

ISO/IEC 23008-3**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

**Fourth edition
2026-02**

This is a preview of ISO/IEC 23008-3:2026. [Click here to purchase the full version from the ANSI store.](#)



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2026

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.....	xii
Introduction.....	xii
1 Scope	1
2 Normative references	1
3 Terms, definitions, symbols, abbreviated terms and conventions.....	1
3.1 Terms and definitions.....	1
3.2 Symbols, abbreviated terms and conventions	2
3.2.1 Symbols and abbreviated terms.....	2
3.2.2 Conventions.....	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, <code>mpegh3daLfeElement()</code>	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

This is a preview of ISO/IEC 23008-3:2026. [Click here to purchase the full version from the ANSI store.](#)

5.4.3	HOA rendering matrix configuration	94
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.....	229
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	238
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.....	247
7.4.3	Object metadata decoding with low delay	255

This is a preview of ISO/IEC 23008-3:2026. [Click here to purchase the full version from the ANSI store.](#)

7.4.4	Enhanced object metadata	260
8	Object rendering	263
8.1	Description	263
8.2	Terms and definitions	263
8.3	Input data	264
8.4	Processing	265
8.4.1	General remark	265
8.4.2	Imaginary loudspeakers	265
8.4.3	Dividing the loudspeaker setup into a triangle mesh	266
8.4.4	Rendering algorithm	268
9	SAOC 3D	272
9.1	Description	272
9.2	Definitions	272
9.3	Delay and synchronization	274
9.4	Syntax	274
9.4.1	Payloads for SAOC 3D	274
9.4.2	Definition of SAOC 3D payloads	278
9.5	SAOC 3D processing	280
9.5.1	Compressed data stream decoding and dequantization of SAOC 3D data	280
9.5.2	Time/frequency transforms	280
9.5.3	Signals and parameters	281
9.5.4	SAOC 3D decoding	283
9.5.5	Dual mode	288
10	Generic loudspeaker rendering/format conversion	288
10.1	Description	288
10.2	Definitions	290
10.2.1	General remarks	290
10.2.2	Variable definitions	290
10.3	Processing	290
10.3.1	Application of transmitted downmix matrices	290
10.3.2	Application of transmitted equalizer settings	295
10.3.3	Downmix processing involving multiple channel groups	295
10.3.4	Initialization of the format converter	296
10.3.5	Audio signal processing	312
11	Immersive loudspeaker rendering/format conversion	318
11.1	Description	318
11.2	Syntax	320
11.3	Definitions	320
11.3.1	General remarks	320
11.3.2	Variable definitions	321
11.4	Processing	322
11.4.1	Initialization of the format converter	322
11.4.2	Audio signal processing	364
12	Higher order ambisonics (HOA)	372
12.1	Technical overview	372
12.1.1	Block diagram	372
12.1.2	Overview of the decoder tools	373
12.2	Syntax	374
12.2.1	Configuration of HOA elements	374
12.2.2	Payloads of HOA elements	378
12.3	Data structure	391
12.3.1	Definitions of HOA Config	391

This is a preview of ISO/IEC 23008-3:2026. [Click here to purchase the full version from the ANSI store.](#)

