FINAL
DRAFT

AMENDMENT

ISO/IEC 13818-1:2021 FDAM 1

ISO/IEC JTC **1**/SC **29**

Secretariat: **JISC**

Voting begins on:
2022-07-25

Voting terminates on:
2022-09-19

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

Part 1: Systems

AMENDMENT 1: Carriage of LCEVC and other improvements

Partie 1: Systèmes

AMENDEMENT 1: Titre manque

Reference number
ISO/IEC 13818-1:2021/FDAM 1:2022(E)

COPYRIGHT PROTECTED DOCUMENT

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

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.

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 ITU-T as ITU-T H.222.0 (08/2018) and drafted in accordance with its editorial rules, in collaboration with Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*.

A list of all parts in the ISO/IEC 13818 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.

INTERNATIONAL STANDARD

ITU-T RECOMMENDATION

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

Amendment 1

Carriage of LCEVC and other improvements

1) Clause 1.2.3

In 1.2.3, Additional References, add the following references:

- ISO/IEC 23094-2:20XX, Information technology – General Video Coding – Part 2: Low Complexity Enhancement Video Coding

- BCP47: IETF RFC 4647 Matching of Language Tags September 2006 / IETF RFC 5646 Tags for Identifying Languages September 2009

2) Clauses 2.1.90bis to 2.1.90ter

After clause 2.1.90, add the following clauses 2.1.90bis to 2.1.90ter:

2.1.90bis LCEVC access unit (system): An access unit as defined for byte streams in ISO/IEC 23094-2 with the constraints specified in clause 2.25.1.

2.1.90ter LCEVC video sequence (system): coded video sequence as defined in ISO/IEC 23094-2.

3) Clause 2.4.2.7

In clause 2.4.2.7, replace the fourth paragraph:

The delay of any data through the System Target Decoder buffers shall be less than or equal to one second except for still picture video data, ISO/IEC 14496 streams, ISO/IEC 23008-2 streams, ISO/IEC 23090-3 streams and ISO/IEC 23094-1 streams. Specifically: $td_n(j) - t(i) \leq 1$ second for all j, and all bytes i in access unit $A_n(j)$.

with:

The delay of any data through the System Target Decoder buffers shall be less than or equal to one second except for still picture video data, ISO/IEC 14496, ISO/IEC 23008-2, ISO/IEC 23090-3, ISO/IEC 23094-1 and ISO/IEC 23094-2 streams. Specifically: $td_n(j) - t(i) \leq 1$ second for all j, and all bytes i in access unit $A_n(j)$.

Further replace the seventh paragraph:

For ISO/IEC 14496, ISO/IEC 23008-2, ISO/IEC 23090-3 and ISO/IEC 23094-1 streams, the delay is constrained by $td_n(j) - t(i) \leq 10$ seconds for all j, and all bytes i in access unit $A_n(j)$.

with:

For ISO/IEC 14496, ISO/IEC 23008-2, ISO/IEC 23090-3, ISO/IEC 23094-1 and ISO/IEC 23094-2 streams, the delay is constrained by $td_n(j) - t(i) \leq 10$ seconds for all j, and all bytes i in access unit $A_n(j)$.

4) Clause 2.4.3.7

In clause 2.4.3.7, in Table 2-22, replace:

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, Rec. ITU-T H.265 | ISO/IEC 23008-2, Rec. ITU-T H.26X | ISO/IEC 23090-3 or ISO/IEC 23094-1 video stream number xxxx

with:

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, Rec. ITU-T H.265 | ISO/IEC 23008-2, Rec. ITU-T H.26X | ISO/IEC 23090-3, ISO/IEC 23094-1 or ISO/IEC 23094-2 video stream number xxxx

Further, in the section specifying the PTS (presentation time stamp), insert the following new paragraph after the paragraph starting with "For EVC video streams …":

For LCEVC video streams, the PTS shall be present in the PES packet header and it shall refer to only one LCEVC access unit that commences in this PES packet. To achieve consistency between the STD model and the HRD model defined in Annex C of ISO/IEC 23094-2, for each LCEVC access unit the PTS value in the STD shall, within the accuracy of their respective clocks, indicate the same instant in time as the nominal DPB output time in the HRD, as defined in Annex C of ISO/IEC 23094-2.

