Recommendation
ITU-T H.222.0 (2021) Amd. 1 (12/2022)

SERIES H: Audiovisual and multimedia systems

Infrastructure of audiovisual services – Transmission multiplexing and synchronization

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

Amendment 1 – Carriage of LCEVC and other improvements

ITU-T H-SERIES RECOMMENDATIONS

AUDIOVISUAL AND MULTIMEDIA SYSTEMS

For further details, please refer to the list of ITU-T Recommendations.

INTERNATIONAL STANDARD ISO/IEC 13818-1
RECOMMENDATION ITU-T H.222.0

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

Amendment 1

Carriage of LCEVC and other improvements

Summary

Rec. ITU-T H.222.0 | ISO/IEC 13818-1 was developed in 1994 to principally support the combination and synchronization of video and audio coding methods defined in ISO/IEC 13818 Part 2 (Rec. ITU-T H.262) and Part 3. It has since then been extended to support additional video and audio coding specifications as well as multiple kinds of metadata.

Amendment 1 to ITU-T H.222.0 (2021) | ISO/IEC 13818-1:2022 extends the specification by defining how LCEVC (ISO/IEC 23094-2) is carried over MPEG-2 systems. It also defines an additional descriptor signalling the kind of media service and its usage. Further, it includes clarifications for the specification of carriage of JPEG XS.

All of this is done in a way that is compatible with the existing support for other codecs.

History

Edition	Recommendation	Approval	Study Group	Unique ID*
1.0	ITU-T H.222.0	1995-07-10	15	[11.1002/1000/1071](http://handle.itu.int/11.1002/1000/1071)
1.1	ITU-T H.222.0 (1995) Amd. 1	1996-11-11	16	[11.1002/1000/3834](http://handle.itu.int/11.1002/1000/3834)
1.2	ITU-T H.222.0 (1995) Amd. 2	1996-11-11	16	[11.1002/1000/4096](http://handle.itu.int/11.1002/1000/4096)
1.3	ITU-T H.222.0 (1995) Technical Cor. 1	1998-02-06	16	[11.1002/1000/4532](http://handle.itu.int/11.1002/1000/4532)
1.4	ITU-T H.222.0 (1995) Amd. 3	1998-02-06	16	[11.1002/1000/4228](http://handle.itu.int/11.1002/1000/4228)
1.5	ITU-T H.222.0 (1995) Amd. 4	1998-02-06	16	[11.1002/1000/4229](http://handle.itu.int/11.1002/1000/4229)
1.6	ITU-T H.222.0 (1995) Amd. 5	1999-05-27	16	[11.1002/1000/4498](http://handle.itu.int/11.1002/1000/4498)
1.7	ITU-T H.222.0 (1995) Amd. 6	1999-05-27	16	[11.1002/1000/4671](http://handle.itu.int/11.1002/1000/4671)
1.8	ITU-T H.222.0 (1995) Amd. 7	2000-02-17	16	[11.1002/1000/4926](http://handle.itu.int/11.1002/1000/4926)
2.0	ITU-T H.222.0	2000-02-17	16	[11.1002/1000/5142](http://handle.itu.int/11.1002/1000/5142)
2.1	ITU-T H.222.0 (2000) Technical Cor. 1	2001-03-01	16	[11.1002/1000/5419](http://handle.itu.int/11.1002/1000/5419)
2.2	ITU-T H.222.0 (2000) Technical Cor. 2	2002-03-29	16	[11.1002/1000/5675](http://handle.itu.int/11.1002/1000/5675)
2.3	ITU-T H.222.0 (2000) Amd. 1	2002-12-14	16	[11.1002/1000/6190](http://handle.itu.int/11.1002/1000/6190)
2.4	ITU-T H.222.0 (2000) Amd. 1/Cor. 1	2003-06-29	16	[11.1002/1000/6449](http://handle.itu.int/11.1002/1000/6449)
2.5	ITU-T H.222.0 (2000) Amd. 2	2003-06-29	16	[11.1002/1000/6363](http://handle.itu.int/11.1002/1000/6363)
2.6	ITU-T H.222.0 (2000) Amd. 3	2004-03-15	16	[11.1002/1000/7208](http://handle.itu.int/11.1002/1000/7208)
2.7	ITU-T H.222.0 (2000) Technical Cor. 3	2005-01-08	16	[11.1002/1000/7435](http://handle.itu.int/11.1002/1000/7435)
2.8	ITU-T H.222.0 (2000) Amd. 4	2005-01-08	16	[11.1002/1000/7436](http://handle.itu.int/11.1002/1000/7436)
2.9	ITU-T H.222.0 (2000) Amd. 5	2005-01-08	16	[11.1002/1000/7437](http://handle.itu.int/11.1002/1000/7437)
2.10	ITU-T H.222.0 (2000) Technical Cor. 4	2005-09-13	16	[11.1002/1000/8560](http://handle.itu.int/11.1002/1000/8560)
3.0	ITU-T H.222.0	2006-05-29	16	[11.1002/1000/8802](http://handle.itu.int/11.1002/1000/8802)

