

Third edition
2022-08

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

Part 3: 3D audio

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

Partie 3: Audio 3D



Reference number
ISO/IEC 23008-3:2022(E)

© ISO/IEC 2022



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2022

All rights reserved. Unless otherwise specified, or required in the context of its implementation, 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
CP 401 • Ch. de Blandonnet 8
CH-1214 Vernier, Geneva
Phone: +41 22 749 01 11
Email: copyright@iso.org
Website: www.iso.org

Published in Switzerland

Contents

Page

Foreword	xiii
Introduction	xiv
1 Scope	1
2 Normative references	1
3 Terms, definitions, symbols, abbreviated terms and mnemonics	2
3.1 Terms, definitions, symbols and abbreviated terms	2
3.2 Mnemonics	2
4 Technical overview	2
4.1 Decoder block diagram	2
4.2 Overview over the codec building blocks	3
4.3 Efficient combination of decoder processing blocks in the time domain and QMF domain.....	6
4.4 Rule set for determining processing domains	9
4.4.1 Audio core codec processing domain	9
4.4.2 Mixing	10
4.4.3 DRC-1 Operation domains (DRC in rendering context)	10
4.4.4 Audio core codec interface domain to rendering	10
4.4.5 Rendering context	10
4.4.6 Post-processing context	11
4.4.7 End-of-chain context.....	11
4.5 Sample rate converter	11
4.6 Decoder delay	11
4.7 Contribution mode of MPEG-H 3D audio.....	12
4.8 MPEG-H 3D audio profiles and levels	12
4.8.1 General	12
4.8.2 Profiles	13
5 MPEG-H 3D audio core decoder	27
5.1 Definitions.....	27
5.1.1 Joint stereo.....	27
5.1.2 MPEG surround based stereo (MPS 212)	28
5.2 Syntax.....	28
5.2.1 General	28
5.2.2 Decoder configuration	28
5.2.3 MPEG-H 3D audio core bitstream payloads	51
5.3 Data structure	72
5.3.1 General	72
5.3.2 General configuration data elements.....	72
5.3.3 Loudspeaker configuration data elements	75
5.3.4 Core decoder configuration data elements.....	77
5.3.5 Downmix matrix data elements	81
5.3.6 HOA rendering matrix data elements	84
5.3.7 Signal group information elements	87
5.3.8 Low frequency enhancement (LFE) channel element, mpeg3daLfeElement()	87
5.3.9 Compatible profile and levels sets.....	88
5.4 Configuration element descriptions	88
5.4.1 General	88
5.4.2 Downmix configuration.....	88
5.4.3 HOA rendering matrix configuration	94

This is a preview of "ISO/IEC 23008-3:2022". Click [here](#) to purchase the full version from the ANSI store.

5.5	Tool descriptions	98
5.5.1	General.....	98
5.5.2	Quad channel element	98
5.5.3	Transform splitting.....	100
5.5.4	MPEG surround for mono to stereo upmixing.....	107
5.5.5	Enhanced noise filling	110
5.5.6	Audio pre-roll	134
5.5.7	Fullband LPD	137
5.5.8	Time-domain bandwidth extension.....	148
5.5.9	LPD stereo coding	161
5.5.10	Multichannel coding tool.....	169
5.5.11	Filterbank and block switching	179
5.5.12	Frequency domain prediction.....	180
5.5.13	Long-term postfilter.....	183
5.5.14	Tonal component coding.....	188
5.5.15	Internal channel on MPS212 for low complexity format conversion	198
5.5.16	High resolution envelope processing (HREP) tool	210
5.6	Buffer requirements.....	216
5.6.1	Minimum decoder input buffer	216
5.6.2	Bit reservoir.....	216
5.6.3	Maximum bit rate.....	217
5.7	Stream access point requirements and inter-frame dependency.....	217
6	Dynamic range control and loudness processing	218
6.1	General.....	218
6.2	Description.....	218
6.3	Syntax.....	219
6.3.1	Loudness metadata	219
6.3.2	Dynamic range control metadata.....	219
6.3.3	Data elements.....	220
6.4	Decoding process	222
6.4.1	General.....	222
6.4.2	Dynamic range control.....	224
6.4.3	Usage of downmixId in MPEG-H	224
6.4.4	DRC set selection process	225
6.4.5	DRC-1 for SAOC 3D Content.....	227
6.4.6	DRC-1 for HOA content.....	228
6.4.7	Loudness normalization.....	229
6.4.8	Peak limiter.....	230
6.4.9	Time-synchronization of DRC gains	230
6.4.10	Default parameters	230
7	Object metadata decoding	230
7.1	General.....	230
7.2	Description.....	230
7.3	Syntax.....	231
7.3.1	Object metadata configuration	231
7.3.2	Top level object metadata syntax	232
7.3.3	Subsidiary payloads for efficient object metadata decoding	233
7.3.4	Subsidiary payloads for object metadata decoding with low delay	239
7.3.5	Enhanced object metadata configuration	244
7.4	Data structure	247
7.4.1	Definition of ObjectMetadataConfig() payloads	247
7.4.2	Efficient object metadata decoding	248
7.4.3	Object metadata decoding with low delay	256

