INTERNATIONAL STANDARD

ISO/IEC 13818-1

Fifth edition
2015-07-01

AMENDMENT 5
2016-10-15

Information technology — Generic coding of moving pictures and associated audio information —

Part 1: Systems

AMENDMENT 5: Carriage of MPEGH 3D audio over MPEG2 systems

Technologies de l'information — Codage générique des images animées et du son associé —
Partie 1: Systèmes

AMENDMENT 5: Transport de l'audio MPEGH 3D à travers les systèmes MPEG2

Reference number
ISO/IEC 13818-1:2015/Amd.5:2016(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) 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.

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).

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).

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 WTO principles in the Technical Barriers to Trade (TBT) see the following URL: Foreword - Supplementary information

Amendment 5 to ISO/IEC 13818-1:2015 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*, in collaboration with ITU-T. The identical text is published as Rec. ITU-T H.222.0 (07/2016)/Amd.5.

CONTENTS

INTERNATIONAL STANDARD
ITU-T RECOMMENDATION

Information technology – Generic coding of moving pictures and associated audio information: Systems

Amendment 5

Carriage of MPEG-H 3D audio over MPEG-2 systems

1) Clause 1.2.3

In clause 1.2.3, add:

- ISO/IEC 23001-8:2016, *Information technology – MPEG systems technologies – Part 8: Coding-independent code-points*.
- ISO/IEC 23003-3:2012, *Information technology – MPEG audio technologies – Part 3: Unified speech and audio coding*.
- ISO/IEC 23003-4:2015, *Information technology – MPEG audio technologies – Part 4: Dynamic Range Control*.
- ISO/IEC 23008-3:2105, *Information technology – High efficiency coding and media delivery in heterogeneous environments – Part 3: 3D audio*.

2) Table 2-22

Replace 2-22 with the following:

Table 2-22 – Stream_id assignments

Stream_id	Note	stream coding
1011 1100	1	program_stream_map
1011 1101	2,9	private_stream_1
1011 1110		padding_stream
1011 1111	3	private_stream_2
110x xxxx		ISO/IEC 13818-3 or ISO/IEC 11172-3 or ISO/IEC 13818-7 or ISO/IEC 14496-3 or ISO/IEC 23008-3 audio stream number x xxxx
1110 xxxx		Rec. ITU-T H.262 | ISO/IEC 13818-2, ISO/IEC 11172-2, ISO/IEC 14496-2, Rec. ITU-T H.264 | ISO/IEC 14496-10 or Rec. ITU-T H.265 | ISO/IEC 23008-2 video stream number xxxx
1111 0000	3	ECM_stream
1111 0001	3	EMM_stream
1111 0010	5	Rec. ITU-T H.222.0 | ISO/IEC 13818-1 Annex A or ISO/IEC 13818-6_DSMCC_stream
1111 0011	2	ISO/IEC_13522_stream
1111 0100	6	Rec. ITU-T H.222.1 type A
1111 0101	6	Rec. ITU-T H.222.1 type B
1111 0110	6	Rec. ITU-T H.222.1 type C
1111 0111	6	Rec. ITU-T H.222.1 type D
1111 1000	6	Rec. ITU-T H.222.1 type E
1111 1001	7	ancillary_stream
1111 1010		ISO/IEC 14496-1_SL-packetized_stream
1111 1011		ISO/IEC 14496-1_FlexMux_stream
1111 1100		metadata stream
1111 1101	8	extended_stream_id

Table 2-22 – Stream_id assignments

Stream_id	Note	stream coding
1111 1110		reserved data stream
1111 1111	4	program_stream_directory
The notation x means that the values '0' or '1' are both permitted and results in the same stream type. The stream number is given by the values taken by the x's. NOTE 1 – PES packets of type program_stream_map have unique syntax specified in 2.5.4.1. NOTE 2 – PES packets of type private_stream_1 and ISO/IEC_13552_stream follow the same PES packet syntax as those for Rec. ITU-T H.262 | ISO/IEC 13818-2 video and ISO/IEC 13818-3 audio streams. NOTE 3 – PES packets of type private_stream_2, ECM_stream and EMM_stream are similar to private_stream_1 except no syntax is specified after PES_packet_length field. NOTE 4 – PES packets of type program_stream_directory have a unique syntax specified in 2.5.5. NOTE 5 – PES packets of type DSM-CC_stream have a unique syntax specified in ISO/IEC 13818-6. NOTE 6 – This stream_id is associated with stream_type 0x09 in Table 2-34. NOTE 7 – This stream_id is only used in PES packets, which carry data from a program stream or an ISO/IEC 11172-1 System Stream, in a transport stream (refer to 2.4.3.8). NOTE 8 – The use of stream_id 0xFD (extended_stream_id) identifies that this PES packet employs an extended syntax to permit additional stream types to be identified. NOTE 9 – JPEG 2000 video streams (stream_type = 0x21) are carried using the same PES packet syntax as private_stream_1.		

