INTERNATIONAL STANDARD

ISO/IEC 13818-1

Third edition
2007-10-15

AMENDMENT 1
2007-11-01

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

AMENDMENT 1: Transport of MPEG-4 streaming text and MPEG-4 lossless audio over MPEG-2 systems

Technologies de l'information — Codage générique des images animées et des informations sonores associées: Systèmes

AMENDEMENT 1: Transport de texte en flux MPEG-4 et d'audio sans perte MPEG-4 sur des systèmes MPEG-2

Reference number
ISO/IEC 13818-1:2007/Amd.1:2007(E)

© ISO/IEC 2007

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS

Not for Resale

PDF disclaimer

This PDF file may contain embedded typefaces. In accordance with Adobe's licensing policy, this file may be printed or viewed but shall not be edited unless the typefaces which are embedded are licensed to and installed on the computer performing the editing. In downloading this file, parties accept therein the responsibility of not infringing Adobe's licensing policy. The ISO Central Secretariat accepts no liability in this area.

Adobe is a trademark of Adobe Systems Incorporated.

Details of the software products used to create this PDF file can be found in the General Info relative to the file; the PDF-creation parameters were optimized for printing. Every care has been taken to ensure that the file is suitable for use by ISO member bodies. In the unlikely event that a problem relating to it is found, please inform the Central Secretariat at the address given below.



COPYRIGHT PROTECTED DOCUMENT

© ISO/IEC 2007

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing 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

--`,,```,,,,````-`-`,,`,,`,`,,`---

::

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS

© ISO/IEC 2007 – All rights reserved

Not for Resale

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.

Amendement 1 to ISO/IEC 13818-1 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 ITU-T Rec. H.222.0 (2006)/Amd.1 (01/2007).

© ISO/IEC 2007 – All rights reserved

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

--`,,```,,,,````-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS

Not for Resale

INTERNATIONAL STANDARD
ITU-T RECOMMENDATION

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

Amendment 1

Transport of MPEG-4 streaming text and MPEG-4 lossless audio over MPEG-2 systems

1) Subclause 1.2.3

Add the following references to subclause 1.2.3:

- ISO/IEC 14496-3:2005/Amd.1:2007, *Low delay AAC profile*.
- ISO/IEC 14496-17:2006, *Information technology – Coding of audio-visual objects – Part 17: Streaming text format*.

2) Subclause 2.1.1

Replace, in the definition for access unit in subclause 2.1.1:

In the case of audio, an access unit is the coded representation of an audio frame.

by:

In the case of audio, an access unit is the coded representation of an audio frame, whereby each audio frame carries data from one or more audio channels; an audio frame may carry for example one mono channel, or two stereo channels or seven surround sound channels.

In the case of an ISO/IEC 14496-17 text stream, see ISO/IEC 14496-17 for the definition of an access unit.

3) New subclause 2.4.2.9

Add, after subclause 2.4.2.8:

2.4.2.9 T-STD extensions for carriage of ISO/IEC 14496-17 text streams

To define the decoding in the T-STD of ISO/IEC 14496-17 text streams carried in a Transport Stream, the T-STD model needs to be extended. The T-STD extension and T-STD parameters for decoding of ISO/IEC 14496-17 text streams are defined in 2.15.3.1.

4) Subclause 2.4.3.5

Replace, in the semantics of "discontinuity indicator" under subclause 2.4.3.5 starting from the 5th paragraph:

For the purpose of this clause, an elementary stream access point is defined as follows:

- ISO/IEC 11172-2 video and ITU-T Rec. H.262 | ISO/IEC 13818-2 video – The first byte of a video sequence header.
- ISO/IEC 14496-2 visual – The first byte of the visual object sequence header.
- ITU-T Rec. H.264 | ISO/IEC 14496-10 video – The first byte of an AVC access unit. The SPS and PPS parameter sets referenced in this and all subsequent AVC access units in the coded video stream shall be provided after this access point in the byte stream and prior to their activation.
- Audio – The first byte of an audio frame.

--`,,```,,,,````-`-`,,`,,`,`,,`---
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

by:

For the purpose of this clause, an elementary stream access point is defined as follows:

- ISO/IEC 11172-2 video and ITU-T Rec. H.262 | ISO/IEC 13818-2 video – The first byte of a video sequence header.
- ISO/IEC 14496-2 visual – The first byte of the visual object sequence header.
- ITU-T Rec. H.264 | ISO/IEC 14496-10 video – The first byte of an AVC access unit. The SPS and PPS parameter sets referenced in this and all subsequent AVC access units in the coded video stream shall be provided after this access point in the byte stream and prior to their activation.
- Audio – The first byte of an audio frame.
- ISO/IEC 14496-17 text stream – The first byte of a text access unit. In case in-band sample descriptions are used, each in-band sample description shall be provided in the ISO/IEC 14496-17 stream after this access point and prior to its use by an access unit.