This is a preview of "ISO/IEC 23008-3:2022". Click [here](#) to purchase the full version from the ANSI store.

7.4.4	Enhanced object metadata	261
8	Object rendering.....	264
8.1	Description.....	264
8.2	Terms and definitions.....	264
8.3	Input data	264
8.4	Processing	265
8.4.1	General remark	265
8.4.2	Imaginary loudspeakers	266
8.4.3	Dividing the loudspeaker setup into a triangle mesh	267
8.4.4	Rendering algorithm	269
9	SAOC 3D.....	273
9.1	Description.....	273
9.2	Definitions.....	273
9.3	Delay and synchronization.....	275
9.4	Syntax.....	275
9.4.1	Payloads for SAOC 3D.....	275
9.4.2	Definition of SAOC 3D payloads.....	279
9.5	SAOC 3D processing	281
9.5.1	Compressed data stream decoding and dequantization of SAOC 3D data	281
9.5.2	Time/frequency transforms.....	281
9.5.3	Signals and parameters	282
9.5.4	SAOC 3D decoding.....	284
9.5.5	Dual mode	289
10	Generic loudspeaker rendering/format conversion	289
10.1	Description.....	289
10.2	Definitions.....	291
10.2.1	General remarks	291
10.2.2	Variable definitions	291
10.3	Processing	291
10.3.1	Application of transmitted downmix matrices.....	291
10.3.2	Application of transmitted equalizer settings	296
10.3.3	Downmix processing involving multiple channel groups.....	296
10.3.4	Initialization of the format converter	297
10.3.5	Audio signal processing.....	313
11	Immersive loudspeaker rendering/format conversion	319
11.1	Description.....	319
11.2	Syntax.....	321
11.3	Definitions.....	321
11.3.1	General remarks	321
11.3.2	Variable definitions	322
11.4	Processing	323
11.4.1	Initialization of the format converter	323
11.4.2	Audio signal processing.....	366
12	Higher order ambisonics (HOA).....	373
12.1	Technical overview	373
12.1.1	Block diagram	373
12.1.2	Overview of the decoder tools	374
12.2	Syntax.....	376
12.2.1	Configuration of HOA elements	376
12.2.2	Payloads of HOA elements.....	380
12.3	Data structure	393
12.3.1	Definitions of HOA Config	393

This is a preview of "ISO/IEC 23008-3:2022". Click [here](#) to purchase the full version from the ANSI store.