3) Table 2-34

In Table 2-34 add the following:

Table 2-34 – Stream type assignments

Value	Description
0x2D	ISO/IEC 23008-3 Audio with MHAS transport syntax – main stream
0x2E	ISO/IEC 23008-3 Audio with MHAS transport syntax – auxiliary stream
0x2F-0x7E	Rec. ITU-T H.222.0 | ISO/IEC 13818-1 Reserved

4) Clause 2.6.90

Replace Table 2-105 with:

Table 2-105 – Extension descriptor

Syntax	No. of bits	Mnemonic
Extension_descriptor () {		
descriptor_tag	**8**	**uimsbf**
descriptor_length	**8**	**uimsbf**
extension_descriptor_tag	**8**	**uimsbf**
if (extension_descriptor_tag == 0x02) {		
ObjectDescriptorUpdate()		
}		
else if (extension_descriptor_tag == 0x03) {		
HEVC_timing_and_HRD_descriptor()		
}		
else if (extension_descriptor_tag == 0x04) {		

Table 2-105 – Extension descriptor

Syntax	No. of bits	Mnemonic
af_extension_descriptor()		
}		
else if (extension_descriptor_tag == 0x05) {		
HEVC_operation_point_descriptor()		
}		
else if (extension_descriptor_tag == 0x06) {		
HEVC_hierachy_extension_descriptor()		
}		
else if (extension_descriptor_tag == 0x07) {		
Green_extension_descriptor ()		
}		
else if (extension_descriptor_tag == 0x08) {		
MPEG-H_3dAudio_descriptor()		
}		
else if (extension_descriptor_tag == 0x09) {		
MPEG-H_3dAudio_config_descriptor()		
}		
else if (extension_descriptor_tag == 0x0A) {		
MPEG-H_3dAudio_scene_descriptor()		
}		
else if (extension_descriptor_tag == 0x0B) {		
MPEG-H_3dAudio_text_label_descriptor()		
}		
else if (extension_descriptor_tag == 0x0C) {		
MPEG-H_3dAudio_multi-stream_descriptor()		
}		
else if (extension_descriptor_tag == 0x0D) {		
MPEG-H_3dAudio_drc_loudness_descriptor()		
}		
else if (extension_descriptor_tag == 0x0E) {		
MPEG-H_3dAudio_command_descriptor()		
}		
else {		
for (i=0; i<N; i++) {		
reserved	**8**	**bslbf**
}		
}		
}		

5) Clause 2.6.91

Add the following immediately before Table 2-106:

MPEG-H_3dAudio_descriptor() – This structure is defined in 2.6.106 and 2.6.107.

MPEG-H_3dAudio_config_descriptor() – This structure is defined in 2.6.108 and 2.6.109.

MPEG-H_3dAudio_scene_descriptor() – This structure is defined in 2.6.110 and 2.6.111.

MPEG-H_3dAudio_text_label_descriptor() – This structure is defined in 2.6.112 and 2.6.113.

MPEG-H_3dAudio_multi-stream_descriptor() – This structure is defined in 2.6.114 and 2.6.115.

MPEG-H_3dAudio_drc_loudness_descriptor() – This structure is defined in 2.6.116 and 2.6.117.

MPEG-H_3dAudio_command_descriptor() – This structure is defined in 2.6.118.

Replace Table 2-106 with the following:

Table 2-106 – Extension descriptor tag values

Extension_descriptor_tag	TS	PS	Identification
0	n/a	n/a	Reserved
1	n/a	X	Forbidden
2	X	X	ODUpdate_descriptor
3	X	n/a	HEVC_timing_and_HRD_descriptor()
4	X	n/a	af_extensions_descriptor()
5	X	n/a	HEVC_operation_point_descriptor()
6	X	n/a	hierarchy_extension_descriptor()
7	X	n/a	Green_extension_descriptor()
8	X	n/a	MPEG-H_3dAudio_descriptor()
9	X	n/a	MPEG-H_3dAudio_config_descriptor()
0x0A	X	n/a	MPEG-H_3dAudio_scene_descriptor()
0x0B	X	n/a	MPEG-H_3dAudio_text_label_descriptor()
0x0C	X	n/a	MPEG-H_3dAudio_multi-stream_descriptor()
0x0D	X	n/a	MPEG-H_3dAudio_drc_loudness_descriptor()
0x0E	X	n/a	MPEG-H_3dAudio_command_descriptor()
0x0F-0xFF	n/a	n/a	Rec. ITU-T H.222.0 | ISO/IEC 13818-1 Reserved