---

* To access the Recommendation, type the URL http://handle.itu.int/ in the address field of your web browser, followed by the Recommendation's unique ID. For example, http://handle.itu.int/11.1002/1000/11830-en.

3.1	ITU-T H.222.0 (2006) Amd. 1	2007-01-13	16	11.1002/1000/9024
3.2	ITU-T H.222.0 (2006) Amd. 2	2007-08-29	16	11.1002/1000/9214
3.3	ITU-T H.222.0 (2006) Cor. 1	2008-06-13	16	11.1002/1000/9471
3.4	ITU-T H.222.0 (2006) Cor. 2	2009-03-16	16	11.1002/1000/9692
3.5	ITU-T H.222.0 (2006) Amd. 3	2009-03-16	16	11.1002/1000/9691
3.6	ITU-T H.222.0 (2006) Cor. 3	2009-12-14	16	11.1002/1000/10621
3.7	ITU-T H.222.0 (2006) Cor. 4	2009-12-14	16	11.1002/1000/10622
3.8	ITU-T H.222.0 (2006) Amd. 4	2009-12-14	16	11.1002/1000/10623
3.9	ITU-T H.222.0 (2006) Amd. 5	2011-05-14	16	11.1002/1000/11287
3.10	ITU-T H.222.0 (2006) Amd. 6	2011-05-14	16	11.1002/1000/11288
4.0	ITU-T H.222.0	2012-06-29	16	11.1002/1000/11655
4.1	ITU-T H.222.0 (2012) Amd. 1	2014-01-13	16	11.1002/1000/12054
4.2	ITU-T H.222.0 (2012) Amd. 2	2014-01-13	16	11.1002/1000/12055
4.3	ITU-T H.222.0 (2012) Amd. 3	2014-01-13	16	11.1002/1000/12056
4.4	ITU-T H.222.0 (2012) Amd. 4	2014-01-13	16	11.1002/1000/12057
4.5	ITU-T H.222.0 (2012) Amd. 5	2014-10-14	16	11.1002/1000/12306
5.0	ITU-T H.222.0	2014-10-14	16	11.1002/1000/12359
5.1	ITU-T H.222.0 (2014) Amd. 1	2015-04-29	16	11.1002/1000/12452
5.2	ITU-T H.222.0 (2014) Amd. 1 Cor. 1	2015-11-29	16	11.1002/1000/12625
5.3	ITU-T H.222.0 (2014) Amd. 2	2015-12-14	16	11.1002/1000/12632
5.4	ITU-T H.222.0 (2014) Amd. 3	2015-12-14	16	11.1002/1000/12633
5.5	ITU-T H.222.0 (2014) Amd. 1 Cor. 2	2016-07-14	16	11.1002/1000/12899
5.5	ITU-T H.222.0 (2014) Cor. 1	2016-07-14	16	11.1002/1000/12903
5.7	ITU-T H.222.0 (2014) Amd. 4	2016-07-14	16	11.1002/1000/12900
5.8	ITU-T H.222.0 (2014) Amd. 5	2016-07-14	16	11.1002/1000/12901
5.9	ITU-T H.222.0 (2014) Amd. 6	2016-07-14	16	11.1002/1000/12902
5.10	ITU-T H.222.0 (2014) Amd. 3 Cor. 1	2017-03-01	16	11.1002/1000/13184
5.10	ITU-T H.222.0 (2014) Cor. 2	2017-03-01	16	11.1002/1000/13188
5.12	ITU-T H.222.0 (2014) Amd. 7	2017-03-01	16	11.1002/1000/13186
5.13	ITU-T H.222.0 (2014) Amd. 8	2017-03-01	16	11.1002/1000/13187
6.0	ITU-T H.222.0	2017-03-01	16	11.1002/1000/13269
6.1	ITU-T H.222.0 (2017) Amd. 1	2017-12-14	16	11.1002/1000/13431
7.0	ITU-T H.222.0	2018-08-29	16	11.1002/1000/13664
7.1	ITU-T H.222.0 (2018) Cor. 1	2019-11-29	16	11.1002/1000/14106
7.2	ITU-T H.222.0 (2018) Amd. 1	2019-11-29	16	11.1002/1000/14105
8.0	ITU-T H.222.0 (V8)	2021-06-13	16	11.1002/1000/14658
8.1	ITU-T H.222.0 (2021) Cor. 1	2022-12-14	16	11.1002/1000/15172
8.2	ITU-T H.222.0 (2021) Amd. 1	2022-12-14	16	11.1002/1000/15171