5) Subclause 2.4.3.7

a) *In subclause 2.4.3.7, replace Table 2-27*:

Table 2-27 – Stream_id_extension assignments

stream_id_extension	Note	stream coding
000 0000	1	IPMP Control Information stream
000 0001	2	IPMP stream
000 0010 … 011 1111		reserved_data_stream
100 0000 … 111 1111		private_stream
NOTE 1 – PES packets of stream_id_extension 0b000 0000 (IPMP Control Information Stream) have a unique syntax specified in ISO/IEC 13818-11 (MPEG-2 IPMP). NOTE 2 – PES packets of stream_id_extension 0b000 0001 (IPMP Stream) have a unique syntax specified in ISO/IEC 13818-11 (MPEG-2 IPMP).		

by:

Table 2-27 – Stream_id_extension assignments

stream_id_extension	Note	stream coding
000 0000	1	IPMP Control Information stream
000 0001	2	IPMP stream
000 0010 … 000 1111		ISO/IEC 14496-17 text stream
001 0000 … 011 1111		reserved_data_stream
100 0000 … 111 1111		private_stream
NOTE 1 – PES packets of stream_id_extension 0b000 0000 (IPMP Control Information Stream) have a unique syntax specified in ISO/IEC 13818-11 (MPEG-2 IPMP). NOTE 2 – PES packets of stream_id_extension 0b000 0001 (IPMP Stream) have a unique syntax specified in ISO/IEC 13818-11 (MPEG-2 IPMP).		

b) *Replace, in the semantics of PTS in subclause 2.4.3.7*:

The presentation time $t_{pn}(k)$ shall be equal to the decoding time $t_{dn}(k)$ for:

- audio access units;
- access units in ITU-T Rec. H.262 | ISO/IEC 13818-2 or ISO/IEC 14496-2 low delay video sequences;
- B-pictures in ISO/IEC 11172-2, ITU-T Rec. H.262 | ISO/IEC 13818-2 or ISO/IEC 14496-2 video streams.

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

by:

In the case of an ISO/IEC 14496-17 text stream, if a PTS is present in PES packet header, it shall refer to the first text access unit commencing in the PES packet. A text access unit commences in a PES packet if the first byte of the text access unit is present in the PES packet.

The presentation time $t_{pn}(k)$ shall be equal to the decoding time $t_{dn}(k)$ for:

- audio access units;
- access units in ITU-T Rec. H.262 | ISO/IEC 13818-2 or ISO/IEC 14496-2 low delay video sequences;
- B-pictures in ISO/IEC 11172-2, ITU-T Rec. H.262 | ISO/IEC 13818-2 or ISO/IEC 14496-2 video streams;
- text access units in ISO/IEC 14496-17.

6) Subclause 2.4.4.9

a) *Replace Table 2-34 in subclause 2.4.4.9 with the following*:

Table 2-34 – Stream type assignments

Value	Description
0x00	ITU-T | ISO/IEC Reserved
0x01	ISO/IEC 11172-2 Video
0x02	ITU-T Rec. H.262 | ISO/IEC 13818-2 Video or ISO/IEC 11172-2 constrained parameter video stream
0x03	ISO/IEC 11172-3 Audio
0x04	ISO/IEC 13818-3 Audio
0x05	ITU-T Rec. H.222.0 | ISO/IEC 13818-1 private_sections
0x06	ITU-T Rec. H.222.0 | ISO/IEC 13818-1 PES packets containing private data
0x07	ISO/IEC 13522 MHEG
0x08	ITU-T Rec. H.222.0 | ISO/IEC 13818-1 Annex A DSM-CC
0x09	ITU-T Rec. H.222.1
0x0A	ISO/IEC 13818-6 type A
0x0B	ISO/IEC 13818-6 type B
0x0C	ISO/IEC 13818-6 type C
0x0D	ISO/IEC 13818-6 type D
0x0E	ITU-T Rec. H.222.0 | ISO/IEC 13818-1 auxiliary
0x0F	ISO/IEC 13818-7 Audio with ADTS transport syntax
0x10	ISO/IEC 14496-2 Visual
0x11	ISO/IEC 14496-3 Audio with the LATM transport syntax as defined in ISO/IEC 14496-3/Amd.1
0x12	ISO/IEC 14496-1 SL-packetized stream or FlexMux stream carried in PES packets
0x13	ISO/IEC 14496-1 SL-packetized stream or FlexMux stream carried in ISO/IEC 14496_sections
0x14	ISO/IEC 13818-6 Synchronized Download Protocol
0x15	Metadata carried in PES packets
0x16	Metadata carried in metadata_sections
0x17	Metadata carried in ISO/IEC 13818-6 Data Carousel
0x18	Metadata carried in ISO/IEC 13818-6 Object Carousel
0x19	Metadata carried in ISO/IEC 13818-6 Synchronized Download Protocol
0x1A	IPMP stream (defined in ISO/IEC 13818-11, MPEG-2 IPMP)
0x1B	AVC video stream as defined in ITU-T Rec. H.264 | ISO/IEC 14496-10 Video
0x1C	ISO/IEC 14496-3 Audio, without using any additional transport syntax, such as DST, ALS and SLS
0x1D	ISO/IEC 14496-17 Text
0x1E-0x7E	ITU-T Rec. H.222.0 | ISO/IEC 13818-1 Reserved
0x7F	IPMP stream
0x80-0xFF	User Private