Further, in the section specifying the DTS (decoding time stamp), insert the following new paragraph after the paragraph starting with "For EVC video streams …":

For LCEVC video streams, DTS shall not be present in the PES packet header, because the LCEVC decoding process is strictly in presentation order. This restriction on the strict coincidence of decoding order and presentation order of the access units refers exclusively to the LCEVC video stream, while the base video stream that the LCEVC video stream enhances has no such restriction, that is, the decoding order and presentation order can be different.

5) Clause 2.4.4.10

In 2.4.4.10 in Table 2-34, replace:

0x36 .. 0x7E	Rec. ITU-T H.222.0 | ISO/IEC 13818-1 reserved

with:

0x36	LCEVC video stream conforming to one or more profiles defined in ISO/IEC 23094-2
0x37 .. 0x7E	Rec. ITU-T H.222.0 | ISO/IEC 13818-1 reserved

6) Clause 2.6.90

In 2.6.90 in Table 2-109, replace:

```
else if ( extension_descriptor_tag == 0x16) {
    EVC_timing_and_HRD_descriptor()
}
```

with:

```
else if ( extension_descriptor_tag == 0x16) {
    EVC_timing_and_HRD_descriptor()
}
else if ( extension_descriptor_tag == 0x17) {
    LCEVC_video_descriptor ()
}
else if ( extension_descriptor_tag == 0x18) {
    LCEVC_linkage_descriptor()
}
else if ( extension_descriptor_tag == 0x19) {
    Media_service_kind_descriptor()
}
```

7) Clause 2.6.91

In clause 2.6.91, before Table 2-110, add:

LCEVC_video_descriptor () – This structure is defined in 2.6.137 and 2.6.138.

LCEVC_linkage_descriptor () – This structure is defined in 2.6.139 and 2.6.140.

Media_service_kind_descriptor() – This structure is defined in 2.6.141 and 2.6.142

Further, in Table 2-110, replace:

0x17 .. 0xFF	n/a	n/a	Rec. ITU-T H.222.0 | ISO/IEC 13818-1 Reserved

with

0x17	X	X	LCEVC_video_descriptor
0x18	X	X	LCEVC_linkage_descriptor
0x19	X	X	Media_service_kind_descriptor
0x1A .. 0xFF	n/a	n/a	Rec. ITU-T H.222.0 | ISO/IEC 13818-1 Reserved

8) Clause 2.6.127

In clause 2.6.127, in Table 2-132, remove the **descriptor_tag** *and* **descriptor_length** *fields.*

9) Clause 2.6.137 to 2.6.142

After 2.6.136, add the following clauses 2.6.137 to 2.6.142:

2.6.137 LCEVC video descriptor

Table 2-141bis– LCEVC video descriptor

Syntax	No. Of bits	Mnemonic
LCEVC_video_descriptor() {		
lcevc_stream_tag	**8**	**uimsbf**
profile_idc	**4**	**uimsbf**
level_idc	**4**	**uimsbf**
sublevel_idc	**2**	**uimsbf**
processed_planes_type_flag	**1**	**bslbf**

picture_type_bit_flag	**1**	**bslbf**
field_type_bit_flag	**1**	**bslbf**
reserved	**3**	**bslbf**
HDR_WCG_idc	**2**	**uimsbf**
reserved_zero_2bit	**2**	**bslbf**
video_properties_tag	**4**	**uimsbf**
}		

2.6.138 Semantic definition of fields of LCEVC video descriptor

lcevc_stream_tag – This is a 8-bit field specifying the identifier of an association between base and enhancement streams of an LCEVC encoding.

profile_idc, level_idc, sublevel_idc – These fields shall be coded according to the semantics defined in ISO/IEC 23094-2 for profile_idc, level_idc, sublevel_idc, in clause 7.3.4 (SC, sequence_configuration).

processed_planes_type_flag – This field shall be coded according to the semantics defined in ISO/IEC 23094-2 for processed_planes_type_flag, in clause 7.3.5 (GC, global_configuration).

picture_type_bit_flag, field_type_bit_flag – These fields shall be coded according to the semantics defined in ISO/IEC 23094-2 for picture_type_bit_flag, field_type_bit_flag in clause 7.3.6 (PC, picture_configuration).