CONTENTS

INTERNATIONAL STANDARD ISO/IEC 13818-1
RECOMMENDATION ITU-T H.222.0

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

Amendment 1

Carriage of LCEVC and other improvements

Scope

This amendment extends 13818-1 by defining:

- how LCEVC (ISO/IEC 23094-2) is carried over MPEG-2 systems (ITU-T H.222.0 | ISO/IEC 13818-1);

- an additional descriptor signalling the kind of media service and its usage.

It does this in a compatible way with existing support for other codecs.

It includes clarifications for the specification of carriage of JPEG XS.

1)   Clause 1.2.3

In 1.2.3, Additional references, add the following references:

–   ISO/IEC 23094-2:2021, *Information technology – General video coding – Part 2: Low complexity enhancement video coding.*

–   IETF BCP 47: IETF RFC 4647 (2006) *Matching of Language Tags* combined with IETF RFC 5646 (2009), *Tags for Identifying Languages.*

–   ANSI/SCTE 35 (2019), *Digital Program Insertion Cueing Message for Cable.*

2)   Clauses 2.1.90*bis* to 2.1.90*ter*

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

2.1.90*bi*s   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.

2.1.90*ter*   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 sixth 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 eighth 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.266	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.266	ISO/IEC 23090-3, ISO/IEC 23094-1 or ISO/IEC 23094-2 video stream number xxxx

*Further, in the **PTS (presentation time stamp)** section, 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 its accuracy, 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 **DTS (decoding time stamp)** section, insert the following new paragraph after the paragraph starting with "For EVC video streams ...":*

For LCEVC video streams, the 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. On the other hand, the base video stream that the LCEVC video stream enhances has no such restriction, which means that in this case the decoding order and the 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, right 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)   Clauses 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 provides a description of the 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 an 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 for profile_idc, level_idc, sublevel_idc in clause 7.3.4 (SC, sequence_configuration) of ISO/IEC 23094-2.

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

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

HDR_WCG_idc – This 2-bit field indicates the presence or absence of high dynamic range (HDR) and/or wide colour 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 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 messages defined for LCEVC because the ones from the base encoding are used. For this reason, LCEVC timing and HRD descriptors are not needed.

2.6.139   LCEVC linkage descriptor

Table 2-141ter provides a description of the 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 at an elementary level, then the elementary stream media_service_kind description shall take precedence.

Table 2-141quater provides a description of the media service kind descriptor.

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 the values "0-unknown", "1-Video", "2-Audio" or "3-Text", as indicated in 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. See Table 2-141sexies.

