INTERNATIONAL STANDARD

ISO/IEC 23008-3

First edition
2015-10-15

AMENDMENT 1
2016-08-01

Information technology — High efficiency coding and media delivery in heterogeneous environments —

Part 3: 3D audio

AMENDMENT 1: MPEG-H, 3D audio profile and levels

Technologies de l'information — Codage à haute efficacité et livraison des medias dans des environnements hétérogènes —

Partie 3: Audio 3D

AMENDEMENT 1: Niveaux et profil audio 3D MPEG-H

Reference number
ISO/IEC 23008-3:2015/Amd.1:2016(E)

ISO copyright office
Ch. de Blandonnet 8 • CP 401
CH-1214 Vernier, Geneva, Switzerland
Tel. +41 22 749 01 11
Fax +41 22 749 09 47
copyright@iso.org
www.iso.org

Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

The procedures used to develop this document and those intended for its further maintenance are described in the ISO/IEC Directives, Part 1. In particular the different approval criteria needed for the different types of ISO documents should be noted. This document was drafted in accordance with the editorial rules of the ISO/IEC Directives, Part 2 (see www.iso.org/directives).

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights. Details of any patent rights identified during the development of the document will be in the Introduction and/or on the ISO list of patent declarations received (see www.iso.org/patents).

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation on the meaning of ISO specific terms and expressions related to conformity assessment, as well as information about ISO's adherence to the World Trade Organization (WTO) principles in the Technical Barriers to Trade (TBT) see the following URL: www.iso.org/iso/foreword.html.

Amendment 1 to ISO/IEC 23008-3:2015 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

Information technology — High efficiency coding and media delivery in heterogeneous environments —

Part 3:
3D audio

AMENDMENT 1: MPEG-H, 3D audio profile and levels

Page 346

Add the following section after Clause 18.

19 MPEG-H 3D Audio Profile Definition

19.1 Profile: Main Profile

The Main Profile for MPEG-H 3D Audio contains all normative bitstream elements and normative decoder tools defined in MPEG-H 3D Audio specification.

That means that the following tools will be included Main Profile decoders:

— MPEG-H 3D Audio Core Decoder

— HOA Rendering

— SAOC 3D Renderer

— Static object metadata (MAE) and rendering

— Dynamic object metadata (OAM) and rendering

— Generic Loudspeaker Rendering/Format Conversion

— Immersive Loudspeaker Rendering/Format Conversion

— Binaural Rendering Time Domain and/or Frequency Domain

— H2B Binaural Rendering

— Loudness Metadata

— DRC processing

The following table specifies the levels of the *Main Profile.*

Mpegh3daProfile LevelIndication	Applicable Notes	Max. number of core channels	Max. sampling rate of core	Max number of loudspeaker output channels	Max. PCU in wMOPS[a]	Max. RCU
1		8	48000	8	**138**	
2		16	48000	16	**265**	
3	1) 2) 3)	32	48000	24	**448**	
4	1) 2) 3)	64	48000	24	**830**	
5	1) 2) 3)	128	96000	64	**3223**	
General restrictions for all levels:						
	— HOA: The number of active predictions must not be larger than $ceil\left(\frac{(N+1)^2}{4}\right)$ (NumActivePred in Table 127 Syntax of HOAPredictionInfo(DirSigChannelIds, NumOfDirSigs)). N is the HOA order. For the definition of global HOA parameters refer to 12.4.1.1. — The HOA order must not be larger than 3 for Level 1, 4 for Level 2, 5 for Level 3, 6 for Level 4 and 7 for Level 5 (see HoaOrder in Table 119 Syntax of HOAConfig()). — The number of input objects (for SAOC encoding) must not be larger than 2 times the maximum number of core coder channels — The number of predominant sounds of HOA must not be larger than 8 for Level 1, 10 for Level 2, 12 for Level 3, 14 for Level 4, and 16 for Level 5.					
Restrictions for specific levels:						
	1) SAOC: The maximum number of SAOC downmix channels is 32. SAOC objects must be grouped, i.e. a set of SAOC objects is mixed into a group of maximum of 8 downmix channels and not to any other downmix channel. IOCs must not be transmitted between SAOC objects different groups. 2) The maximum number of channels in each group with SignalGroupTypeChannels is 24, multiple such groups can exist 3) For DRC-1 and DRC-3 the maximum number of channel groups for each is 16.					
Note: Also, it is assumed that the both Binaural Renderers (TD and FD) are implemented. The total complexity may increase if only a single Binaural Renderer is available. The numbers for binaural processing are calculated on the basis of BRIR filters of 1 second length measured in a BS.1116 compliant room. [a] The maximum PCU numbers are based on theoretical calculations and estimations of the number of operations. They represent worst case total complexity numbers. All PCU figures are provided as informative data.						