6) Clauses 2.6.106 to 2.6.118

Add the following clauses after 2.6.105:

2.6.106 MPEG-H 3D audio descriptor

The MPEG-H 3D audio descriptor provides information on basic coding information in the associated ISO/IEC 23008-3 stream. This descriptor shall be present in the associated PMT for MPEG-H 3D audio content with stream_type equal to 0x2D.

Table 2-111*sexies*– MPEG-H 3D audio descriptor

Syntax	No of bits	Mnemonic
MPEG-H_3dAudio_descriptor() {		
mpegh3daProfileLevelIndication	**8**	**uimsbf**
interactivityEnabled	**1**	**bslbf**
reserved	**9**	**bslbf**
referenceChannelLayout	**6**	**uimsbf**
for (i=0; i<N; i++) {		
reserved	**8**	**bslbf**
}		
}		

2.6.107 Semantics for MPEG-H 3D audio descriptor

mpegh3daProfileLevelIndication – The audio profile and level of the associated ISO/IEC 23008-3 audio stream, encoded as specified for the mpegh3daProfileLevelIndication field in clause 5.3.2 in ISO/IEC 23008-3.

referenceChannelLayout – Reference channel configuration value as defined as "ChannelConfiguration" in ISO/IEC 23001-8 ("Codec Independent Code Points").

interactivityEnabled – If set to 1, this flag indicates that the 3D audio stream contains elements with associated metadata which enables user interactivity. If this flag is set to 0, no user interactivity of any kind is available. This flag may be used to determine the need for initializing the user interactivity interface in the Systems decoder.

2.6.108 MPEG-H 3D audio config descriptor

The MPEG-H 3D audio config descriptor provides information on the complete configuration data of one ISO/IEC 23008-3 stream.

Table 2-111*septies* – MPEG-H 3D audio config descriptor

Syntax	No of bits	Mnemonic

MPEG-H_3dAudio_config_descriptor() {
mpegh3daConfig()
}		

2.6.109 Semantics for MPEG-H 3D audio config descriptor

mpegh3daConfig() – The mpegh3daConfig() of the associated ISO/IEC 23008-3 audio stream, as specified in clause 5.2.2.1 in ISO/IEC 23008-3.

2.6.110 MPEG-H 3D audio scene descriptor

The MPEG-H 3D audio scene descriptor provides information on user selectable and/or modifiable audio objects in an ISO/IEC 23008-3 stream.

Table 2-111*octies* – MPEG-H 3d audio scene descriptor

Syntax	No of bits	Mnemonic
MPEG-H_3dAudio_scene_descriptor() {		
groupDefinitionPresent	**1**	**bslbf**
switchGroupDefinitionPresent	**1**	**bslbf**
presetGroupDefinitionPresent	**1**	**bslbf**
reserved	**5**	**bslbf**
3dAudioSceneInfoID	**8**	**bslbf**
if (groupDefinitionPresent) {		
reserved	**1**	**bslbf**
numGroups	**7**	**uimsbf**
for (i=0; i < numGroups; i++) {		
reserved	**1**	**bslbf**
mae_groupID	**7**	**uimsbf**
reserved	**3**	**bslbf**
mae_allowOnOff	**1**	**bslbf**
mae_defaultOnOff	**1**	**bslbf**
mae_allowPositionInteractivity	**1**	**bslbf**
mae_allowGainInteractivity	**1**	**bslbf**
mae_hasContentLanguage	**1**	**bslbf**
reserved	**4**	**bslbf**
mae_contentKind	**4**	**uimsbf**
if (mae_allowPositionInteractivity) {		
reserved	**1**	**bslbf**
mae_interactivityMinAzOffset	**7**	**uimsbf**
reserved	**1**	**bslbf**
mae_interactivityMaxAzOffset	**7**	**uimsbf**
reserved	**3**	**bslbf**
mae_interactivityMinElOffset	**5**	**uimsbf**
reserved	**3**	**bslbf**