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 1 – This parameter can originate from a list of SCTE UPID identifiers list. Alternatively, identifiers can be reserved outside of the UPID identifiers 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
0x03	supplementary
0x04	emergency
0x05	description
0x06	enhanced-audio-intelligibility
0x07	dub
0x08	primary commentary
0x09	primary
0x0A	native
0x0B	Music and effects
0x0C	dialogue
0x0D	voice-over
0x0E	sign
0x0F	multi-view
0x10	karaoke

Table 2-141decies – Media service type values

Value	Description
0x11	caption
0x12	subtitle
0x13	forced-subtitle
0x14	metadata
0x15	non-primary
0x16	substitution
0x17	alternate commentary
0x18	stadium sound
0x19 .. 0xEF	Reserved
0xF0 .. 0xFF	User private

media_service_type value semantics

[General]

undefined: This is a default value that indicates that the purpose of the stream is unknown. It is different from the language code "und", which would indicate that the language of the audio is unknown. If "main" is not assigned in the descriptor for the stream, then this defaults to "main".

main: Main media component(s) that is/are intended for presentation if no other information is provided.

alternate: Media content component(s) that is/are an alternative to (a) main media content component(s) of the same media component type.

supplementary: Media content component that is supplementary to a media content component of a different media component type.

emergency: Experience that provides information about a current emergency. It is intended to enable the protection of life, health, safety and property, and may also include critical details regarding the emergency and how to respond to the emergency.

substitution: Indicates that the media content component is continuous but may have a substitution in language or media_service_type that defaults to the main type for that media content component when not available. This often occurs in linear channels where Spanish or audio description may not continuously exist for all programs on the channel, but the linear channel track does need to be continuous.

[Video specific]

sign: Visual media component representing a sign-language interpretation of an audio component.

multi-view: This value indicates that the video stream is an alternate view of the main video feed. This can be used for different sports angle feed in a football game.

caption: Textual representation of the audio track usually including both dialogue and all sounds. This text is often embedded in the video track intended for an audience with loss or partial loss of hearing and is in the same language as the dialogue.

forced-subtitle: Textual information meant for display when no other text representation is selected. It is used to clarify dialogue, alternative languages, texted graphics or location/person IDs that are not otherwise covered in the dubbed/localized audio.

[Audio specific]

description: This value indicates that the referenced program element is prepared for the visually impaired viewer. Alternate labels for this are "audio descriptive services" (ADS) or "descriptive video services" (DVS) or "visual impaired commentary".

enhanced-audio-intelligibility: Experience containing an element for improved intelligibility of the dialogue. An alternative label for this is "hearing impaired".

dub: Experience that contains an element that is presented in a language different from the original (e.g., dubbed audio, translated captions).

primary commentary: This value indicates the referenced program element is prepared with narration in addition to the natural sound of the content. This narration can be an announcer to a sporting event or the director's comments on a film or program, and is the default commentary for the content.

alternate commentary: This value indicates the referenced program element is prepared as bonus or alternate narration in addition to the natural sound of the content. This additional commentary is of the same language as the primary commentary and can be used for cases such as an additional away announcer to a sporting event or producer's comments of the film or program.

primary: This value indicates the default recommended language of the channel or program. This is a separate value from the language code itself. If the language code is Welsh with an undefined media_service_type, it means "this is Welsh". When the language code is welsh with media_service_type 0x09, it means "this is Welsh, and it is the primary language of this program."

non-primary: This value indicates a secondary language of the channel or program. This is a separate value from the language code itself, and indicates a secondary priority of language in the channel or program. Secondary languages may have substitutions in its track when that language is not available in order to provide a continuous track for the channel.

substitution: This value indicates the language in the channel or program may be replaced by another language if the language is sometimes unavailable, but the track needs to be continuous with the program. An example of this is a Spanish SAP (secondary audio program) where the commercial does not have a Spanish language track so an English language track is substituted until the advertisement is done.

native: This value indicates the language in which the program was produced. An absence of this value indicates that it is either undefined or a dubbed language.