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

b) *Add below Table 2-34 the following clarifying text*:

In the above table various stream types are assigned for carriage of audio signals, with or without a transport syntax. Typically, the transport syntax is used for providing sync words. The use of a specific transport syntax, if at all, is specified in the clauses in this Specification specifying the transport of the various audio signals.

7) Subclause 2.5.2.4

Replace, in subclause 2.5.2.4 "PES streams":

Buffer sizes BS_n in the PES-STD model are defined as follows:

– *For ITU-T Rec. H.262 | ISO/IEC 13818-2 video*:

$$BS_n = VBV_{max}[profile, level] + BS_{oh}$$

BS_{oh} = (1/750) seconds × R_{max}[profile, level], where VBV_{max}[profile, level] and R_{max}[profile, level] are the maximum VBV size and bit rate per profile, level, and layer as defined in Tables 8-13 and 8-14, respectively, of ITU-T Rec. H.262 | ISO/IEC 13818-2. BS_{oh} is allocated for PES packet header overhead.

– *For ISO/IEC 11172-2 video*:

$$BS_n = VBV_{max} + BS_{oh}$$

BS_{oh} = (1/750) seconds × R_{max}, where R_{max} and vbv_max refer to the maximum bitrate and maximum vbv_buffer_size for a constrained parameter bitstream in ISO/IEC 11172-2 respectively.

– *For ISO/IEC 11172-3 or ISO/IEC 13818-3 audio*:

$$BS_n = 2848 \text{ bytes}$$

– *For ITU-T Rec. H.264 | ISO/IEC 14496-10 video*:

$$BS_n = 1200 \times MaxCPB[level] + BS_{oh}$$

where MaxCPB[level] is defined in Table A.1 (Level Limits) in ITU-T Rec. H.264 | ISO/IEC 14496-10 for each level.

by:

As a PES stream only carries a single elementary stream, the buffer sizes in the PES-STD do not account for multiplexing with other elementary streams, but only for multiplexing of the elementary stream carried in the PES stream with PES headers, pack headers and system headers. The buffer sizes BS_n in the PES-STD model are defined as follows:

– *For ITU-T Rec. H.262 | ISO/IEC 13818-2 video*:

$$BS_n = VBV_{max}[profile, level] + BS_{oh}$$

BS_{oh} = (1/750) seconds × R_{max}[profile, level], where VBV_{max}[profile, level] and R_{max}[profile, level] are the maximum VBV size and bit rate per profile, level, and layer as defined in Tables 8-13 and 8-14, respectively, of ITU-T Rec. H.262 | ISO/IEC 13818-2. BS_{oh} is allocated for PES packet header overhead.

– *For ISO/IEC 11172-2 video*:

$$BS_n = VBV_{max} + BS_{oh}$$

BS_{oh} = (1/750) seconds × R_{max}, where R_{max} and vbv_max refer to the maximum bitrate and maximum vbv_buffer_size for a constrained parameter bitstream in ISO/IEC 11172-2 respectively.

– *For ISO/IEC 11172-3 or ISO/IEC 13818-3 audio*:

$$BS_n = 2848 \text{ bytes}$$

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

– *For ISO/IEC 13818-7 ADTS audio*:

BS_n = 2848 bytes if 1-2 channels
BS_n = 7200 bytes if 3-8 channels
BS_n = 10800 bytes if 9-12 channels
BS_n = 43200 bytes if 13-48 channels

Note that the above numbers differ from the BS_n numbers specified in 2.4.3.2 due to the fact that a PES stream carries a single elementary stream only.