12.3.2	Syntax of <code>getSubbandBandwidths()</code>	395
12.3.3	Definitions of HOA payload	396
12.4	HOA tool description	403
12.4.1	HOA frame converter	403
12.4.2	Spatial HOA decoding	419
12.4.3	HOA renderer	448
12.4.4	Layered coding for HOA	457
13	Binaural renderer	460
13.1	General	460
13.2	Frequency-domain binaural renderer	460
13.2.1	General	460
13.2.2	Definitions	462
13.2.3	Parameterization of binaural room impulse responses	466
13.2.4	Frequency-domain binaural processing	478
13.3	Time-domain binaural renderer	485
13.3.1	General	485
13.3.2	Definitions	486
13.3.3	Parameterization of binaural room impulse responses	488
13.3.4	Time-domain binaural processing	492
14	MPEG-H 3D audio stream (MHAS)	493
14.1	Overview	493
14.2	Syntax	494
14.2.1	Main MHAS syntax elements	494
14.2.2	Subsidiary MHAS syntax elements	496
14.3	Semantics	496
14.3.1	<code>mpeghAudioStreamPacket()</code>	496
14.3.2	<code>MHASPacketPayload()</code>	497
14.3.3	Subsidiary MHAS packets	499
14.4	Description of <code>MHASPacketTypes</code>	499
14.4.1	<code>PACTYP_FILLDATA</code>	499
14.4.2	<code>PACTYP_MPEGH3DACFG</code>	499
14.4.3	<code>PACTYP_MPEGH3DAFRAME</code>	499
14.4.4	<code>PACTYP_SYNC</code>	500
14.4.5	<code>PACTYP_SYNCGAP</code>	500
14.4.6	<code>PACTYP_MARKER</code>	500
14.4.7	<code>PACTYP_CRC16</code> and <code>PACTYP_CRC32</code>	501
14.4.8	<code>PACTYP_DESCRIPTOR</code>	501
14.4.9	<code>PACTYP_USERINTERACTION</code>	501
14.4.10	<code>PACTYP_LOUDNESS_DRC</code>	501
14.4.11	<code>PACTYP_BUFFERINFO</code>	502
14.4.12	<code>PACTYP_GLOBAL_CRC16</code> and <code>PACTYP_GLOBAL_CRC32</code>	502
14.4.13	<code>PACTYP_AUDIOTRUNCATION</code>	502
14.4.14	<code>PACTYP_AUDIOSCENEINFO</code>	503
14.4.15	<code>PACTYP_EARCON</code>	503
14.4.16	<code>PACTYP_PCMCONFIG</code>	504
14.4.17	<code>PACTYP_PCMDATA</code>	504
14.4.18	<code>PACTYP_LOUDNESS</code>	504
14.4.19	MHASPacketType specific requirements for <code>MHASPacketLabel</code>	504
14.5	Application examples	505
14.5.1	Light-weighted broadcast	505
14.5.2	MPEG-2 transport stream	506
14.5.3	CRC error detection	506
14.5.4	Audio sample truncation	507
14.6	Multi-stream delivery and interface	507

This is a preview of ISO/IEC 23008-3:2026. [Click here to purchase the full version from the ANSI store.](#)

14.7	Carriage of generic data	510
14.7.1	Syntax.....	510
14.7.2	Semantics.....	511
14.7.3	Processing at the MPEG-H 3D audio decoder	512
15	Metadata audio elements (MAE)	512
15.1	General	512
15.2	Syntax.....	513
15.3	Semantics.....	522
15.4	Definition of mae_metaDataElementIDs	534
15.5	Loudness compensation after gain interactivity.....	535
16	Loudspeaker distance compensation.....	537
17	Interfaces to the MPEG-H 3D audio decoder.....	538
17.1	General	538
17.2	Interface for local setup information	538
17.2.1	General	538
17.2.2	WIRE output	538
17.2.3	Syntax for local setup information	539
17.2.4	Semantics for local setup information	539
17.3	Interface for local loudspeaker setup and rendering.....	540
17.3.1	General	540
17.3.2	Syntax for local loudspeaker signalling.....	540
17.3.3	Semantics for local loudspeaker signalling.....	541
17.4	Interface for binaural room impulse responses (BRIRs).....	542
17.4.1	General	542
17.4.2	Syntax of binaural renderer interface	542
17.4.3	Semantics.....	547
17.5	Interface for local screen size information	551
17.5.1	General	551
17.5.2	Syntax.....	551
17.5.3	Semantics.....	551
17.6	Interface for signaling of local zoom area.....	552
17.6.1	General	552
17.6.2	Syntax.....	552
17.6.3	Semantics.....	552
17.7	Interface for user interaction.....	553
17.7.1	General	553
17.7.2	Definition of user interaction categories.....	553
17.7.3	Definition of an interface for user interaction.....	554
17.7.4	Syntax of interaction interface	555
17.7.5	Semantics of interaction interface	556
17.8	Interface for loudness normalization and dynamic range control (DRC).....	558
17.9	Interface for scene displacement data.....	558
17.9.1	General	558
17.9.2	Definition of an interface for scene-displacement data	559
17.9.3	Syntax of the scene displacement interface	560
17.9.4	Semantics of the scene displacement interface	560
17.10	Interfaces for channel-based, object-based, and HOA metadata and audio data.....	561
17.10.1	General	561
17.10.2	Expectations on external renderers	561
17.10.3	Object-based metadata and audio data (object output interface)	561
17.10.4	Channel-based metadata and audio data	569
17.10.5	HOA metadata and audio data	573
17.10.6	Audio PCM data	577