Music and effects: This value indicates that the referenced program element has no language. This can also be known as "clean effects".

dialogue: An isolated language track intended to be combined with a music and effects track.

stadium sound: Audio track for an event (e.g., sporting event) devoid of any additional commentary. It would be the natural sound one would hear if physically attending the event at the stadium. An alternative term for this would be natural sound.

voice-over: An isolated narration track intended to be combined with audio that reflects the natural sound of the scene content.

[Text/Data specific]

dub: Experience that contains an element that is presented in a language different from the original (e.g., dubbed audio, translated captions).

commentary: Experience that contains a commentary (e.g., director's commentary, sports announcer).

primary: This value indicates the default recommended textual language of the channel or program. This is a separate value from the language code itself. If the language code is Welsh with an undefined media_service_type, it means "this is Welsh subtitles". When the language code is Welsh with media_service_type 0x09, it means "this is Welsh, and it is the primary subtitles language of this program".

non-primary: This value indicates a secondary textual language of the channel or program. This is a separate value from the language code itself and indicates a secondary priority of textual language in the channel or program. Secondary textual languages may have substitutions in its track when that textual language is not available in order to provide a continuous track for the channel.

native: This value indicates the language in which the program was produced. An absence of this value indicates that it is either an undefined or a dubbed language.

karaoke: Textual representation of a song's lyrics, usually in the same language as the associated songs.

subtitle: Textual representation of the audio track. It usually represents only the dialogue and often in a language other than the audio track dialogue. This is intended for a foreign language audience.

metadata: Media component containing information intended to be processed by application specific elements.

> NOTE 2 – Similar corresponding description values can be found in the definition of the ISO_639_language descriptor and in other standards/specifications ISO/IEC 23009-1 (MPEG-DASH), SMPTE ST 2067-2 (IMF Core Constraints), or DASH-IF interoperability points.

Figure 2-3bis and Figure 2-3ter show examples of the Media Service Kind Descriptor.

```
// Program – Media Service Kind Descriptor
i=0 // Identify Program
    media_description_flag →self
    identifier_flag→ true
    lang_pairs → 0
    media_type_idc → content
    ID_length_code→ 4 (ID_len is 12)
    ID_type → 0x20A (EIDR-Content)
    media_ID_field → 10.5240/0E4F-892E-442F-6BD4-15B0-1
i=1 // Identify Video
    media_description_flag → self
    identifier_flag→ false
    lang_pairs →2
    media_type_idc → video
    j=0
        configuration_type → complete
        lang_purpose_cnt→0
        lang_len_idc→2 (lang_len = 3)
        BCP_47_language_code. →zxx
        default // media_service_type → und → main
    j=1 // identify Closed Captioning English
        configuration_type→complete combination
        lang_purpose_cnt→1
        lang_len_idc→2 (lang_len = 3)
        BCP_47_language_code. → eng
        k=1 // media_service_type → caption

i=2 // Identify Main Audio Track of Linear Channel
    media_description_flag → self
    identifier_flag → false
    lang_pairs → 1
    media_type_idc → audio
    j=0 // Identify as English Main-Primary-Native
        configuration_type → complete
        lang_purpose_cnt → 3
        len_idc → 2 (lang_len = 3)
        BCP_47_language_code. →eng
        k=1 // media_service_type → main
        k=2 // media_service_type → primary
        k=3 // media_service_type → native
i=3 // Identify 2nd Audio Track as Spanish
    media_description_flag → self
    identifier_flag→ false
    lang_pairs →1
    media_type_idc → audio
    j=0 // Identify as Spanish Main
        configuration_type → complete
        lang_purpose_cnt→0
        len_idc→2 (lang_len = 3)
        BCP_47_language_code → spa
        default // media_service_type → und → main
```