HDR_WCG_idc – This 2-bit field indicates the presence or absence of high dynamic range (HDR) and/or wide color gamut (WCG) video components in the associated PID according to Table 2-134.

video_properties_tag – This 4-bit field is used to indicate specific widely used video property CICP combinations as indicated by Rec. ITU-T H.Sup 19 | ISO/IEC 23091-4 for SDR, WCG, or HDR/WCG streams depending on the value of HDR_WCG_idc. When HDR_WCG_idc is equal to 0, Table 2-135 applies. When HDR_WCG_idc is equal to 1, Table 2-136 applies. When HDR_WCG_idc is equal to 2, Table 2-137 applies. When HDR_WCG_idc is equal to 3, Table 2-138 applies.

NOTE 1 – Since the LCEVC enhancement stream and the base video stream that is enhanced are encoded separately, there can be separate values for HDR_WCG_idc and HDR_WCG_idc in the base and in the enhancement streams, and such values are not necessarily identical.

NOTE 2 – There is no Buffering Period SEI and Picture Timing SEI message defined for LCEVC, because the ones from the base encoding are used. For this reason, a LCEVC Timing and HRD Descriptor is not needed.

2.6.139 LCEVC linkage descriptor

Table 2-141ter– LCEVC linkage descriptor

Syntax	No. Of bits	Mnemonic
LCEVC_linkage_descriptor() {		
num_lcevc_stream_tags	**8**	**uimsbf**
for (i=0; i<num_lcevc_stream_tags; i++) {		
lcevc_stream_tag	**8**	**uimsbf**
}		
}		

2.6.140 Semantic definition of fields of LCEVC linkage descriptor

num_lcevc_stream_tags – The number of lcevc_stream_tag items that will follow.

lcevc_stream_tag – Tag value that allows indicating this video elementary stream as the base of an LCEVC video stream that carries the same tag value in its LCEVC video descriptor.

2.6.141 Media service kind descriptor

The media_service_kind_descriptor can read on the Program or PID ES stream itself or can read on auxiliary PID ES that are dependent on the principal PID media component ES (i.e. in most cases the principal media component is video). The media service kind descriptor can also read upon the entire program to describe content with a

single principal media component. Multiple data elements (N) can be included in the descriptor each with one or more language code/media_service_type pairs.

If the same media component is described at a Program Level and an Elementary Level, then the Elementary stream media_service_kind description shall take precedence.

Table 2-141quater – Media service kind descriptor

Syntax	No. of bits	Mnemonic
Media_service_kind_descriptor() {		
for (i=0; i<N; i++) {		
media_description_flag	**1**	**bslbf**
identifier_flag	**1**	**bslbf**
lang_pairs	**3**	**uimbsf**
media_type_idc	**2**	**uimbsf**
reserved	**1**	**bslbf**
if (identifier_flag) {		
ID_length_code	**3**	**uimbsf**
ID_type	**13**	**uimbsf**
if (ID_length_code = = '7') {		
ID_len	**8**	**uimbsf**
}		
media_ID_field	**ID_len*8**	**bslbf**
}		
for (j= 0; j<lang_pairs; j++) {		
configuration_type	**2**	**uimsbf**
lang_purpose_cnt	**3**	**uimsbf**
lang_len_idc	**2**	**uimsbf**
reserved	**1**	**bslbf**
if (lang_len_idc = = 0) {		
lang_len	**8**	**uimsbf**
}		
IETF_BCP_47_language_code	**lang_len*8**	**bslbf**
for (k=0; k<lang_purpose_cnt; k++) {		
media_service_type	**8**	**uimsbf**
}		
}		
}		
}		

2.6.142 Semantic definition of fields in media service kind descriptor

media_description_flag – This flag indicates if the specific entry in the descriptor applies to the elementary stream or program itself or indicates an entry for another stream associated with the elementary stream/program which is needed to produce a specific purpose of the program. The flag will have the values of "1-self" or "2-associate". In the first loop, M would stand for the number of self or auxiliary media_type described in the kind-media-service descriptor on that specific elementary stream/program. Associate descriptions can indicate properties of dependent media components to the principal media component that can be often in a demuxed elementary stream. For example, a program or video elementary stream could associate a dependent audio stream that indicates the native language of the produced work. Given a candidate selection of demuxed audio streams, the best audio-video synchronization experience for audio and text could then be selected. The principal video elementary stream can also indicate dependent video elementary streams such as a video signing language track that maybe needed as an overlay to the video track.

identifier_flag – A flag to indicate if the media_service_kind descriptor data element will carry an identifier.

lang_pairs – This field indicates the number of language code / media_service_type pairs carried with the data element.

media_type_idc – This indicator would have values in the following table "1-Video, 2-Audio, 3-Text" as indicated by table 2-141quinquies.

