 Proposal on Audio Coding

DiBEG
Digital Broadcasting Experts Group

The laboratory tests of Audio Coding in Ph-2 is conducted by Brazil side. DiBEG and Fraunfofer will support these laboratory tests.

DiBEG conduct the following items;

(1) Provision of Audio Test Contents
DiBEG have provided necessary audio contents for the laboratory tests.

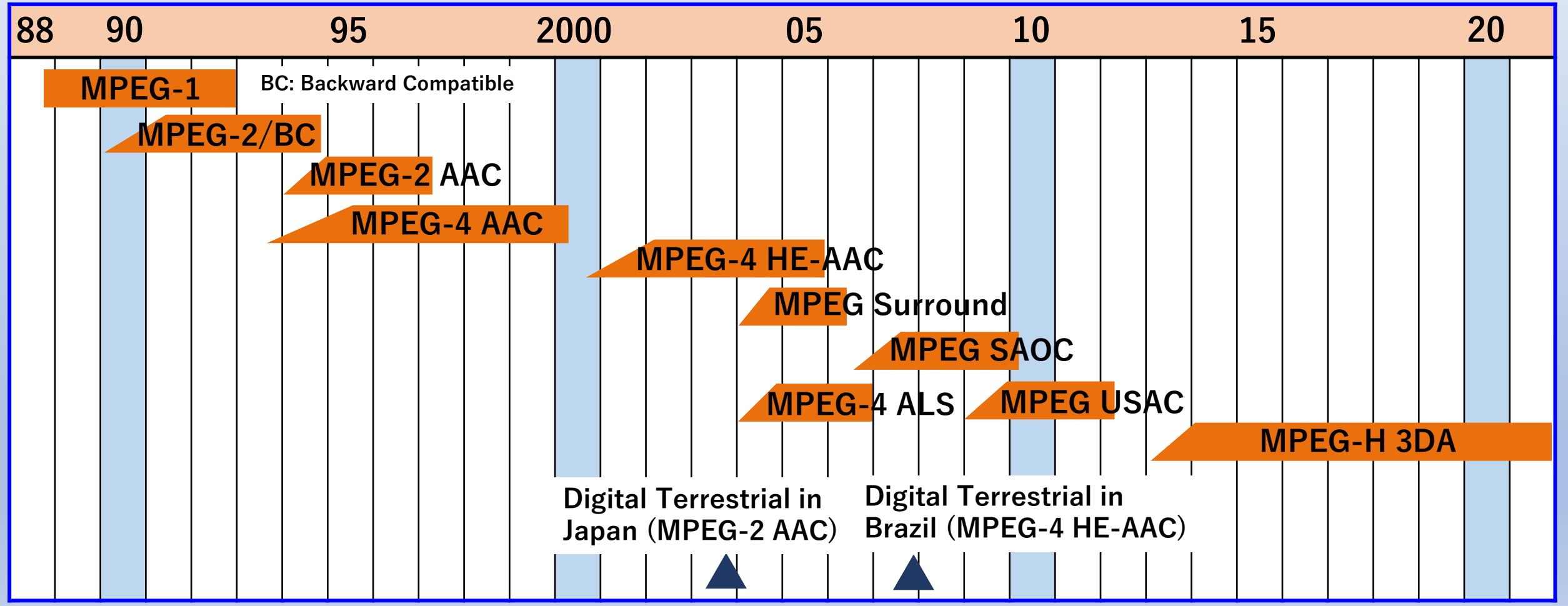
(2) Provision of further technical report (planned)
MPEG-H 3D audio coding will be applied to Advanced ISDB-T. DiBEG are studying MPEG-H 3DA technologies, and will support the evaluation test in Brazil. DiBEG are studying further technical report, if necessary.

(3) Development of MPEG-H Encoder and Decoder
We are developing the MPEG-H 3DA Encoder and Decoder for Advanced ISDB-T. The outline of MPEG-H 3DA technology describes in this document.

# History of MPEG Audio Coding

DiBEG
Digital Broadcasting Experts Group

We describe the outline and the history of MPEG audio coding technologies.
The road map of MPEG audio coding is shown.



