INTERNATIONAL STANDARD

ISO/IEC 13818-1

First edition
1996-04-15

AMENDMENT 6
2000-04-15

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

AMENDMENT 6

Technologies de l'information — Codage générique des images animées et du son associé: Systèmes
AMENDEMENT 6

Reference number
ISO/IEC 13818-1:1996/Amd.6:2000(E)

ISO copyright office
Case postale 56 • CH-1211 Geneva 20
Tel. + 41 22 749 01 11
Fax + 41 22 734 10 79
E-mail copyright@iso.ch
Web www.iso.ch

Printed in Switzerland

Foreword

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

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 3.

In the field of information technology, ISO and IEC have established a joint technical committee, ISO/IEC JTC 1. 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 Amendment may be the subject of patent rights. ISO and IEC shall not be held responsible for identifying any or all such patent rights.

Amendment 6 to International Standard ISO/IEC 13818-1:1996 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/Amd.6.

INTERNATIONAL STANDARD

ITU-T RECOMMENDATION

INFORMATION TECHNOLOGY – GENERIC CODING OF MOVING PICTURES AND ASSOCIATED AUDIO INFORMATION: SYSTEMS

AMENDMENT 6

1) Subclause 2.4.2.3

a) *Replace the following:*

– For audio:

$$Rx_n = 2 \times 10^6 \text{ bits per second}$$

with:

– For ISO/IEC 13818-7 ADTS audio:

Number of channels	Rx_n [bit/s]
1-2	2 000 000
3-8	5 529 600
9-12	8 294 400
13-48	33 177 600

Channels: channels that require their own decoder buffer in this elementary stream (i.e. individual channel streams in a single channel element or channel pair element and independently switched coupling channel elements).

– For other audio:

$$Rx_n = 2 \times 10^6 \text{ bits per second}$$

b) *After the line* The main buffer sizes BS_1 through BS_n are defined as follows, *replace the following:*

Audio

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

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

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

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.

with:

Audio

For ISO/IEC 13818-7 ADTS audio:

Number of channels	BS_n [bytes]
1-2	3 584
3-8	8 976
9-12	12 804
13-48	51 216

Channels: channels that require their own decoder buffer in this elementary stream (i.e. individual channel streams in a single channel element or channel pair element and independently switched coupling channel elements).

For other audio:

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

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.

2) Subclause 2.4.3.5

Add the following table:

Table 2-16-4 – Splice parameters Table 14

4:2:2 Profile High Level Video

Splice_type	Conditions
0000	splice_decoding_delay = 45 ms; max_splice_rate = 300.0×10^6 bit/s
0001	splice_decoding_delay = 90 ms; max_splice_rate = 300.0×10^6 bit/s
0010-0011	Reserved
0100	splice_decoding_delay = 250 ms; max_splice_rate = 180.0×10^6 bit/s
0101-1011	Reserved
1100-1111	User-defined

3) Subclause 2.7.9

Add the following as the last paragraph of Decoder Buffer Size*:*

In the case of ISO/IEC 13818-7 ADTS audio elementary stream in a CSPS, the following applies to support 8 channels:

$$BS_n \leq 8976 \text{ bytes}$$

4) New Annex Q

Add the following after Annex P:

Annex Q

T-STD and P-STD buffer models for ISO/IEC 13818-7 ADTS

(This annex does not form an integral part of this Recommendation | International Standard)

Q.1 Introduction

The Transport Stream system target decoder model for audio streams is defined in 2.4.2. In this annex, the buffer model for ISO/IEC 13818-7 ADTS is described.

ISO/IEC 13818-7 ADTS audio streams can be recognized in an ITU-T Rec. H.222.0 | ISO/IEC 13818-1 multiplex through the presence of stream_id = 0x110yyyyy ('y' = "don't care") and stream_type = 0x0F as defined in Amendment 5 to ITU-T Rec. H.222.0 | ISO/IEC 13818-1.

Q.2 Leak rate from transport buffer

For audio except ISO/IEC 13818-7 ADTS, the leak rate from Transport Buffer is 2 Mbit/s. This rate is, however, lower than the maximum rate of ISO/IEC 13818-7 ADTS. Therefore, the leak rate for the ISO/IEC 13818-7 ADTS stream is set to a different value from ISO/IEC 11172-3 and ISO/IEC 13818-3 audio streams.

ISO/IEC 13818-7 ADTS elementary stream consists of one or more channels. The maximum rate of each channel is 576 kbit/s where the sampling frequency is 96 kHz. Therefore, the leak rate for ISO/IEC 13818-7 ADTS is calculated as following equation:

$$Rx_n = 1.2 \times R_{max} \times N \text{ bits per second}$$

where:

R_{max} is a constant 576 kbit/s as defined in 3.2.2 of ISO/IEC 13818-7. It is an upper bound of the bit rate per channel of AAC ADTS stream corresponding to the maximum value of sampling frequency (i.e. $F_s = 96$ kHz).

and where:

N is the number of audio channels that require their own decoder buffer in this elementary stream (i.e. individual channel streams in a single channel element or channel pair element and independently switched coupling channel elements).