This is a preview of ISO/IEC 23008-3:2026. [Click here to purchase the full version from the ANSI store.](#)

17.11	Interface for positional scene displacement data	577
17.11.1	General	577
17.11.2	Syntax of the positional scene displacement interface.....	578
17.11.3	Semantics of the positional scene displacement interface	578
17.11.4	Processing.....	578
18	Application and processing of local setup information and interaction data and scene displacement data	579
18.1	Element metadata preprocessing	579
18.1.1	General information	579
18.1.2	Initialization	580
18.1.3	Processing loop.....	581
18.1.4	Element routing.....	585
18.2	Interactivity limitations and restrictions	585
18.2.1	General information	585
18.2.2	WIRE interactivity	585
18.2.3	Position interactivity.....	586
18.2.4	Screen-related element remapping and object remapping for zooming	586
18.2.5	Closest loudspeaker layout.....	587
18.3	Screen-related element remapping	587
18.4	Screen-related adaptation and zooming for higher order ambisonics (HOA).....	590
18.5	Object remapping for zooming	592
18.6	Determination of the closest loudspeaker	594
18.7	Determination of a list of loudspeakers for conditioned closest loudspeaker playback	594
18.8	Processing of scene displacement angles for channels and objects (CO).....	596
18.9	Processing of scene displacement angles for scene-based content (HOA)	598
18.10	Determination of a reduced reproduction layout based on excluded sectors	600
18.11	Diffuseness rendering.....	601
19	MPEG-H 3D audio profile definition.....	603
20	Carriage of MPEG-H 3D audio in ISO base media file format	603
20.1	General.....	603
20.2	Random access and stream access	603
20.3	Overview of new box structures.....	603
20.4	MHA decoder configuration record.....	604
20.4.1	Definition	604
20.4.2	Syntax.....	604
20.4.3	Semantics.....	604
20.5	MPEG-H audio sample entry	604
20.5.1	Definition	604
20.5.2	Syntax.....	605
20.5.3	Semantics.....	605
20.6	MPEG-H audio MHAS sample entry	606
20.6.1	Definition	606
20.6.2	Syntax.....	607
20.7	MHA dynamic range control and loudness.....	607
20.7.1	Definition	607
20.7.2	Syntax.....	607
20.7.3	Semantics.....	609
20.8	MHA multi-stream signalling.....	610
20.8.1	Definition	610
20.8.2	Syntax.....	610
20.8.3	Semantics.....	610
20.9	Audio scene information.....	611

This is a preview of ISO/IEC 23008-3:2026. [Click here to purchase the full version from the ANSI store.](#)

