International Telecommunication Union

ITU-T

TELECOMMUNICATION
STANDARDIZATION SECTOR
OF ITU

H.222.0

Amendment 3

(03/2009)

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 3: Transport of scalable video over Rec. ITU-T H.222.0 | ISO/IEC 13818-1

Recommendation ITU-T H.222.0 (2006) – Amendment 3

ITU-T H-SERIES RECOMMENDATIONS

AUDIOVISUAL AND MULTIMEDIA SYSTEMS

CHARACTERISTICS OF VISUAL TELEPHONE SYSTEMS	H.100–H.199
INFRASTRUCTURE OF AUDIOVISUAL SERVICES	
General	H.200–H.219
Transmission multiplexing and synchronization	**H.220–H.229**
Systems aspects	H.230–H.239
Communication procedures	H.240–H.259
Coding of moving video	H.260–H.279
Related systems aspects	H.280–H.299
Systems and terminal equipment for audiovisual services	H.300–H.349
Directory services architecture for audiovisual and multimedia services	H.350–H.359
Quality of service architecture for audiovisual and multimedia services	H.360–H.369
Supplementary services for multimedia	H.450–H.499
MOBILITY AND COLLABORATION PROCEDURES	
Overview of Mobility and Collaboration, definitions, protocols and procedures	H.500–H.509
Mobility for H-Series multimedia systems and services	H.510–H.519
Mobile multimedia collaboration applications and services	H.520–H.529
Security for mobile multimedia systems and services	H.530–H.539
Security for mobile multimedia collaboration applications and services	H.540–H.549
Mobility interworking procedures	H.550–H.559
Mobile multimedia collaboration inter-working procedures	H.560–H.569
BROADBAND, TRIPLE-PLAY AND ADVANCED MULTIMEDIA SERVICES	
Broadband multimedia services over VDSL	H.610–H.619
Advanced multimedia services and applications	H.620–H.629
IPTV MULTIMEDIA SERVICES AND APPLICATIONS FOR IPTV	
General aspects	H.700–H.719
IPTV terminal devices	H.720–H.729
IPTV middleware	H.730–H.739
IPTV application event handling	H.740–H.749
IPTV metadata	H.750–H.759
IPTV multimedia application frameworks	H.760–H.769
IPTV service discovery up to consumption	H.770–H.779

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 3

Transport of scalable video over Rec. ITU-T H.222.0 | ISO/IEC 13818-1

Summary

Amendment 3 to Recommendation ITU-T H.222.0 | ISO/IEC 13818-1 specifies the transport of bit-streams conforming to one or more profiles defined in Annex G of Recommendation ITU-T H.264 | ISO/IEC 14496-10 (fourth edition) over MPEG-2 Transport Streams as defined in Recommendation ITU-T H.222.0 | ISO/IEC 13818-1. It proposes extensions to Recommendation ITU-T H.222.0 | ISO/IEC 13818-1.

A number of use cases can be supported, if different video sub-bitstreams with different values of dependency_id of the scalable bit-stream are transported as different elementary streams (ES). This allows for de-multiplexing on Transport Stream (TS) level, which is the pre-requisite for selective access to a certain set of layers, selective content protection, or unequal error protection mechanisms. MPEG-2 TS already specifies the transport of Network Abstraction Layer (NAL) units conforming to one or more profiles defined in Annex A of Recommendation ITU-T H.264 | ISO/IEC 14496-10. This amendment makes extensions for supporting NAL units and bit-streams according to Annex G of Recommendation ITU-T H.264 | ISO/IEC 14496-10.

Source

Amendment 3 to Recommendation ITU-T H.222.0 (2006) was approved on 16 March 2009 by ITU-T Study Group 16 (2009-2012) under Recommendation ITU-T A.8 procedure. An identical text is also published as ISO/IEC 13818-1, Amendment 3.

FOREWORD

The International Telecommunication Union (ITU) is the United Nations specialized agency in the field of telecommunications, information and communication technologies (ICTs). The ITU Telecommunication Standardization Sector (ITU-T) is a permanent organ of ITU. ITU-T is responsible for studying technical, operating and tariff questions and issuing Recommendations on them with a view to standardizing telecommunications on a worldwide basis.