Q.3 Buffer size

For audio except ISO/IEC 13818-7 ADTS, the main buffer size is 3584 bytes. This size is, however, smaller than the maximum decoder input buffer size of ISO/IEC 13818-7 ADTS. Therefore, the main buffer size for the ISO/IEC 13818-7 ADTS stream is set to a different value from ISO/IEC 11172-3 and ISO/IEC 13818-3 audio streams.

The main buffer size for ISO/IEC 13818-7 ADTS is calculated as follows:

$$BS_n = BS_{mux} + BS_{dec} + BS_{oh}$$

where BS_{oh}, PES packet overhead buffering, is defined as:

$$BS_{oh} = 528 \text{ bytes}$$

and BS_{mux}, additional multiplexing buffering, is defined as:

$$BS_{mux} = 0.004 \text{ s} \times R_{max} \times N$$

and BS_{dec}, access unit buffering, is defined as:

$$BS_{dec} = 6144 \text{ bits} \times N$$

where:

R_{max} is a constant 576 kbit/s as defined in 3.2.2 of ISO/IEC 13818-7. It is an upper bound of the bit rate per channel of AAC ADTS stream corresponding to the maximum value of sampling frequency (i.e. $F_s = 96$ kHz).

and where:

N is the number of audio channels that require their own decoder buffer in this elementary stream (i.e. individual channel streams in a single channel element or channel pair element and independently switched coupling channel elements).

Q.3.1 TBS_n: same as other audio

In terms of the smoothing buffer, there is no difference in TB_n between ISO/IEC 13818-7 ADTS and other audio streams. Consequently, it is not necessary to change TBS_n, which is size of TB_n.

Q.3.2 BS_{mux}: different from other audio

BS_{mux}, additional multiplexing buffering, shall be changed to accept up to 4 ms of delay jitter. This is similar to the approach taken for other streams in ITU-T Rec. H.222.0 | ISO/IEC 13818-1.

Q.3.3 BS_{dec}: different from other audio

BS_{dec}, access unit buffering, is based on the decoder input buffer size of the elementary stream. As defined in 3.2.2 of ISO/IEC 13818-7, total decoder input buffer size is 6144 bit multiplied by the number of channels which each require their decoder input buffer.

Q.3.4 BS_{oh}: different from other audio

BS_{oh} corresponds to the PES packet header overhead.

In 2.4.2.6:

The delay of any data through the System Target Decoders buffers shall be less than or equal to one second except for still picture video data.

Besides, in 2.7.4:

The Program Stream and Transport Stream shall be constructed so that the maximum difference between coded presentation timestamps referring to each elementary video or audio stream is 0.7 s.

BS_{oh} shall be set to the appropriate size corresponding to the PES packet header overhead when AAC stream is packetized with the above rules. The maximum size of PES packet header is 264 bytes. Therefore, $BS_{oh} = 528$ bytes, i.e. twice the maximum size of PES packet header, assures that at least two PES packet headers can enter the main buffer regardless of the size of PES packet header. It means that PES packet header with PTS can be inserted at less than 0.7 s intervals even when the data of one second will be in the main buffer.

Example: sampling frequency is 48 kHz

The size of PES packet header without any optional fields except PTS is 18 bytes. The number of Access Units of one second is about 47. When the data of one second is in the main buffer (i.e. the worst case), PES packet header overhead can fit to the BS_{oh} with packetizing more than or equal to two Access Units into one packet.

$$\text{number\_of\_AU} = 48 \text{ kHz}/1024 = 46\,875 \text{ per second}$$

$$(\text{number\_of\_AU}/2) \times 18[\text{byte}] = 421\,875 \text{ bytes} < BS_{oh}$$

More frequent PES packet headers can be fit to BS_{oh}, if the delay of any data through the main buffer is shorter than one second.

Q.4 Conclusion

The decoder buffer model should cover the maximum size of buffer, however AAC can handle up to 48 channels and very high bit rate. Therefore the 3 levels of number of channels, 2, 8 and 48, are used to define the leak rate and the main buffer size. In a case of 2, the same leak rate and main buffer size as the conventional values are used to keep the compatibility. In other cases (8 and 48), the proposed formulas are applied.

T-STD leak rate for ISO/IEC 13818-7 ADTS audio

Number of channels	Rx_n [bit/s]
1-2	2 000 000
3-8	5 529 600
9-12	8 294 400
13-48	33 177 600

Channels: channels that require their own decoder buffer in this elementary stream (i.e. individual channel streams in a single channel element or channel pair element and independently switched coupling channel elements).

T-STD main buffer size for ISO/IEC 13818-7 ADTS audio

Number of channels	BS_n [bytes]
1-2	3 584
3-8	8 976
9-12	12 804
13-48	51 216

Channels: channels that require their own decoder buffer in this elementary stream (i.e. individual channel streams in a single channel element or channel pair element and independently switched coupling channel elements).

For Program Stream, the above main buffer size should be set in the P-STD_buffer_scale and P-STD_buffer_size as follows:

Number of channels	P-STD_buffer_scale	P-STD_buffer_size
1-2	0	28
3-8	0	71
9-48	0	401

ICS 35.040

Price based on 5 pages