12.3.2 Syntax of getSubbandBandwidths()	397
12.3.3 Definitions of HOA payload	398
12.4 HOA tool description	406
12.4.1 HOA frame converter	406
12.4.2 Spatial HOA decoding	423
12.4.3 HOA renderer	452
12.4.4 Layered coding for HOA	461
13 Binaural renderer	464
13.1 General	464
13.2 Frequency-domain binaural renderer	464
13.2.1 General	464
13.2.2 Definitions	466
13.2.3 Parameterization of binaural room impulse responses	470
13.2.4 Frequency-domain binaural processing	482
13.3 Time-domain binaural renderer	489
13.3.1 General	489
13.3.2 Definitions	490
13.3.3 Parameterization of binaural room impulse responses	492
13.3.4 Time-domain binaural processing	497
14 MPEG-H 3D audio stream (MHAS)	498
14.1 Overview	498
14.2 Syntax	498
14.2.1 Main MHAS syntax elements	498
14.2.2 Subsidiary MHAS syntax elements	500
14.3 Semantics	501
14.3.1 mpegAudioStreamPacket()	501
14.3.2 MHASPacketPayload()	502
14.3.3 Subsidiary MHAS packets	503
14.4 Description of MHASPacketTypes	503
14.4.1 PACTYP_FILLDATA	503
14.4.2 PACTYP_MPEGH3DACFG	504
14.4.3 PACTYP_MPEGH3DAFRAME	504
14.4.4 PACTYP_SYNC	504
14.4.5 PACTYP_SYNCGAP	504
14.4.6 PACTYP_MARKER	504
14.4.7 PACTYP_CRC16 and PACTYP_CRC32	505
14.4.8 PACTYP_DESCRIPTOR	505
14.4.9 PACTYP_USERINTERACTION	506
14.4.10 PACTYP_LOUDNESS_DRC	506
14.4.11 PACTYP_BUFFERINFO	506
14.4.12 PACTYP_GLOBAL_CRC16 and PACTYP_GLOBAL_CRC32	506
14.4.13 PACTYP_AUDIOTRUNCATION	507
14.4.14 PACTYP_AUDIOSCENEINFO	508
14.4.15 PACTYP_EARCON	508
14.4.16 PACTYP_PCMCONFIG	508
14.4.17 PACTYP_PCMDATA	508
14.4.18 PACTYP_LOUDNESS	509
14.4.19 MHASPacketType specific requirements for MHASPacketLabel	509
14.5 Application examples	510
14.5.1 Light-weighted broadcast	510
14.5.2 MPEG-2 transport stream	510
14.5.3 CRC error detection	511
14.5.4 Audio sample truncation	511

This is a preview of "ISO/IEC 23008-3:2022". Click [here](#) to purchase the full version from the ANSI store.

14.6	Multi-stream delivery and interface	512
14.7	Carriage of generic data	514
14.7.1	Syntax.....	515
14.7.2	Semantics.....	515
14.7.3	Processing at the MPEG-H 3D audio decoder	516
15	Metadata audio elements (MAE)	516
15.1	General	516
15.2	Syntax.....	518
15.3	Semantics.....	527
15.4	Definition of mae_metaDataElementIDs.....	540
15.5	Loudness compensation after gain interactivity.....	541
16	Loudspeaker distance compensation.....	543
17	Interfaces to the MPEG-H 3D audio decoder.....	544
17.1	General	544
17.2	Interface for local setup information	544
17.2.1	General	544
17.2.2	WIRE output.....	544
17.2.3	Syntax for local setup information	545
17.2.4	Semantics for local setup information	545
17.3	Interface for local loudspeaker setup and rendering.....	546
17.3.1	General	546
17.3.2	Syntax for local loudspeaker signalling.....	546
17.3.3	Semantics for local loudspeaker signalling.....	547
17.4	Interface for binaural room impulse responses (BRIRs)	548
17.4.1	General	548
17.4.2	Syntax of binaural renderer interface	548
17.4.3	Semantics.....	553
17.5	Interface for local screen size information	557
17.5.1	General	557
17.5.2	Syntax.....	557
17.5.3	Semantics.....	557
17.6	Interface for signaling of local zoom area.....	558
17.6.1	General	558
17.6.2	Syntax.....	558
17.6.3	Semantics.....	559
17.7	Interface for user interaction.....	559
17.7.1	General	559
17.7.2	Definition of user interaction categories.....	559
17.7.3	Definition of an interface for user interaction.....	560
17.7.4	Syntax of interaction interface	561
17.7.5	Semantics of interaction interface	562
17.8	Interface for loudness normalization and dynamic range control (DRC).....	564
17.9	Interface for scene displacement data.....	565
17.9.1	General	565
17.9.2	Definition of an interface for scene-displacement data	565
17.9.3	Syntax of the scene displacement interface	566
17.9.4	Semantics of the scene displacement interface	566
17.10	Interfaces for channel-based, object-based, and HOA metadata and audio data	567
17.10.1	General.....	567
17.10.2	Expectations on external renderers.....	567
17.10.3	Object-based metadata and audio data (object output interface)	567
17.10.4	Channel-based metadata and audio data	575
17.10.5	HOA metadata and audio data.....	580

This is a preview of "ISO/IEC 23008-3:2022". [Click here to purchase the full version from the ANSI store.](#)