The World Telecommunication Standardization Assembly (WTSA), which meets every four years, establishes the topics for study by the ITU-T study groups which, in turn, produce Recommendations on these topics.

The approval of ITU-T Recommendations is covered by the procedure laid down in WTSA Resolution 1.

In some areas of information technology which fall within ITU-T's purview, the necessary standards are prepared on a collaborative basis with ISO and IEC.

NOTE

In this Recommendation, the expression "Administration" is used for conciseness to indicate both a telecommunication administration and a recognized operating agency.

Compliance with this Recommendation is voluntary. However, the Recommendation may contain certain mandatory provisions (to ensure e.g. interoperability or applicability) and compliance with the Recommendation is achieved when all of these mandatory provisions are met. The words "shall" or some other obligatory language such as "must" and the negative equivalents are used to express requirements. The use of such words does not suggest that compliance with the Recommendation is required of any party.

INTELLECTUAL PROPERTY RIGHTS

ITU draws attention to the possibility that the practice or implementation of this Recommendation may involve the use of a claimed Intellectual Property Right. ITU takes no position concerning the evidence, validity or applicability of claimed Intellectual Property Rights, whether asserted by ITU members or others outside of the Recommendation development process.

As of the date of approval of this Recommendation, ITU had received notice of intellectual property, protected by patents, which may be required to implement this Recommendation. However, implementers are cautioned that this may not represent the latest information and are therefore strongly urged to consult the TSB patent database at http://www.itu.int/ITU-T/ipr/.

CONTENTS

INTERNATIONAL STANDARD
RECOMMENDATION ITU-T

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

Amendment 3

Transport of scalable video over Rec. ITU-T H.222.0 | ISO/IEC 13818-1

1) Subclause 1.2.2

Replace in subclause 1.2.2:

ITU-T Recommendation H.264 (2005), *Advanced video coding for generic audiovisual services.*

ISO/IEC 14496-10:2005, *Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding*.

by:

ITU-T Recommendation H.264 (2007), *Advanced video coding for generic audiovisual services.*

ISO/IEC 14496-10:2008, *Information technology – Coding of audio-visual objects – Part 10: Advanced video coding*.

2) Subclause 2.1 *(Definitions)*

a) Replace subclause 2.1.7:

2.1.7 AVC video stream (system): An ITU-T Rec. H.264 | ISO/IEC 14496-10 stream. An AVC video stream consists of one or more AVC video sequences.

by:

2.1.7 AVC video stream (system): An ITU-T Rec. H.264 | ISO/IEC 14496-10 stream. An AVC video stream consists of one or more AVC video sequences. An AVC video stream may also result from re-assembling video sub-bitstreams.

b) Add after subclause 2.1.76 (New subclauses 2.1.77 – 2.1.81):

2.1.77 video sub-bitstream: A video sub-bitstream is defined to be all VCL NAL units associated with the same value of dependency_id of an AVC video stream which conforms to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10 and all associated non-VCL NAL units in decoding order as defined in ITU-T Rec. H.264 | ISO/IEC 14496-10. Re-assembling video sub-bitstreams in a consecutive order of dependency_id, starting from the dependency_id equal to 0 up to any value of dependency_id, results in an AVC video stream. A video sub-bitstream shall have the AVC byte stream format as defined in Annex B of ITU-T Rec. H.264 | ISO/IEC 14496-10.

2.1.78 AVC video sub-bitstream: The video sub-bitstream that contains the base layer as defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10 and that shall additionally contain NAL units with nal_unit_type equal to 14 (prefix NAL units). The AVC video sub-bitstream contains all VCL NAL units associated with dependency_id equal to 0.

2.1.79 SVC video sub-bitstream: The video sub-bitstream that contains VCL NAL units with nal_unit_type equal to 20 with the same NAL unit header syntax element dependency_id not equal to 0.