Table 2-111*octies* – MPEG-H 3d audio scene descriptor

Syntax	No of bits	Mnemonic
mae_interactivityMaxElOffset	5	uimsbf
mae_interactivityMinDistOffset	4	uimsbf
mae_interactivityMaxDistOffset	4	uimsbf
}		
if (mae_allowGainInteractivity) {		
reserved	2	bslbf
mae_interactivityMinGain	6	uimsbf
reserved	3	bslbf
mae_interactivityMaxGain	5	uimsbf
}		
if (mae_hasContentLanguage) {		
mae_contentLanguage	24	uimsbf
}		
}		
}		
if (switchGroupDefinitionPresent) {		
reserved	3	bslbf
numSwitchGroups	5	uimsbf
for (i=0; i < numSwitchGroups; i++) {		
reserved	1	bslbf
mae_switchGroupID	5	uimsbf
mae_switchGroupAllowOnOff	1	bslbf
mae_switchGroupDefaultOnOff	1	bslbf
reserved	3	bslbf
mae_bsSwitchGroupNumMembers	5	uimsbf
for (i = 0; i < mae_bsSwitchGroupNumMembers + 1; i++) {		
reserved	1	bslbf
mae_switchGroupMemberID	7	uimsbf
}		
reserved	1	bslbf
mae_switchGroupDefaultGroupID	7	uimsbf
}		
}		
if (presetGroupDefinitionPresent) {		
reserved	3	bslbf
mae_numGroupPresets	5	uimsbf
for (i = 0; i < mae_numGroupPresets; i++) {		
reserved	3	bslbf
mae_groupPresetID	5	uimsbf
reserved	3	bslbf
mae_groupPresetKind	5	uimsbf
reserved	4	bslbf
mae_numGroupPresetConditions	4	uimsbf
for (j = 0; j < mae_numGroupPresetConditions+1; j++) {		
mae_groupPresetGroupID	7	uimsbf
mae_groupPresetConditionOnOff	1	bslbf
if (mae_groupPresetConditionOnOff) {		

Table 2-111*octies* – MPEG-H 3d audio scene descriptor

Syntax	No of bits	Mnemonic
reserved	**4**	**bslbf**
mae_groupPresetDisableGainInteractivity	**1**	**bslbf**
mae_groupPresetGainFlag	**1**	**bslbf**
mae_groupPresetDisablePositionInteractivity	**1**	**bslbf**
mae_groupPresetPositionFlag	**1**	**bslbf**
if (mae_groupPresetGainFlag) {		
mae_groupPresetGain	**8**	**uimsbf**
}		
if(mae_groupPresetPositionFlag){		
mae_groupPresetAzOffset	**8**	**uimsbf**
reserved	**2**	**bslbf**
mae_groupPresetElOffset	**6**	**uimsbf**
reserved	**4**	**bslbf**
mae_groupPresetDistFactor	**4**	**uimsbf**
}		
}		
}		
}		
for (i=0; i<N; i++) {		
reserved	**8**	**bslbf**
}		
}		

2.6.111 Semantic definition of fields in MPEG-H 3D audio scene descriptor

groupDefinitionPresent – A one-bit flag signalling the presence of interactivity information of one group in this descriptor.

switchGroupDefinitionPresent – A one-bit flag signalling the presence of switch group information in this descriptor.

presetGroupDefinitionPresent – A one-bit flag signalling the presence of preset group information in this descriptor.

3dAudioSceneInfoID – See ISO/IEC 23008-3, clause 15.3.

numGroups – This field signals the number of groups in the audio scene description. This field can take values between 1 and 127, and shall be less or equal to the value of mae_numGroups present in the associated ISO/IEC 23008-3 stream.

mae_groupID – See ISO/IEC 23008-3, clause 15.3.

mae_allowOnOff – See ISO/IEC 23008-3, clause 15.3.

mae_defaultOnOff – See ISO/IEC 23008-3, clause 15.3.

mae_allowPositionInteractivity – See ISO/IEC 23008-3, clause 15.3.

mae_allowGainInteractivity – See ISO/IEC 23008-3, clause 15.3.

mae_hasContentLanguage – See ISO/IEC 23008-3, clause 15.3.

mae_contentKind – See ISO/IEC 23008-3, clause 15.3.

mae_interactivityMinAzOffset – See ISO/IEC 23008-3, clause 15.3.

mae_interactivityMaxAzOffset – See ISO/IEC 23008-3, clause 15.3.

mae_interactivityMinElOffset – See ISO/IEC 23008-3, clause 15.3.

mae_interactivityMaxElOffset – See ISO/IEC 23008-3, clause 15.3.