17.10.6	Audio PCM data	584
17.11	Interface for positional scene displacement data	585
17.11.1	General	585
17.11.2	Syntax of the positional scene displacement interface	585
17.11.3	Semantics of the positional scene displacement interface	585
17.11.4	Processing	586
18	Application and processing of local setup information and interaction data and scene displacement data	586
18.1	Element metadata preprocessing	586
18.2	Interactivity limitations and restrictions	592
18.2.1	General information	592
18.2.2	WIRE interactivity	592
18.2.3	Position interactivity	593
18.2.4	Screen-related element remapping and object remapping for zooming	593
18.2.5	Closest loudspeaker layout	594
18.3	Screen-related element remapping	594
18.4	Screen-related adaptation and zooming for higher order ambisonics (HOA)	597
18.5	Object remapping for zooming	599
18.6	Determination of the closest loudspeaker	601
18.7	Determination of a list of loudspeakers for conditioned closest loudspeaker playback	601
18.8	Processing of scene displacement angles for channels and objects (CO)	603
18.9	Processing of scene displacement angles for scene-based content (HOA)	605
18.10	Determination of a reduced reproduction layout based on excluded sectors	607
18.11	Diffuseness rendering	608
19	MPEG-H 3D audio profile definition	610
20	Carriage of MPEG-H 3D audio in ISO base media file format	610
20.1	General	610
20.2	Random access and stream access	611
20.3	Overview of new box structures	611
20.4	MHA decoder configuration record	611
20.4.1	Definition	611
20.4.2	Syntax	612
20.4.3	Semantics	612
20.5	MPEG-H audio sample entry	612
20.5.1	Definition	612
20.5.2	Syntax	613
20.5.3	Semantics	613
20.6	MPEG-H audio MHAS sample entry	614
20.6.1	Definition	614
20.6.2	Syntax	615
20.7	MHA dynamic range control and loudness	615
20.7.1	Definition	615
20.7.2	Syntax	615
20.7.3	Semantics	617
20.8	MHA multi-stream signalling	617
20.8.1	Definition	617
20.8.2	Syntax	618
20.8.3	Semantics	618
20.9	Audio scene information	619
20.9.1	MHA group definition	619
20.9.2	MHA switch group definition	620
20.9.3	MHA group preset definition	621
20.9.4	MHA group description text label	623

This is a preview of "ISO/IEC 23008-3:2022". Click [here](#) to purchase the full version from the ANSI store.

20.9.5	MHA scene information	625
20.10	Track references	626
20.11	MPEG-H Audio profile and level compatibility sets	626
20.11.1	Definition	626
20.11.2	Syntax	626
20.11.3	Semantics	626
21	Sub-parameters for the MIME type 'Codecs' parameter	627
21.1	General	627
21.2	'Codecs' parameter for MPEG-H 3D audio	627
22	Timing considerations and decoder behaviour	627
23	Multi-stream handling	627
23.1	Restrictions on extension payloads	627
24	Low complexity generic loudspeaker rendering/format conversion	629
24.1	Description	629
24.2	Definitions	630
24.2.1	General remarks	630
24.2.2	Variable definitions	621
24.3	Processing	631
24.3.1	Application of transmitted downmix matrices	631
24.3.2	Application of transmitted equalizer settings	636
24.3.3	Downmix processing involving multiple channel groups	636
24.3.4	Initialization of the format converter	637
24.3.5	Audio signal processing	654
25	Low complexity immersive loudspeaker rendering/format conversion	656
25.1	Description	656
25.2	Syntax	657
25.3	Definitions	658
25.3.1	General remarks	658
25.3.2	Variable definitions	658
25.4	Processing	659
25.4.1	Initialization of the format converter	659
25.4.2	Audio signal processing	703
26	MPEG surround	706
26.1	Technical overview	706
26.2	Syntax and data structure	707
26.3	Tool description	707
27	Production metadata decoding	707
27.1	General	707
27.1.1	Object distance coding	707
27.1.2	Direct headphone signalling	707
27.1.3	Reference distance coding	708
27.2	Syntax	709
27.2.1	Production metadata configuration	709
27.2.2	Production metadata frame	710
27.3	Semantics	711
27.3.1	Production metadata configuration	711
27.4	Decoding process	713
28	Earcon metadata	713
28.1	General	713
28.2	Syntax	713

This is a preview of "ISO/IEC 23008-3:2022". [Click here to purchase the full version from the ANSI store.](#)