Figure 2-3bis – Media Service Kind Descriptor semantics at program level

```
// An Audio elementary stream demuxed that has an association to a video track or content
i=0 // Identify Audio Elementary Stream
media_description_flag → self
identifier_flag → false
lang_pairs → 1
media_type_idc → audio
j=0 // Identify as English Main-Native
    configuration_type → complete
    lang_purpose_cnt → 2
    len_idc→2 (lang_len = 3)
    BCP_47_language_code → eng
    k=1//media_service_type → main
    k=2//media_service_type → native
i=1 // Associate to Video via identifier
    media_description_flag → associate
identifier_flag → true
lang_pairs → 0
media_type_idc → video
ID_length_code→ 4 (ID_len is 12)
ID_type → 0x20A (EIDR-Video)
media_ID_field → 10.5239/C370-DCA5
```

Figure 2-3ter – –Media Program Kind Descriptor semantics at elementary stream level

10) **Clause 2.25**

After clause 2.24, add the following clause 2.25:

2.25 Carriage of LCEVC

2.25.1 Constraints for the transport of LCEVC

For LCEVC video streams, the following constraints additionally apply:

- An LCEVC video stream shall be an element of an Rec. ITU-T H.222.0 | ISO/IEC 13818-1 program and the stream_type for this elementary stream shall be equal to the value defined in Table 2-34.

- A Rec. ITU-T H.222.0 | ISO/IEC 13818-1 program may include more than one LCEVC video stream.

- Each LCEVC video stream in a Rec. ITU-T H.222.0 | ISO/IEC 13818-1 program shall have an LCEVC video descriptor present in the descriptor loop of the program map section for the LCEVC video stream.

- It is possible that the same video elementary stream constitutes a base stream for more than one LCEVC video stream.

- For each LCEVC video stream in a Rec. ITU-T H.222.0 | ISO/IEC 13818-1 program, a video elementary stream in the same Rec. ITU-T H.222.0 | ISO/IEC 13818-1 program shall exist with an LCEVC linkage descriptor and the array of lcevc_stream_tags of this LCEVC linkage descriptor shall include one with the same value of lcevc_stream_tag of the LCEVC video descriptor of this LCEVC video stream.

- An LCEVC video stream enhances the base video elementary stream, element of the same Rec. ITU-T H.222.0 | ISO/IEC 13818-1 program, that carries the same lcevc_stream_tag, in its LCEVC linkage descriptor, as the lcevc_stream_tag indicated in the LCEVC video descriptor.

- The NAL units, as specified in ISO/IEC 23094-2, necessary for decoding an LCEVC video stream, shall be present within the elementary stream carrying that LCEVC video stream.

2.25.2 Carriage in PES packets

ISO/IEC 23094-2 video is carried in PES packets as PES_packet_data_bytes, using one of the 16 stream_id values assigned to video, while signalling the ISO/IEC 23094-2 video stream by means of the assigned stream-type value in the PMT (see Table 2-34). The highest level that may occur in an LCEVC video stream as well as a profile and tier that the entire stream conforms to should be signalled using the LCEVC video descriptor. If an LCEVC video descriptor is associated with an LCEVC video stream, then this descriptor shall be conveyed in the descriptor loop for the respective elementary stream entry in the program map table.

This Recommendation | International Standard does not specify the presentation of ISO/IEC 23094-2 streams in the context of a program.

For PES packetization, the following data alignment constraints apply:

- Only one access unit shall be carried within each PES packet.

- PES_packet_length of value 0 is allowed and shall indicate to the decoder that it should detect PES boundaries by means of the payload_unit_start_indicator.

For synchronization and STD management, PTSs shall be encoded in the header of each PES packet that carries the ISO/IEC 23094-2 video elementary stream data. For PTS encoding, the constraints and semantics apply as defined in 2.4.3.7 and 2.7.

2.25.3 DPB buffer management

Carriage of a LCEVC video stream over Rec. ITU-T H.222.0 | ISO/IEC 13818-1 does not impact the size of buffer DPB. For decoding of a LCEVC video stream in the STD, the size of DPB is as defined in ISO/IEC 23094-2. The DPB shall be managed as specified in Annex C of ISO/IEC 23094-2 (clauses C.3 and C.5). A decoded LCEVC access unit enters the DPB instantaneously upon decoding of the LCEVC access unit, hence at the CPB removal time of the LCEVC access unit. A decoded LCEVC access unit is presented at the DPB output time. If the base video stream provides insufficient information to determine the CPB removal time and the DPB output time of LCEVC access units, then these time instants shall be determined in the STD model from PTS and DTS timestamps as follows:

- The CPB removal time of LCEVC access unit n is the instant in time indicated by DTS(n), where DTS(n) is the DTS value of base access unit n.