– *For ISO/IEC 14496-3 audio, except for ISO/IEC 14496-3 DST, ALS and SLS*:

BS_n = 2848 bytes if 1-2 channels
BS_n = 7200 bytes if 3-8 channels
BS_n = 10800 bytes if 9-12 channels
BS_n = 43200 bytes if 13-48 channels

Note that the above numbers differ from the BSn numbers specified in 2.11.2.2 due to the fact that a PES stream carries a single elementary stream only.

– *For ISO/IEC 14496-3 DST-64 audio*:

BS_n = 5000 × (number of channels) bytes; hence for stereo BS_n = 10 000 bytes and for 5.1 surround sound audio BS_n = 30 000 bytes

– *For ISO/IEC 14496-3 DST-128 audio*:

BS_n = 10 000 × (number of channels) bytes; hence for stereo BS_n = 20 000 bytes and for 5.1 surround sound audio BS_n = 60 000 bytes

– *For ISO/IEC 14496-3 DST-256 audio*:

BS_n = 20 000 × (number of channels) bytes; hence for stereo BS_n = 40 000 bytes and for 5.1 surround sound audio BS_n = 120 000 bytes

– *For ISO/IEC 14496-3 ALS and SLS audio*:

BS_n = 33 000 × (number of channels) bytes; hence for stereo BS_n = 66 000 bytes and for 5.1 surround sound audio BS_n = 198 000 bytes

– *For ITU-T Rec. H.264 | ISO/IEC 14496-10 video*:

$$BS_n = 1200 \times \text{MaxCPB[level]} + BS_{oh}$$

where MaxCPB[level] is defined in Table A.1 (Level Limits) in ITU-T Rec. H.264 | ISO/IEC 14496-10 for each level.

8) New subclause 2.5.2.8

Add, after subclause 2.5.2.7:

2.5.2.8 P-STD extensions for carriage of ISO/IEC 14496-17 text streams

For decoding of ISO/IEC 14496-17 text streams carried in a Program Stream in the P-STD model, see 2.15.3.2.

--`,,```,,,,````-`-`,,`,,`,`,,`---

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

9) Subclause 2.6.1

Replace Table 2-45 with the following one:

Table 2-45 – Program and program element descriptors

descriptor_tag	TS	PS	Identification
0	n/a	n/a	Reserved
1	n/a	n/a	Reserved
2	X	X	video_stream_descriptor
3	X	X	audio_stream_descriptor
4	X	X	hierarchy_descriptor
5	X	X	registration_descriptor
6	X	X	data_stream_alignment_descriptor
7	X	X	target_background_grid_descriptor
8	X	X	video_window_descriptor
9	X	X	CA_descriptor
10	X	X	ISO_639_language_descriptor
11	X	X	system_clock_descriptor
12	X	X	multiplex_buffer_utilization_descriptor
13	X	X	copyright_descriptor
14	X		maximum_bitrate_descriptor
15	X	X	private_data_indicator_descriptor
16	X	X	smoothing_buffer_descriptor
17	X		STD_descriptor
18	X	X	IBP_descriptor
19-26	X		Defined in ISO/IEC 13818-6
27	X	X	MPEG-4_video_descriptor
28	X	X	MPEG-4_audio_descriptor
29	X	X	IOD_descriptor
30	X		SL_descriptor
31	X	X	FMC_descriptor
32	X	X	external_ES_ID_descriptor
33	X	X	MuxCode_descriptor
34	X	X	FmxBufferSize_descriptor
35	X		multiplexbuffer_descriptor
36	X	X	content_labeling_descriptor
37	X	X	metadata_pointer_descriptor
38	X	X	metadata_descriptor
39	X	X	metadata_STD_descriptor
40	X	X	AVC video descriptor
41	X	X	IPMP_descriptor (defined in ISO/IEC 13818-11, MPEG-2 IPMP)
42	X	X	AVC timing and HRD descriptor
43	X	X	MPEG-2 AAC audio descriptor
44	X	X	FlexMux_Timing_descriptor
45	X	X	MPEG-4_text_descriptor
46	X	X	MPEG-4_audio_extension_descriptor
47-63	n/a	n/a	ITU-T Rec. H.222.0 | ISO/IEC 13818-1 Reserved
64-255	n/a	n/a	User Private

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

10) Subclause 2.6.39

Replace, in subclause 2.6.39:

MPEG-4_audio_profile_and_level – This 8-bit field shall identify the profile and level of the ISO/IEC 14496-3 audio stream corresponding to the Table 2-71.

by:

MPEG-4_audio_profile_and_level – This 8-bit field identifies the profile and level of the ISO/IEC 14496-3 audio stream corresponding to Table 2-71. If encoded with the value 0x0F, then it is signalled that no profile and level is defined for the associated MPEG-4 audio stream. The encoded value 0xFF indicates that the audio profile and level is not specified by the MPEG-4_audio_profile_and_level field; in that case, in addition to the MPEG-4 audio descriptor, an MPEG-4 audio extension descriptor shall be associated to the same MPEG-4 audio stream. In all other cases, next to an MPEG-4 audio descriptor, also an MPEG-4 audio extension descriptor may be associated to the same MPEG-4 audio stream.