28.3	Semantics.....	715
28.4	Decoding process	719
Annex A (normative) Tables for arithmetic decoding of IGF information		721
A.1	cf_se01[27]	721
A.2	cf_se10[27]	721
A.3	cf_se02[7][27].....	721
A.4	short cf_se20[7][27].....	721
A.5	short cf_se11[7][7][27]	722
A.6	cf_off_se01.....	723
A.7	cf_off_se10.....	723
A.8	cf_off_se02[7]	723
A.9	short cf_off_se20[7]	724
A.10	cf_off_se11[7][7]	724
Annex B (normative) SAOC 3D Decorrelator pre-mixing matrices.....		725
B.1	Premixing matrix for 22.2 output configuration.....	725
B.2	Algorithm for generating pre-mixing matrices.....	727
B.2.1	Input to the algorithm and representations	727
B.2.2	Algorithm steps	727
Annex C (informative) Encoder tools.....		732
C.1	General overview	732
C.1.1	Encoder block diagram	732
C.1.2	Overview of the encoder and decoder building blocks.....	732
C.2	Core encoder tools.....	733
C.2.1	Quad channel element	733
C.2.2	Transform splitting.....	734
C.2.3	Calculation of residual signal for MPEG surround with hybrid residual coding	735
C.2.4	Enhanced noise filling	735
C.3	Object metadata encoding.....	736
C.3.1	Pre-processing of the object metadata	736
C.3.2	Efficient object metadata encoding.....	737
C.3.3	Object metadata encoding with low delay	737
C.3.4	Spatially skipping objects	738
C.4	SAOC 3D encoder.....	738
C.4.1	Overview	738
C.4.2	Calculation of the SAOC 3D parameters.....	738
C.4.3	Time/frequency transform	739
C.4.4	Framing	739
C.4.5	Parameter quantization and coding	739
C.5	HOA encoder.....	739
C.5.1	Specification of value ranges for HOA coefficients	739
C.5.2	Encoder block diagram	742
C.5.3	Spatial HOA encoding	743
C.6	MPEG surround encoder tool	779
Annex D (normative) Peak limiter for unguided clipping prevention		781
Annex E (normative) Compact template downmix matrices		782
Annex F (normative) HOA tables.....		783
F.1	HOA format description.....	783
F.1.1	Spherical coordinate system for HOA.....	783
F.1.2	General.....	783
F.1.3	Definition of real valued spherical harmonics.....	785
F.1.4	Definition of the HOA signal matrix	785

This is a preview of "ISO/IEC 23008-3:2022". Click [here](#) to purchase the full version from the ANSI store.