2.1.80 SVC dependency representation: The VCL NAL units of an AVC access unit associated with the same value of dependency_id which is provided as part of the NAL unit header or the associated prefix NAL unit header, and the associated non-VCL NAL units. Re-assembling SVC dependency representations in a consecutive order of dependency_id starting from the lowest value of dependency_id present in the access unit up to any value of dependency_id present in the access unit, while reordering the non-VCL NAL units conforming to the order of NAL units within an access unit as defined in ITU-T Rec. H.264 | ISO/IEC 14496-10, results in an AVC access unit.

2.1.81 SVC slice (system): A byte_stream_nal_unit as defined in ITU-T Rec. H.264 | ISO/IEC 14496-10 with nal_unit_type equal to 20.

3) Subclause 2.4.2.8 *(T-STD extensions for carriage of ITU-T Rec. H.264 | ISO/IEC 14496-10 Video)*

Replace in subclause 2.4.2.8:

To define the decoding in the T-STD of ITU-T Rec. H.264 | ISO/IEC 14496-10 video 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 ITU-T Rec. H.264 | ISO/IEC 14496-10 video streams are defined in 2.14.3.1.

by:

To define the decoding in the T-STD of ITU-T Rec. H.264 | ISO/IEC 14496-10 video 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 AVC video streams conforming to one or more profiles defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10 are defined in 2.14.3.1 and T-STD extension and T-STD parameters for decoding of AVC video streams conforming to one ore more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10 are defined in 2.14.3.5.

4) Subclause 2.4.3.5 *(Semantic definition of fields in adaptation field)*

Replace in Discontinuity_indicator:

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.

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.
- AVC video streams conforming to one or more profiles defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10 – 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.
- Video sub-bitstreams of AVC video streams conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10 – The first byte of an SVC dependency representation is an elementary stream access point if the following conditions are met:
  - The subset sequence parameter sets and picture parameter sets referenced in this and all subsequent SVC dependency representation in the video sub-bitstream shall be provided after this access point in the byte stream and prior to their activation.
  - If this SVC video sub-bitstream access point requires the elementary stream access point of the same AVC access unit, if any, contained in the corresponding elementary stream that needs to be present in decoding order before decoding the elementary stream associated with this elementary stream access point, then the corresponding elementary stream shall also include an elementary stream access point.

    NOTE 1 – If the hierarchy descriptor is present for this SVC video sub-bitstream then the video sub-bitstream of which the hierarchy_layer_index equals the hierarchy_embedded_layer_index of this SVC sub-bitstream should have an elementary stream access point in the same access unit.
- 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) Subclauses 2.4.3.6 *(PES packet)* and 2.4.3.7 *(Semantic definition of fields in PES packet)*

a) TREF

Change Table 2-21 as indicated below:

--- Unchanged initial part of Table 2-21 ---		
---		
if (PES_extension_flag_2 == '1') {		
marker_bit	**1**	**bslbf**
PES_extension_field_length	**7**	**uimsbf**
stream_id_extension_flag	**1**	**bslbf**
if (stream_id_extension_flag == '0') {		
stream_id_extension	**7**	**uimsbf**
}		
else {		
reserved	**6**	**bslbf**
tref_extension_flag	**1**	**bslbf**
if (tref_extension_flag == '0') {		
reserved	**4**	**bslbf**
TREF[32..30]	**3**	**bslbf**
marker_bit	**1**	**bslbf**
TREF[29..15]	**15**	**bslbf**
marker_bit	**1**	**bslbf**
TREF[14..0]	**15**	**bslbf**
marker_bit	**1**	**bslbf**
}		
}		
for (i = 0; i < N3; i++) {		
reserved	**8**	**bslbf**
}		
}		

--- Unchanged trailing part of Table 2-21 ---		
---		

b) stream_id

Replace in subclause 2.4.3.7:

stream_id – In Program Streams, the stream_id specifies the type and number of the elementary stream as defined by the stream_id Table 2-22. In Transport Streams, the stream_id may be set to any valid value which correctly describes the elementary stream type as defined in Table 2-22. In Transport Streams, the elementary stream type is specified in the Program Specific Information as specified in 2.4.4.

by:

stream_id – In Program Streams, the stream_id specifies the type and number of the elementary stream as defined by the stream_id Table 2-22. In Transport Streams, the stream_id may be set to any valid value which correctly describes the elementary stream type as defined in Table 2-22. In Transport Streams, the elementary stream type is specified in the Program Specific Information as specified in 2.4.4.