11) Subclause 2.6.39

Replace Table 2-71 by the following table:

Table 2-71 – MPEG-4_audio_profile_and_level assignment values

Value	Description
0x00-0x0E	Reserved
0x0F	No audio profile and level defined for the associated MPEG-4 audio stream
0x10	Main profile, level 1
0x11	Main profile, level 2
0x12	Main profile, level 3
0x13	Main profile, level 4
0x14-0x17	Reserved
0x18	Scalable Profile, level 1
0x19	Scalable Profile, level 2
0x1A	Scalable Profile, level 3
0x1B	Scalable Profile, level 4
0x1C-0x1F	Reserved
0x20	Speech profile, level 1
0x21	Speech profile, level 2
0x22-0x27	Reserved
0x28	Synthesis profile, level 1
0x29	Synthesis profile, level 2
0x2A	Synthesis profile, level 3
0x2B-0x2F	Reserved
0x30	High quality audio profile, level 1
0x31	High quality audio profile, level 2
0x32	High quality audio profile, level 3
0x33	High quality audio profile, level 4
0x34	High quality audio profile, level 5
0x35	High quality audio profile, level 6
0x36	High quality audio profile, level 7
0x37	High quality audio profile, level 8
0x38	Low delay audio profile, level 1
0x39	Low delay audio profile, level 2
0x3A	Low delay audio profile, level 3
0x3B	Low delay audio profile, level 4
0x3C	Low delay audio profile, level 5

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Table 2-71 – MPEG-4_audio_profile_and_level assignment values

Value	Description
0x3D	Low delay audio profile, level 6
0x3E	Low delay audio profile, level 7
0x3F	Low delay audio profile, level 8
0x40	Natural audio profile, level 1
0x41	Natural audio profile, level 2
0x42	Natural audio profile, level 3
0x43	Natural audio profile, level 4
0x44-0x47	Reserved
0x48	Mobile audio internetworking profile, level 1
0x49	Mobile audio internetworking profile, level 2
0x4A	Mobile audio internetworking profile, level 3
0x4B	Mobile audio internetworking profile, level 4
0x4C	Mobile audio internetworking profile, level 5
0x4D	Mobile audio internetworking profile, level 6
0x4E-0x4F	Reserved
0x50	AAC profile, level 1
0x51	AAC profile, level 2
0x52	AAC profile, level 4
0x53	AAC profile, level 5
0x54-0x57	Reserved
0x58	High efficiency AAC profile, level 2
0x59	High efficiency AAC profile, level 3
0x5A	High efficiency AAC profile, level 4
0x5B	High efficiency AAC profile, level 5
0x5C-0x5F	Reserved
0x60	High efficiency AAC v2 profile, level 2
0x61	High efficiency AAC v2 profile, level 3
0x62	High efficiency AAC v2 profile, level 4
0x63	High efficiency AAC v2 profile, level 5
0x64-0xFE	Reserved
0xFF	Audio profile and level not specified by the MPEG-4_audio_profile_and_level field in this descriptor

12) New subclauses 2.6.70 to 2.6.73

Add, after subclause 2.6.69:

2.6.70 MPEG-4 text descriptor

The MPEG-4 text descriptor carries textConfig() specified in ISO/IEC 14496-17 for the associated ISO/IEC 14496-17 text stream, thereby providing basic information needed for the decoding of the associated ISO/IEC 14496-17 stream. For each ISO/IEC 14496-17 text stream carried in an ITU-T Rec. H.222.0 | ISO/IEC 13818-1 stream, the MPEG-4 text descriptor shall be included in the PMT or in the PSM, if PSM is present in the program stream.

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

Table 2-92.1 – MPEG-4 text descriptor

Syntax	No. of bits	Mnemonic

MPEG-4_text_descriptor () {
descriptor_tag
descriptor_length
textConfig()
}
	
	8
	8
		
		uimsbf
		uimsbf

2.6.71 Semantic definition of fields in MPEG-4 text descriptor

textConfig() – This shall carry the TextConfig() of the associated ISO/IEC 14496-17 text stream, as defined in ISO/IEC 14496-17.