- The DPB output time of LCEVC access unit n is the instant in time indicated by PTS(n), where PTS(n) is the PTS value of base access unit n.

NOTE 1 – LCEVC video sequences in which the low_delay_hrd_flag in the syntax structure hrd_parameters() on the base stream is set to 1 carry sufficient information to determine the DPB output time and the CPB removal time of each LCEVC access unit.

Hence, for base access units for which STD underflow can occur, the CPB removal time and the DPB output time are defined by HRD parameters, and not by DTS and PTS timestamps.

NOTE 2 – An LCEVC video stream may carry information to determine compliance of the LCEVC video stream to the HRD, as specified in Annex C of ISO/IEC 23094-2. The presence of this information can be signalled in a transport stream using the base timing and HRD descriptor with the hrd_management_valid_flag set to '1'. Irrespective of the presence of this information, compliance of a base video stream to the T-STD ensures that HRD buffer management requirements for CPB are met when each byte in the LCEVC video stream is delivered to and removed from CPB in the HRD at exactly the same instant in time at which the byte is delivered to and removed from EB$_n$ in the T-STD.

11)   Annex W, clause W.2

In W.2, replace the last sentence of the first paragraph:

Two codestreams may comprise an access unit when the access unit is an interlaced frame.

with:

An access unit shall contain exactly one codestream for the progressive frame mode, or shall contain two codestreams for the interlaced frame mode (one codestream per interlace field).

Further, replace the second sentence of the third paragraph:

This still picture has an associated PTS and the presentation time of succeeding pictures, if any, is later than that of the still picture by at least two picture periods.

with:

This still picture has an associated PTS, and the presentation time of succeeding pictures, if any, is later than that of the still picture by at least one picture period for progressive frame mode, but at least two picture periods for interlaced frame mode.

12)   Annex W, clause W.3

In W.3, insert, in the section specifying the semantics after the first paragraph and before the specification of the brat field:

jxes_length – This field is used to indicate the length (in bytes) of the jxes_header, which is 30 as fixed value.

13)   Bibliography

Create a bibliography with the following entries:

– ISO 639-3:2007, *Codes for the representation of names of languages – Part 3: Alpha-3 code for comprehensive coverage of languages.*

– ITU-T H-series Recommendations – Supplement 19 (2021), *Usage of video signal type code points*

ISO/IEC 23091-4:2021, *Information technology — Coding-independent code points — Part 4: Usage of video signal type code points.*

– ST 2067-2:2020, *SMPTE Standard – Interoperable Master Format – Core Constraints.*

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	Tariff and accounting principles and international telecommunication/ICT economic and policy issues
Series E	Overall network operation, telephone service, service operation and human factors
Series F	Non-telephone telecommunication services
Series G	Transmission systems and media, digital systems and networks
Series H	**Audiovisual and multimedia systems**
Series I	Integrated services digital network
Series J	Cable networks and transmission of television, sound programme and other multimedia signals
Series K	Protection against interference
Series L	Environment and ICTs, climate change, e-waste, energy efficiency; construction, installation and protection of cables and other elements of outside plant
Series M	Telecommunication management, including TMN and network maintenance
Series N	Maintenance: international sound programme and television transmission circuits
Series O	Specifications of measuring equipment
Series P	Telephone transmission quality, telephone installations, local line networks
Series Q	Switching and signalling, and associated measurements and tests
Series R	Telegraph transmission
Series S	Telegraph services terminal equipment
Series T	Terminals for telematic services
Series U	Telegraph switching
Series V	Data communication over the telephone network
Series X	Data networks, open system communications and security
Series Y	Global information infrastructure, Internet protocol aspects, next-generation networks, Internet of Things and smart cities
Series Z	Languages and general software aspects for telecommunication systems