# Outline of MPEG-H 3DA Coding

DiBEG
Digital Broadcasting Experts Group

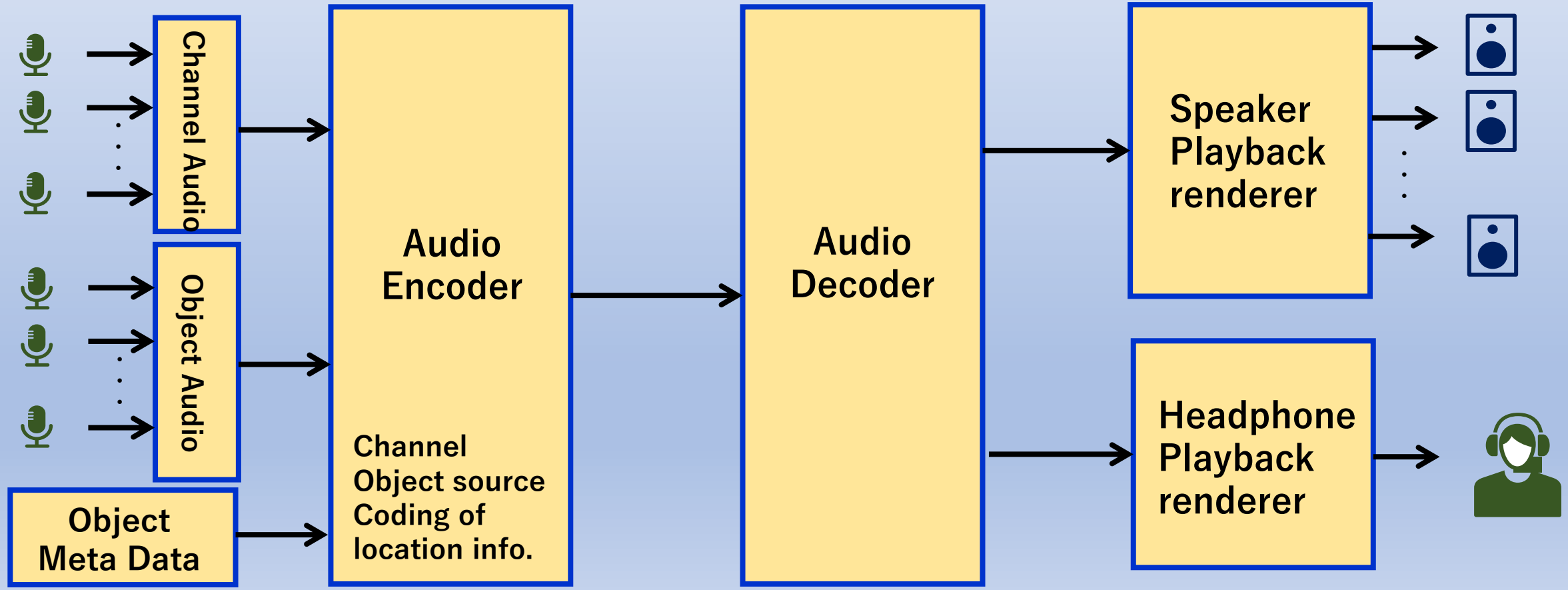
Signal flow between Encoding and Decoding of MPEG-H 3DA is shown.

## 3D Audio Technology

- Immersive sound such like theater
- Several kind of audio can be adjusted.

## Renderer

- Format Conv: 2ch/5.1ch/22.2ch
- Dialog emphasis
- Exchange of Dialog



# Features of MPEG-H 3DA Coding

DiBEG
Digital Broadcasting Experts Group

MPEG-H 3D audio has two standard specifications for broadcasting and streaming applications.

(1) Baseline Profile
(2) Low Complexity Profile

## Baseline Profile

The Baseline Profile is the recommended profile for all broadcast and streaming services. The Baseline profile is the best choice for electronics manufacturers to be incorporated in their devices.

## Low Complexity Profile

The Low Complexity Profile is a super set of the Baseline. It includes additional two coding tools for Higher-Order Ambisonics (HOA) and Linear Prediction Domain (LPD), which are not needed for majority of next generation audio broadcast and streaming applications.

# Features of MPEG-H 3DA Coding

DiBEG
Digital Broadcasting Experts Group

MPEG-H 3D audio has the following features;

## Immersive Audio

Distinguished from surround sound by expanding the sound image in the vertical dimension, immersive sound offers more enveloping and realistic experience.
The immersive sound is carried in three primary ways;

- Traditional channel-based sound where each transmission channel is associated with a studio loudspeaker position.
- Sound carried through audio objects, which may be positioned in three dimensions independently of loudspeaker positions
- Scene-based where a sound scene is represented by a set of coefficient signals that are the linear weights of spatial orthogonal spherical harmonics basis functions.

MPEG-H 3D audio is the codec that can provide immersive sound using any combination of the three well-established audio formats:

- Channel-based
- Object-based
- Higher-Order Ambisonics

# Features of MPEG-H 3DA Coding

DiBEG
Digital Broadcasting Experts Group

MPEG-H 3D audio has the following features;

## Personalization and Interactivity

The use of audio objects, usually in combination with channel-based audio, enables viewers to interact in new ways with the content and create a personalized listening experience. The MPEG-H 3D audio metadata carries all the information needed to enable viewers to manipulate the audio objects by attenuating or increasing their level, disabling them, or changing their position in the three-dimensional space. Additionally, the MPEG-H 3D audio metadata structures empower broadcasters to enable or disable interactivity options and to strictly set the limits in which the user can interact with the content.

## Advanced Accessibility

Using object-based audio, MPEG-H 3D audio offers advanced and improved accessibility services for enabling healing-impaired and visual-impaired audience to experience advanced next-generation audio features.

# Features of MPEG-H 3DA Coding

DiBEG
Digital Broadcasting Experts Group

MPEG-H 3D audio has the following features;

## Universal Delivery

The traditional TV broadcast environment uses a well-defined end-to-end solution to deliver audio content to the end user. Accordingly, it has been a good compromise to define a particular target loudness and dynamic range for this specific delivery channel and well-known type of sound reproduction system of the receiving device.
However, new types of delivery platforms and infrastructures have become significant and are constantly evolving. In a multi-platform environment, the same MPEG-H 3D audio stream is delivered through different distribution networks and is consumed on a variety of devices in different environments.

DiBEG
Digital Broadcasting Experts Group

# Muito obrigado!

**ARIB / DiBEG extend technical cooperation to Brazil adopting Next Generation Broadcasting Standard!**

di-jim3 @ arib.or.jp

ARIB
ASSOCIATION OF RADIO INDUSTRIES AND BUSINESSES

DiBEG
Digital Broadcasting Experts Group