2.6.72 MPEG-4 audio extension descriptor

The MPEG-4 audio extension descriptor carries zero or more audioProfileLevelIndication fields and zero or one audioSpecificConfig() field, both encoded as specified in ISO/IEC 14496-3. Note that for each ISO/IEC 14496-3 audio stream carried in an ITU-T Rec. H.222.0 | ISO/IEC 13818-1 stream, it is required that the MPEG-4 audio descriptor be included in the PMT or in the PSM, if PSM is present in the program stream, while the MPEG-4 audio extension descriptor may be present too, providing additional information. If in the MPEG-4 audio descriptor the MPEG-4_audio_profile_and_level field is encoded with the value 0xFF, indicating that the audio profile and level is not specified in the MPEG-4 audio descriptor, then the MPEG-4 audio extension descriptor shall be present in the same PMT or PSM as the MPEG-4 audio descriptor. Note that this descriptor allows to provide the audioSpecificConfig out of band, so as to allow receivers to retrieve information about the associated audio stream without accessing the stream itself. The descriptor also allows to associate an audioSpecificConfig to an audio stream.

Table 2-92.2 – MPEG-4 audio extension descriptor

Syntax	No. of bits	Mnemonic

MPEG-4_audio_extension_descriptor () {
descriptor_tag
descriptor_length
ASC_flag
reserved
num_of_loops
for (i=0; i<num_of_loops; i++) {
audioProfileLevelIndication
}
if (ASC_flag == '1') {
ASC_size
audioSpecificConfig()
}
}
	
	8
	8
	1
	3
	4
	
	8
	
	8
		
		uimsbf
		uimsbf
		bslbf
		bslbf
		uimsbf
		
		uimsbf
		
		uimsbf

2.6.73 Semantic definition of fields in MPEG-4 audio extension descriptor

ASC_flag – A one-bit flag signalling the presence of the ASC_size field in this descriptor.

num_of_loops – A 4-bit field specifying the number of immediately following audioprofileLevelIndication fields in this descriptor. This field may be encoded with the value zero.

audioProfileLevelIndication – The audio profile and level of the associated ISO/IEC 14496-3 audio stream, encoded as specified for the audioprofileLevelIndication field in subclause 1.5.2.1 in ISO/IEC 14496-3. Note that a single ISO/IEC 14496-3 audio stream may comply to more than one audio profile and level, and that this descriptor is designed to convey up to 15 different audioprofileLevelIndication values.

ASC_size – The number of bytes of the immediately following AudioSpecificConfig().

audioSpecificConfig() – The audioSpecificConfig() of the associated ISO/IEC 14496-3 audio stream, as specified in subclause 1.6.2.1 in ISO/IEC 14496-3.

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

13) Subclause 2.11.1

Add the following paragraph at the end of subclause 2.11.1:

Carriage of ISO/IEC 14496-17 text streams over ITU-T Rec. H.222.0 | ISO/IEC 13818-1 streams is specified in 2.15.

14) Subclause 2.11.2.1

a) *Replace, in subclause 2.11.2.1*:

In case of an ISO/IEC 14496-3 elementary stream, before PES packetization the elementary stream data shall be first encapsulated in the LATM/LOAS AudioSyncStream() transport syntax defined in ISO/IEC 14496-3. If a PTS is present in the PES packet header, it shall refer to the first audio frame that follows the first syncword that commences in the payload of the PES packet.

by:

In case of an ISO/IEC 14496-3 elementary stream signalled by a stream_type value of 0x11, before PES packetization the elementary stream data shall be first encapsulated in the LATM transport syntax defined in ISO/IEC 14496-3. In such case, If a PTS is present in the PES packet header it shall refer to the first audio frame that follows the first syncword that commences in the payload of the PES packet.

In case of an ISO/IEC 14496-3 elementary stream signalled by a stream_type value of 0x1C, the first byte of each audio frame shall be the first byte of the payload of a PES packet; prior to PES packetization no encapsulation in any additional transport syntax shall be applied. An audio frame from an ISO/IEC 14496-3 elementary stream signalled by a stream_type value of 0x1C may be fragmented for carriage in multiple PES packets. In the PES packet header the data_alignment_indicator shall be set to '1' in each PES packet that carries a complete such audio frame or the first fragment thereof. The data_alignment_indicator shall be set to '0' for PES packets carrying subsequent (non-first) fragments of an audio frame.

b) *Add, after the last paragraph of subclause 2.11.2.1*:

Carriage of an individual ISO/IEC 14496-17 text streams in PES packets shall be identified by appropriate stream_id and stream_type values, indicating the use of ISO/IEC 14496-17 text.