19.1.1 Examples for Level 1 of Main Profile

Example 1:

8 input channels as a 7.1 mix are carried as channels and coded at a low bitrate. In the decoder a downmix is performed to 5.1 channels. Finally, a multi-band dynamic range compression is applied to the 6 loudspeaker output signals.

Decoder building block	Core Coder channels	Rendering	Domain switch	DRC	Post-processing	Total PCU in wMOPS
Description	8 (incl all tools) = 4 CPEs	8 ch -> 6 ch	6 ch FD-> TD	multi-band DRC 2	-none-	
PCU	46	5	9	2.2		62

Example 2:

A 2nd order HOA signal is carried in 8 core coder channels and is decoded to produce 9 HOA components.

The H2B binaural processing is applied to render the signal for a headphone output. Single band dynamic range compression is applied to the output.

Decoder building block	Core Coder channels	Rendering	Domain switch	DRC	Post-processing	Total PCU in wMOPS
Description	8 (including all tools) = 4SCE + 2CPE	4 Amb + 4 PS (HOA rendering matrix 9x8 not applied)	8 ch FD-> TD (if SBR, otherwise not applied)	DRC 2 full band	H2B-Binaural Rendering of 9 HOA compo-nents	
PCU	12.6+21.6 = 34.2	15	22/0	0.5	21	82/70

19.1.2 Examples for Level 2 of Main Profile

Example 1:

A 4th Order Higher Order Ambisonics (HOA) signal is coded at about 500 kbit/s, so no SBR is applied. The output domain of the core decoder is time domain so no domain switch is necessary. The HOA spatial decoder reproduces a 4th order HOA signal which is rendered to a 11.1 loudspeaker setup.

Decoder building block	Core Coder channels	Rendering	Domain switch	DRC	Post-processing	Total PCU in wMOPS
Description	8 (including all tools) = 2CPE + 4 SCE	4 Amb + 4 PS (HOA Decoding + Rendering to 11 Speakers)	-	DRC 2 full band		
PCU	8+19.4=27.4	24 + 13 = 37	0	0.5	0	65

Example 2:

A 4th Order Higher Order Ambisonics (HOA) signal is coded at about 250 kbit/s, so SBR is applied. The output domain of the core decoder is frequency domain and a domain switch is necessary. The HOA spatial decoder reproduces a 4th order HOA signal which is rendered to a 11.1 loudspeaker setup. Additionally 2 dialogue objects accompany the HOA scene.

Decoder building block	Core Coder channels	Rendering	Domain switch	DRC	Post-processing	Total PCU in wMOPS
Description	8 (HOA) = 2 CPE + 4 SCE plus 2 (Objects) = 2 SCE	4 Amb + 4 PS (HOA Decoding + Rendering to 11 Speakers) + 2 Objects	10 ch FD to TD	DRC 2 full band		
PCU	12.6 + 21.6 + 6.3 = 40.5	24 + 13 + 2 = 39	15	0.5	0	95

19.1.3 Examples for Level 3 of Main Profile

Example 1:

A 4th Order Higher Order Ambisonics (HOA) signal is coded at about 250 kbit/s, so SBR is applied. The output domain of the core decoder is frequency domain and a domain switch on the core coder transport channels is necessary. The HOA spatial decoder reproduces a 4th order HOA signal which is rendered to a 22.2 loudspeaker setup.

Decoder building block	Core Coder channels	Rendering	Domain switch	DRC	Post-processing	Total PCU in wMOPS
Description	8 (HOA) = 2 CPE + 4 SCE plus	4 Amb + 4 PS (HOA Decoding + Rendering to 22 Speakers)	8 ch FD to TD	DRC 2 full band		
PCU	12.6 + 21.6 = 34.2	24 + 26 = 50	12	1	0	97