20.9.1	MHA group definition.....	611
20.9.2	MHA switch group definition.....	613
20.9.3	MHA group preset definition.....	614
20.9.4	MHA group description text label.....	615
20.9.5	MHA scene information.....	617
20.10	Track references.....	618
20.11	MPEG-H Audio profile and level compatibility sets	618
20.11.1	Definition	618
20.11.2	Syntax	618
20.11.3	Semantics	618
21	Sub-parameters for the MIME type ‘Codecs’ parameter.....	619
21.1	General.....	619
21.2	‘Codecs’ parameter for MPEG-H 3D audio.....	619
22	Timing considerations and decoder behaviour	619
23	Multi-stream handling.....	619
23.1	Restrictions on extension payloads.....	619
24	Low complexity generic loudspeaker rendering/format conversion.....	621
24.1	Description.....	621
24.2	Definitions.....	622
24.2.1	General remarks	622
24.2.2	Variable definitions	623
24.3	Processing	623
24.3.1	Application of transmitted downmix matrices.....	623
24.3.2	Application of transmitted equalizer settings	628
24.3.3	Downmix processing involving multiple channel groups.....	628
24.3.4	Initialization of the format converter	629
24.3.5	Audio signal processing.....	646
25	Low complexity immersive loudspeaker rendering/format conversion	648
25.1	Description.....	648
25.2	Syntax.....	649
25.3	Definitions.....	650
25.3.1	General remarks	650
25.3.2	Variable definitions	650
25.4	Processing	651
25.4.1	Initialization of the format converter	651
25.4.2	Audio signal processing.....	695
26	MPEG surround	698
26.1	Technical overview.....	698
26.2	Syntax and data structure.....	699
26.3	Tool description	699
27	Production metadata decoding	699
27.1	General	699
27.1.1	Object distance coding.....	699
27.1.2	Direct headphone signalling.....	699
27.1.3	Reference distance coding.....	700
27.2	Syntax.....	701
27.2.1	Production metadata configuration	701
27.2.2	Production metadata frame.....	702
27.3	Semantics.....	703
27.3.1	Production metadata configuration	703
27.4	Decoding process.....	705

This is a preview of ISO/IEC 23008-3:2026. [Click here to purchase the full version from the ANSI store.](#)

28	Earcon metadata	705
28.1	General.....	705
28.2	Syntax.....	705
28.3	Semantics.....	707
28.4	Decoding process	711
Annex A (normative)	Tables for arithmetic decoding of IGF information	713
Annex B (normative)	SAOC 3D Decorrelator pre-mixing matrices.....	717
Annex C (informative)	Encoder tools.....	724
Annex D (normative)	Peak limiter for unguided clipping prevention	772
Annex E (normative)	Compact template downmix matrices	773
Annex F (normative)	HOA tables.....	774
Annex G (informative)	Low complexity HOA rendering	816
Annex H (informative)	Information on delay and complexity of time-domain binauralization	830
Annex I (informative)	Determination of a rotation matrix for processing of scene displacement data	835
Annex J (informative)	Decorrelation filtering for ‘diffuseness’ processing.....	837
Annex K (informative)	Distance and depth spread rendering	838
Annex L (informative)	HREP encoder description.....	840
Annex M (informative)	Screen-related adaptation of HOA content in complexity constrained implementations	844
Annex N (normative)	Retaining original file length with MPEG-H 3D audio.....	845
Annex O (normative)	Codebook tables used to de-quantize high band time domain bandwidth extension parameters.....	847
Annex P (informative)	Implementation and usage guidelines for signalling of profile and level compatibility sets	856
Bibliography	856

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).

ISO and IEC draw attention to the possibility that the implementation of this document may involve the use of (a) patent(s). ISO and IEC take no position concerning the evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, ISO and IEC had received notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the patent database available at www.iso.org/patents and <https://patents.iec.ch>. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

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 fourth edition cancels and replaces the third edition (ISO/IEC 23008-3:2022), which has been technically revised.

The main changes are as follows:

- Fixes and updates normative references;
- Fixes literal and formatting errors;
- Increases comprehensibility regarding random access points, configuration changes, bitrate adaptation, hybrid delivery, IGF, DRC and loudness handling.

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).

This is a preview of ISO/IEC 23008-3:2026. [Click here to purchase the full version from the ANSI store.](#)

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 14496-12, *Information technology — Coding of audio-visual objects — Part 12: ISO base media file format*

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 23003-4:2025, *Information technology — MPEG audio technologies — Part 4: Dynamic range control*

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

3 Terms, definitions, symbols, abbreviated terms and conventions

3.1 Terms and definitions

For the purposes of this document, the terms and definitions given in ISO/IEC 14496-3 and ISO/IEC 23003-3 apply.