F.1.5	Definition of the mode matrix.....	785
F.2	Uniformly distributed positions $\Omega q(N = 1) = \theta qN = 1, \phi qN = 1, 1 \leq j \leq O = 4$	786
F.3	Uniformly distributed positions $\Omega q(N = 2) = \theta qN = 2, \phi qN = 2, 1 \leq j \leq O = 9$	786
F.4	Uniformly distributed positions $\Omega q(N = 3) = \theta qN = 3, \phi qN = 3, 1 \leq j \leq O = 16$	786
F.5	Uniformly distributed positions $\Omega q(N = 4) = \theta qN = 4, \phi qN = 4, 1 \leq j \leq O = 25$	787
F.6	32 Uniformly distributed positions in spherical coordinates.....	787
F.7	Uniformly distributed positions $\Omega q(N = 5) = \theta qN = 5, \phi qN = 5, 1 \leq j \leq O = 36$	788
F.8	Uniformly distributed positions $\Omega q(N = 6) = \theta qN = 6, \phi qN = 6, 1 \leq j \leq O = 49$	789
F.9	Uniformly distributed positions $\Omega q(N = 29) = \theta qN = 29, \phi qN = 29, 1 \leq j \leq O = 900$	790
F.10	Uniformly distributed positions $\Omega q(N = 17) = \theta qN = 17, \phi qN = 17, 1 \leq j \leq O = 324$	797
F.11	Uniformly distributed positions $\Omega q(N = 15) = \theta qN = 15, \phi qN = 15, 1 \leq j \leq O = 256$	800
F.12	Table of loudspeaker directions	802
F.13	Table of 64 horizontal-only directions	803
F.14	Table of 256x8 weighting values, WeightValCdbk	804
F.15	Huffman tables for HuffTabIndex=6.....	810
F.16	Huffman tables for HuffTabIndex=7	810
F.17	Huffman tables for HuffTabIndex=8.....	810
F.18	Huffman tables for HuffTabIndex=9.....	810
F.19	Huffman tables for HuffTabIndex=10	811
F.20	Huffman tables for HuffTabIndex=11	811
F.21	Huffman tables for HuffTabIndex=12	811
F.22	Huffman tables for HuffTabIndex=13	812
F.23	Huffman tables for HuffTabIndex=14	812
F.24	Huffman tables for HuffTabIndex=15	812
F.25	HOA Spherical grid for DRC DSHT for order N=1	813
F.26	Spherical grid for DRC DSHT for order N=2	813
F.27	Spherical grid for DRC DSHT for order N=3	813
F.28	Spherical grid for DRC DSHT for order N=4	814
F.29	Spherical grid for DRC DSHT for order N=5	814
F.30	Spherical grid for DRC DSHT for order N=6	815
F.31	Spherical grid for DRC DSHT for order N=7	816
F.32	Spherical grid for DRC DSHT for order N=8	817
F.33	Spherical grid for DRC DSHT for order N=9	818
F.34	Huffman table for decoding HuffmanMagDiffNoSbr	820
F.35	Huffman table for decoding HuffmanMagDiffSbr	820
F.36	Huffman table for decoding DecTableAngleDiff.....	821
F.37	Huffman table for decoding ParHuffmanMagDiffNoSbr	821
F.38	Huffman table for decoding ParHuffmanMagDiffSbr	822
F.39	Huffman table for decoding ParDecTableAngleDiff.....	822
F.40	Table for ParDecorrSigsSelectionTableIdx referring to NumOfDecorrSigsPerParSubbandTable and ParSelectedDecorrSigsIdxMatrixTable	823
F.41	Table for ParDecorrSigsSelectionTableIdx referring to ParPermlIdxVectorTable.....	824
F.42	Table for ParDecorrSigsSelectionTableIdx referring to NumOfDecorrSigsPerFirstOrderParSubbandTable and ParFirstOrderSelectedDecorrSigsIdxMatrixTable	824
F.43	Table for ParDecorrSigsSelectionTableIdx referring to ParFirstOrderPermlIdxVectorTable.....	824
Annex G (informative)	Low complexity HOA rendering.....	825
G.1	General	825
G.2	Detailed description	827
G.2.1	Realization of combined HOA synthesis and rendering.....	828
G.2.2	Further remarks	837

This is a preview of "ISO/IEC 23008-3:2022". Click here to purchase the full version from the ANSI store.

Annex H (informative) Information on delay and complexity of time-domain	
binauralization	839
H.1 Complexity and latency.....	839
H.1.1 Algorithm description.....	839
H.1.2 Complexity.....	840
H.1.3 Latency.....	840
H.2 Experimental results	841
H.3 Alternative low-delay implementations.....	842
Annex I (informative) Determination of a rotation matrix for processing of scene	
displacement data	844
Annex J (informative) Decorrelation filtering for ‘diffuseness’ processing.....	846
Annex K (informative) Distance and depth spread rendering	847
Annex L (informative) HREP encoder description.....	849
L.1 Computation of the gains and the beta factor.....	849
L.2 Computation of the LP part and the HP part.....	850
L.3 Computation of the output signal	850
L.4 Encoding of gains using arithmetic coding.....	851
Annex M (informative) Screen-related adaptation of HOA content in complexity	
constrained implementations.....	853
Annex N (normative) Retaining original file length with MPEG-H 3D audio.....	854
N.1 General.....	854
N.2 Avoiding leading zero samples	854
N.3 Avoiding trailing zero samples	855
Annex O (normative) Codebook tables used to de-quantize high band time domain	
bandwidth extension parameters.....	856
0.1 Codebook table for theLSFCB1_7b	856
0.2 Codebook table for theLSFCB2_7b	858
0.3 Codebook table for theExcFilterCB1_7b	860
0.4 Codebook table for theExcFilterCB2_4b	863
0.5 Codebook table for SHBCB_SubGain5bit	863
0.6 Codebook table for SHBCB_GainFrame5bit.....	863
0.7 Codebook table for win_flatten.....	863
0.8 Codebook table for RVEC.....	864
Annex P (informative) Implementation and usage guidelines for signalling of profile and	
level compatibility sets	865
P.1 Baseline profile signalling.....	865
Bibliography	867

This is a preview of "ISO/IEC 23008-3:2022". Click here to purchase the full version from the ANSI store.

