

DRAFT AMENDMENT

ISO/IEC 23008-3: DAM 1

ISO/IEC JTC 1/SC 29

Secretariat: JISC

Voting begins on:
2015-05-11

Voting terminates on:
2015-08-11

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 médias dans des environnements hétérogènes

AMENDEMENT 1

ICS: 35.040

iTeh STANDARD PREVIEW
(standards.iteh.ai)
Full standard:
<https://standards.iteh.ai/catalog/standards/sist/97c9c2d8-a290-41aa-aaef-56e082f74de6/iso-iec-23008-3-2015-am1>
1-2016

THIS DOCUMENT IS A DRAFT CIRCULATED FOR COMMENT AND APPROVAL. IT IS THEREFORE SUBJECT TO CHANGE AND MAY NOT BE REFERRED TO AS AN INTERNATIONAL STANDARD UNTIL PUBLISHED AS SUCH.

IN ADDITION TO THEIR EVALUATION AS BEING ACCEPTABLE FOR INDUSTRIAL, TECHNOLOGICAL, COMMERCIAL AND USER PURPOSES, DRAFT INTERNATIONAL STANDARDS MAY ON OCCASION HAVE TO BE CONSIDERED IN THE LIGHT OF THEIR POTENTIAL TO BECOME STANDARDS TO WHICH REFERENCE MAY BE MADE IN NATIONAL REGULATIONS.

RECIPIENTS OF THIS DRAFT ARE INVITED TO SUBMIT, WITH THEIR COMMENTS, NOTIFICATION OF ANY RELEVANT PATENT RIGHTS OF WHICH THEY ARE AWARE AND TO PROVIDE SUPPORTING DOCUMENTATION.



Reference number
ISO/IEC 23008-3:2015/DAM 1:2015(E)

© ISO/IEC 2015

iTeh STANDARD PREVIEW
(standards.iteh.ai)
Full standard:
<https://standards.iteh.ai/catalog/standards/sist/97c9e3b8-a290-41aa-aaef-56e082f74de6/iso-iec-23008-3-2015-amd-1-2016>



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2015

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized otherwise in any form or by any means, electronic or mechanical, including photocopying, or posting on the internet or an intranet, without prior written permission. Permission can be requested from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 749 09 47
E-mail copyright@iso.org
Web www.iso.org

Published in Switzerland

Foreword

ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission) form the specialized system for worldwide standardization. National bodies that are members of ISO or IEC participate in the development of International Standards through technical committees established by the respective organization to deal with particular fields of technical activity. ISO and IEC technical committees collaborate in fields of mutual interest. Other international organizations, governmental and non-governmental, in liaison with ISO and IEC, also take part in the work. In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of the joint technical committee is to prepare International Standards. Draft International Standards adopted by the joint technical committee are circulated to national bodies for voting. Publication as an International Standard requires approval by at least 75 % of the national bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Amendment 1 to ISO/IEC 23008-3:201x was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

iTeh STANDARD PREVIEW
(standards.iteh.ai)

Full standard:
<https://standards.iteh.ai/catalog/standards/sist/97c9e3b8-a290-41aa-aaef-56e082f74de6/iso-iec-23008-3-2015-amd-1-2016>

Information technology — High efficiency coding and media delivery in heterogeneous environments — Part 3: 3D audio, AMENDMENT 1: MPEG-H, 3D Audio Profiles and Levels

Add the following to Section

1 MPEG-H 3D Audio Profile Definition

1.1 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 DIS [1].

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 Level/Indication	Applicable Notes	Max. number of core channels	Max. sampling rate of core	Max number of loudspeaker output channels	Max. PCU in wMOPS *)	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 $\text{ceil}\left(\frac{(N+1)^2}{4}\right)$ (NumActivePred in Table 127 Syntax of HOAPredictionInfo(DirSigChannelIds, NumOfDirSigs)). N is the HOA order. For the definition of global HOA parameters refer to Section 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.

*) 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.

Editors Notes:

It should be studied whether further restrictions can be applied to reduce the worst case complexity.

1.1.1 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 channels	Coder	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	-none-		

PCU	DRC 2				
	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 channels	Coder	Rendering	Domain switch	DRC	Post-processing	Total PCU in wMOPS
Description	8 (including all tools) = 4SCE + 2CPE	4 Amb PS rendering matrix 9x8 not applied	+ 4 (HOA > TD (if SBR, otherwise not applied))	8 ch FD- (if full band)	DRC 2	H2B-Binaural Rendering of 9 HOA components	
PCU	12.6+21.6 = 34.2	15		12/0	0.5	21	82/70

1.1.2 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 channels	Coder	Rendering	Domain switch	DRC	Post-processing	Total PCU in wMOPS
Description	8 (including all tools) = 2CPE + 4 SCE	4 Amb PS Decoding + Rendering to 11 Speakers)	+ 4 (HOA > TD (if SBR, otherwise not applied))	-	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 channels	Coder	Rendering	Domain switch	DRC	Post-processing	Total PCU in wMOPS
Description	8 (HOA) = 2CPE + 4 SCE plus 2 (Objects) = 2 SCE	4 Amb PS Decoding + Rendering to 11 Speakers)	+ 4 (HOA + 11 Speakers)	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

1.1.3 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 channels	Coder	Rendering	Domain switch	DRC	Post-processing	Total PCU in wMOPS
Description	8 (HOA) = 2CPE + 4 SCE plus	4 Amb PS Decoding + Rendering to 22 Speakers)	+ 4 (HOA + 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

1.1.4 Examples for Level 4 of Main Profile

1.1.5 Examples for Level 5 of Main Profile

2 References

[1] N14747, ISO/IEC JTC1/SC29/WG11 “Text of ISO/IEC 23008-3/DIS, 3D audio”

iTeh STANDARD PREVIEW
(standards.iteh.ai)
Full standard:
<https://standards.iteh.ai/catalog/standards/sist/97c9e3b8-a290-41aa-aaef-56e082f74de6/iso-iec-23008-3-2015-amd-1-2016>