Table 2-141quinquies – Media type indicator

media_type_idc	Media Type Value
0	unknown
1	video
2	audio
3	text/data

ID_length_code – This parameter assigns to ID_len the number of bytes from the list of potential byte numbers needed for the carriage of the optional identifier. When the parameter has a value of 7 then the ID_len can be configured to other byte length numbers.

Table 2-141sexies – ID_length_code

Value	Length, bytes (*ID_len)*
0	1
1	2
2	4
3	8
4	12
5	16
6	20
7	Explicit value provided in field ID_len

ID_type- parameter to indicate the type of identifier used according to Table 2-141septies.

NOTE – This parameter can originate from a list of SCTE UPID identifier list. Alternatively, identifiers can be reserved outside of the UPID identifier list.

Table 2-141septies – ID_type

Range	Definition
0x0000 .. 0x01FF	Reserved for MPEG standardization
0x0200 .. 0x02FF	Value defined in ANSI/SCTE 35 Table 22 (segmentation_upid_type) + 0x200
0x0300 .. 0x0FFF	Reserved for MPEG standardization
0x1000 .. 0x1FFF	User Private types

ID_len- the number of bytes needed to carry an optional identifier in the data element of the descriptor**.**

media_ID_field- the parameter carrying the identifier in the data element of the descriptor.

configuration_type – This parameter as defined in Table 2-141octies indicates if the elementary stream or associate stream is complete (value= 0 e.g. a 5.1 audio stream) and playable as an independent media component or is partial (value= 1 e.g. M&E or dialogue only) and not playable alone as the media component stream. A complete combination (value= 2 e.g. 5.1/2.0 stream with multiple dialogue languages) contains several variants with each version playable as an independent media component. Internal coding configurations if defined are deferred to the elementary specific format metadata.

Table 2-141octies – configuration type values

configuration_type	configuration type value
0	complete
1	partial
2	complete combination
3	reserved

configuration_type value semantics
complete: describes a media component stream that is a single version playable video/audio/text experience otherwise known as complete or complete main

partial: describes a media component stream that is not a playable video/audio/text experience without additional streams

complete combination: describes a media component stream that contains multiple playable video/audio/text experiences which can be extracted through different parameter settings

lang_purpose_cnt– This parameter indicates the number of media_service_types (up to 6) associated with a single language code. If the count is zero it means to use the default media_service_type of "und" and if no "main" exists for the track then this defaults to "main". A value of 7 is reserved for future purposes,

lang_len_idc– Indicator to determine if a default length is used to indicate the IETF_BCP_47_language_code field length. If indicator is 0 then char length is assignable by a byte value. When set to 1 or 2, the value of lang_len is given in Table 2-141nonies. A value of 3 is reserved for future use.

Table 2-141nonies – lang_len_indicator

lang_length_idc	lang_len
0	No default length. Explicit value provided in field lang_len
1	2
2	3
3	reserved

IETF_BCP47_language_code – Identifies the language or languages including sign language used by the associated program element. The IETF_BCP47_language_code contains a multi-character code as specified by IETF BCP 47 where the number of characters is determined by the value in the preceding lang_len field. IETF BCP 47 incorporates ISO 639-3 language code with additional description. Each character is coded into 8 bits according to ISO 8859-1 and inserted in order into this field. In the case of multilingual audio streams the sequence of ISO_639_language_code fields shall reflect the content of the audio stream or designated as 'mul' three-character subtag for dynamic multilingual cases but in the case that the descriptor needs to remain static. In the case of video components that are not sign language,'zxx' can be used as a three-character primary language subtag.

media_service_type – This 8-bit field specifies the purpose(s) of the stream(s) described in the descriptor as defined in Table 2-141decies.

Table 2-141decies – Media service type values

Value	Description
0x00	undefined
0x01	main
0x02	alternate