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.

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 document 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 or www.iec.ch/members_experts/refdocs).

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. 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) or the IEC list of patent declarations received (see <https://patents.iec.ch>).

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

For an explanation of the voluntary nature of standards, 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 www.iso.org/iso/foreword.html. In the IEC, see www.iec.ch/understanding-standards.

This document was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

This third edition cancels and replaces the second edition (ISO/IEC 23008-3:2019), which has been technically revised. It also incorporates Amendments ISO/IEC 23008-3:2019/Amd 1:2019 and ISO/IEC 23008-3:2019/Amd 2:2020.

The main changes are as follows:

- specification of audio metadata enhancements;
- carriage of Earcon metadata and PCM data in MHAS packets;
- specification of baseline profile;
- signaling of compatibility to multiple profiles and levels.

A list of all parts in the ISO/IEC 23008 series can be found on the ISO and IEC websites.

Any feedback or questions on this document should be directed to the user's national standards body. A complete listing of these bodies can be found at www.iso.org/members.html and www.iec.ch/national-committees.

Introduction

3D sound systems are able to realize a significantly enhanced sound experience relative to current widespread 5.1 channel audio programs and playback systems. These systems demand high quality audio coding and error-free transmission in order to keep the timbre, sound localization and sound envelopment of the original audio program. Presentation over headphones with suitable spatialization are also considered.

This document provides means for all scenarios where there is a need to compress a multi-channel audio program (e.g. 22.2 channel program) and to render it to the native target number of loudspeakers. In order to reach a wide market, a 3D audio program is able to be downmixed to a lower hierarchy of loudspeakers, for example 10.1 or 8.1 channels. In addition, all scenarios support a level of random access to facilitate broadcast break-in, and “trick modes” such as fast forward when playing from stored media.

This document focuses on applications such as audio for home theatres where the audio presentation is immersive, involving many loudspeakers (e.g. from 10 to more than 20) surrounding the listener and placed below, at and above ear-level. Moreover, applications as personal TV, TV for smartphones and multi-channel audio-only programs are envisioned. These require that 3D audio encoding/decoding systems are able to operate at low bitrates appropriate for efficient transmission over a cellular channel. At the same time, the sense of envelopment and accurate sonic localization even for systems having a tablet-sized visual displays with loudspeakers built into the device and headphone listening are maintained. The definition of the main profile, its associated bitstream syntax, semantics, and decoding process description was provided in the first edition of this document (ISO/IEC 23008-3:2015).

The International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC) draw attention to the fact that it is claimed that compliance with this document may involve the use of a patent.

ISO and IEC take no position concerning the evidence, validity and scope of this patent right.

The holder of this patent right has assured ISO and IEC that he/she is willing to negotiate licences under reasonable and non-discriminatory terms and conditions with applicants throughout the world. In this respect, the statement of the holder of this patent right is registered with ISO and IEC. Information may be obtained from the patent database available at www.iso.org/patents or patents.iec.ch.

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

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

Part 3: 3D audio

1 Scope

This document specifies technology that supports the efficient transmission of immersive audio signals and flexible rendering for the playback of immersive audio in a wide variety of listening scenarios. These include home theatre setups with 3D loudspeaker configurations, 22.2 loudspeaker systems, automotive entertainment systems and playback over headphones connected to a tablet or smartphone.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

ISO/IEC 13818-1, *Information technology — Generic coding of moving pictures and associated audio information — Part 1: Systems*

ISO/IEC 14496-3:2019, *Information technology — Coding of audio-visual objects — Part 3: Audio*

ISO/IEC 14496-11, *Information technology — Coding of audio-visual objects — Part 11: Scene description and application engine*

ISO/IEC 23091-3, *Information technology — Coding-independent code-points — Part 3: Audio*

ISO/IEC 23003-1:2007, *Information technology — MPEG audio technologies — Part 1: MPEG Surround*

ISO/IEC 23003-2, *Information technology — MPEG audio technologies — Part 2: Spatial Audio Object Coding (SAOC)*

ISO/IEC 23003-3:2020, *Information technology — MPEG audio technologies — Part 3: Unified speech and audio coding*

ISO/IEC 23003-4:2020, *Information technology — MPEG audio technologies — Part 4: Dynamic range control*

IETF RFC 4122, *A Universally Unique IDentifier (UUID) URN Namespace*