15) Subclause 2.11.2.2

Replace, in subclause 2.11.2.2:

In case of carriage of an ISO/IEC 14496-3 stream:

Size BS_n of Buffer B_n for ISO/IEC 14496-3 AAC audio,

$$\text{else } BS_n = BS_{mux} + BS_{dec} + BS_{oh} = 3584 \text{ bytes}$$

In this case the size of the access unit decoding buffer BS_{dec}, and the PES packet overhead buffer BS_{oh} are constrained by:

$$BS_{dec} + BS_{oh} \leq 2848 \text{ bytes}$$

A portion (736 bytes) of the 3584 byte buffer is allocated for buffering to allow multiplexing. The rest, 2848 bytes, are shared for access unit buffering BS_{dec}, BS_{oh} and additional multiplexing.

Rate Rx_n for ISO/IEC 14496-3 AAC audio same as defined for ISO/IEC 13818-7 ADTS audio in 2.4.2.3,

$$\text{else } Rx_n = 2\,000\,000 \text{ bit/s}$$

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

by:

For carriage of an ISO/IEC 14496-3 audio stream the following applies.

Size BS_n of Buffer B_n, whereby $BS_n = BS_{mux} + BS_{dec} + BS_{oh}$:

For ISO/IEC 14496-3 audio, except for ISO/IEC 14496-3 DST, ALS and SLS:

BS_n = 3584 bytes if 1-2 channels

Here, the size of the access unit decoding buffer BS_{dec}, and the PES packet overhead buffer BS_{oh} are constrained by: $BS_{dec} + BS_{oh} \leq 2848$ bytes; a portion (736 bytes) of the 3584 byte buffer is allocated for buffering to allow multiplexing. The rest, 2848 bytes, are shared for access unit buffering BS_{dec}, BS_{oh} and additional multiplexing.

BS_n = 8976 bytes if 3-8 channels

BS_n = 12804 bytes if 9-12 channels

BS_n = 51216 bytes if 13-48 channels

For ISO/IEC 14496-3 DST-64, ALS and SLS audio:

if number of audio channels <= 8 then BS_n = 1 600 000 bytes,

else BS_n = 200 000 × (number of audio channels) bytes.

For ISO/IEC 14496-3 DST-128 audio:

BS_n = 400 000 × (number of audio channels) bytes.

For ISO/IEC 14496-3 DST-256 audio:

BS_n = 800 000 × (number of audio channels) bytes.

Rate Rx_n:

For ISO/IEC 14496-3 audio, except for ISO/IEC 14496-3 DST, ALS and SLS:

Rx_n = 2 000 000 bit/s if 1-2 channels

Rx_n = 5 529 600 bit/s if 3-8 channels

Rx_n = 8 294 400 bit/s if 9-12 channels

Rx_n = 33 177 600 bit/s if 13-48 channels

For ISO/IEC 14496-3 DST-64, ALS and SLS audio:

if number of audio channels <= 8 then Rx_n = 30 000 000 bit/s,

else Rx_n = 120 000 000 bit/s.

For ISO/IEC 14496-3 DST-128 and DST-256 audio:

Rx_n = 120 000 000 bit/s.

16) New subclause 2.15

Add the following new clause:

2.15 Carriage of ISO/IEC 14496-17 text streams

2.15.1 Introduction

This specification defines the carriage of ISO/IEC 14496-17 elementary text streams within ITU-T Rec. H.222.0 | ISO/IEC 13818-1 systems, both for program and transport streams. Typically, an ISO/IEC 14496-17 text stream will be an element of an ISO/IEC 13818-1 program, as defined by the PMT in a Transport Stream and the PSM in a Program Stream. The carriage and buffer management of ISO/IEC 14496-17 text streams is defined using existing parameters from ITU-T Rec. H.222.0 | ISO/IEC 13818-1 such as PTS and DTS, as well as information from the ISO/IEC 14496-17 text stream. For this purpose, clause 2.15.3 specifies the decoding of ISO/IEC 14496-17 streams within the T-STD and P-STD models, using the Hypothetical Text Decoder (HTD) defined in ISO/IEC 14496-17.

When an ISO/IEC 14496-17 text stream is carried in an ITU-T Rec. H.222.0 | ISO/IEC 13818-1 stream, the ISO/IEC 14496-17 coded data shall be contained in PES packets, as defined in 2.15.2. Information needed to decode the ISO/IEC 14496-17 text stream is provided by the MPEG-4 text descriptor, specified in 2.6.70.

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

Applications that wish to transmit font streams as specified in ISO/IEC 14496-18, can reference the streamed fonts as specified in Annex A1 of ISO/IEC 14496-17. To transmit ISO/IEC 14496-18 font streams, applications can use the data or object carousel as defined in ISO/IEC 13818-6.

2.15.2 Carriage in PES packets

ISO/IEC 14496-17 text streams are carried in PES packets, using one of the assigned 14 stream_id_extension values. If the textFormat field in TextConfig() is coded with the value 0x01, indicating Timed Text as specified in 3GPP TS 26.245, then the text stream consists of a concatenation of 3GPP Text Access Units in display order, whereby each Text Access Units consists of one or more so-called TTUs. For carriage in PES packets, alignment between PES packets and TTUs is required; i.e., the first byte in the payload of a PES packet carrying a ISO/IEC 14496-17 text stream shall be the first byte of a TTU, which is the first byte of the TTU header specified in clause 7.4.2 in ISO/IEC 14496-17. There is no further need for alignment; hence, the first TTU in a PES packet may contain for example a non-first fragment of a Text Access unit or a Sample Description.

A PTS shall be coded in the PES packet header of each PES packet carrying an ISO/IEC 14496-17 text stream. The PTS shall refer to the first Text Access Unit that commences in the PES packet. A Text Access Unit commences in a PES packet if a TTU[1] with a complete text sample or the TTU[2] with the first fragment of a text sample is present in the PES packet. For identifying whether a TTU[2] carries the first or a subsequent text sample fragment, see clauses 7.3.2.2 and 7.4.5 in ISO/IEC 14496-17.

2.15.3 STD extensions

2.15.3.1 T-STD extensions

The T-STD model includes a transport buffer TB_n and a multiplex buffer B_n prior to the Hypothetical Text Decoder defined in ISO/IEC 14496-17. Transport Stream packets with ISO/IEC 14496-17 data, as indicated by its PID, enter the buffer TB_n. Bytes are removed from TB_n to enter Buffer B_n at the rate Rx_n. Delivery of bytes from B_n to enter the TTU Decoder in the HTD is at rate Rbx_n. See also Figure 2-13.

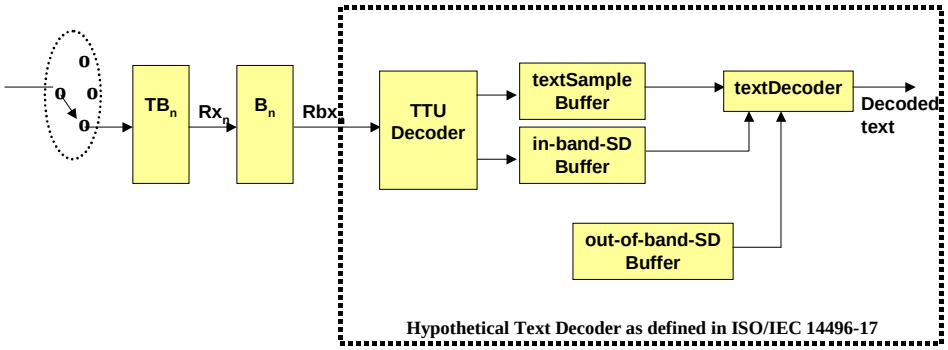


Figure 2-13 – T-STD model extensions for ISO/IEC 14496-17 text streams

Buffer management

The size TBS_n of TB_n is equal to 512 bytes. For ISO/IEC 14496-17 text streams, the size BS_n of B_n is equal to 4096 bytes. The rate Rx_n is equal to 2 000 000 bits/s. The rate Rbx_n meets the delivery schedule defined for ISO/IEC 14496-17 data at the input of the HTD in clause 7.7 of ISO/IEC 14496-17, i.e., Rbx_n = R[profile, level] if the textSampleBuffer is not full and Rbx_n = 0 if the textSampleBuffer is full, where R[profile, level] is defined as the profile and level specific rate R in clause 7.8 of ISO/IEC 14496-17.

Each of the following requirements shall apply:

- Buffer TB_n shall not overflow.
- Buffer B_n shall not overflow.
- Each HTD requirement specified in clause 7.7 of ISO/IEC 14496-17 shall be met.

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS Not for Resale

2.15.3.2 P-STD extensions

The P-STD model for the decoding of an ISO/IEC 14496-17 text stream includes a multiplex buffer B_n prior to the Hypothetical Text Decoder defined in ISO/IEC 14496-17. For each ISO/IEC 14496-17 text stream n, the size BS_n of buffer B_n in the P-STD is defined by the P-STD_buffer_size field in the PES packet header.

Each of the following requirements shall apply:

- Buffer B_n shall not overflow.
- Each HTD requirement specified in clause 7.7 of ISO/IEC 14496-17 shall be met.

17) Bibliography

Add Bibliography and the following:

- ISO/IEC 14496-18:2004, *Information technology – Coding of audio-visual objects – Part 18: Font compression and streaming*.

Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale

ICS 35.040

Price based on 13 pages

© ISO/IEC 2007 – All rights reserved
Copyright International Organization for Standardization
Provided by IHS under license with ISO
No reproduction or networking permitted without license from IHS
Not for Resale