For AVC video streams conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10, all video sub-bitstreams of the same AVC video stream shall have the same stream_id value.

c) PTS (presentation time stamp)

Replace in subclause 2.4.3.7:

For ITU-T Rec. H.264 | ISO/IEC 14496-10 video, if a PTS is present in the PES packet header, it shall refer to the first AVC access unit that commences in this PES packet. An AVC access unit commences in a PES packet if the first byte of the AVC access unit is present in the PES packet. To achieve consistency between the STD model and the HRD model defined in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10, for each decoded AVC 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, defined herein as $t_{o,n,dpb}(n) = t_{r,n}(n) + t_c * dpb\_output\_delay(n)$, where $t_{r,n}(n)$, t_c, and dpb_output_delay(n) are defined as in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10.

by:

For AVC video streams conforming to one or more profiles defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10, if a PTS is present in the PES packet header, it shall refer to the first AVC access unit that commences in this PES packet. An AVC access unit commences in a PES packet if the first byte of the AVC access unit is present in the PES packet. To achieve consistency between the STD model and the HRD model defined in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10, for each decoded AVC 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, defined herein as $t_{o,n,dpb}(n) = t_{r,n}(n) + t_c * dpb\_output\_delay(n)$, where $t_{r,n}(n)$, t_c, and dpb_output_delay(n) are defined as in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10.

For video sub-bitstreams of AVC video streams conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10, if a PTS is present in the PES packet header, it shall refer to the first SVC dependency representation that commences in this PES packet. An SVC dependency representation commences in a PES packet if the first byte of the SVC dependency representation is present in the PES packet. To achieve consistency between the STD model and the HRD model defined in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10, for each re-assembled and decoded AVC 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, defined herein as $t_{o,n,dpb}(n) = t_{r,n}(n) + t_c * dpb\_output\_delay(n)$, where $t_{r,n}(n)$, t_c, and dpb_output_delay(n) are defined as in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10.

d) DTS (decoding time stamp)

Replace in subclause 2.4.3.7:

For ITU-T Rec. H.264 | ISO/IEC 14496-10 video, if a DTS is present in the PES packet header, it shall refer to the first AVC access unit that commences in this PES packet. An AVC access unit commences in a PES packet if the first byte of the AVC access unit is present in the PES packet. To achieve consistency between the STD model and the HRD model defined in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10, for each AVC access unit the DTS value in the STD shall, within the accuracy of their respective clocks, indicate the same instant in time as the nominal CPB removal time $t_{r,n}(n)$ in the HRD, as defined in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10.

by:

For AVC video streams conforming to one or more profiles defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10, if a DTS is present in the PES packet header, it shall refer to the first AVC access unit that commences in this PES packet. An AVC access unit commences in a PES packet if the first byte of the AVC access unit is present in the PES packet. To achieve consistency between the STD model and the HRD model defined in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10, for each AVC access unit the DTS value in the STD shall, within the accuracy of their respective clocks, indicate the same instant in time as the nominal CPB removal time $t_{r,n}(n)$ in the HRD, as defined in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10.

For video sub-bitstreams of AVC video streams conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10, if a DTS is present in the PES packet header, it shall refer to the first SVC dependency representation that commences in this PES packet. An SVC dependency representation commences in a PES packet if the first byte of the SVC dependency representation is present in the PES packet. To achieve consistency between the STD model and the HRD model defined in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10, for each re-assembled AVC access unit the DTS value in the STD shall, within the accuracy of their respective clocks, indicate the same instant in time as the nominal CPB removal time $t_{r,n}(n)$ in the HRD, as defined in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10.

e) P-STD_buffer_size

Replace in subclause 2.4.3.7:

The size BS_n shall be larger than or equal to the size of the CPB signalled by the CpbSize[cpb_cnt_minus1] specified by the NAL hrd_parameters() in the AVC video stream. If the NAL hrd_parameters() are not present in the AVC video stream, then BS_n shall be larger than or equal to the size of the NAL CPB for the byte stream format defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10 as 1200 × MaxCPB for the applied level.

by:

For AVC video streams conforming to one or more profiles defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10, the size BS_n shall be larger than or equal to the size of the CPB signalled by the CpbSize[cpb_cnt_minus1] specified by the NAL hrd_parameters() in the AVC video stream. If the NAL hrd_parameters() are not present in the AVC video stream, then BS_n shall be larger than or equal to the size of the NAL CPB for the byte stream format defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10 as 1200 × MaxCPB for the applied level.

For video sub-bitstreams of AVC video streams conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10, the size BS_n shall be larger than or equal to the size of the CPB signalled by the CpbSize[cpb_cnt_minus1] specified by the NAL hrd_parameters() for the video sub-bitstream carried in elementary stream ES_n as defined in 2.14.3.6. If the NAL hrd_parameters() are not present in the video sub-bitstream, the size BS_n shall be larger than or equal to the size of the NAL CPB for the byte stream format defined in ITU-T Rec. H.264 | ISO/IEC 14496-10 as 1200 × MaxCPB for the applied level for the elementary stream ES_n.

f) TREF semantics

Add in subclause 2.4.3.7:

tref_extension_flag – A 1-bit flag, which when set to '0' indicates that a TREF field is present in the PES packet header. The value of '1' for this flag is reserved.

TREF (timestamp reference) – The TREF is a 33-bit number coded in three separate fields. It indicates the decoding time value, $td_n(j)$, in the system target decoder as indicated by the DTS, or in absence of the DTS, by the PTS of the PES header of the same j-th access unit in a corresponding elementary stream n.

6) Subclause 2.4.4.9 *(Semantic definition of fields in Transport Stream program map section)*

Replace Table 2-34 by:

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
0x12	ISO/IEC 14496-1 SL-packetized stream or FlexMux stream carried in PES packets

Table 2-34 – Stream type assignments

Value	Description
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 conforming to one or more profiles defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10 or AVC video sub-bitstream as defined in 2.1.78
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	Auxiliary video stream as defined in ISO/IEC 23002-3
0x1F	SVC video sub-bitstream of an AVC video stream conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10
0x20-0x7E	ITU-T Rec. H.222.0 | ISO/IEC 13818-1 Reserved
0x7F	IPMP stream
0x80-0xFF	User Private

7) Subclause 2.5.2.7 *(P-STD extensions for carriage of ITU-T Rec. H.264 | ISO/IEC 14496-10 Video)*

Replace in subclause 2.5.2.7:

For decoding of ITU-T Rec. H.264 | ISO/IEC 14496-10 video streams carried in a Program Stream in the P-STD model, see 2.14.3.2.

by:

For decoding of AVC video streams conforming to one or more profiles defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10 carried in a Program Stream in the P-STD model, see 2.14.3.2 and for decoding of AVC video streams conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10 carried in a Program Stream in the P-STD model, see 2.14.3.6

8) Subclause 2.5.3.6 *(Semantic definition of fields in system header)*

Replace in subclause 2.5.3.6, in the semantic definition of the field **system_video_lock_flag***:*

For ITU-T Rec. H.264 | ISO/IEC 14496-10 video streams, the frequency of the AVC time base is defined by the AVC parameter time_scale. If the system_video_lock_flag is set to '1' for an AVC video stream, then the frequency of the AVC time base shall be locked to the STC and shall be exactly equal to N times system_clock_frequency divided by K, with N and K integers that have a fixed value within each AVC video sequence, with K greater than or equal to N.

by:

For ITU-T Rec. H.264 | ISO/IEC 14496-10 video streams, the frequency of the AVC time base is defined by the AVC parameter time_scale. If the system_video_lock_flag is set to '1' for an AVC video stream or for a video sub-bitstream, then the frequency of the AVC time base shall be locked to the STC and shall be exactly equal to N times system_clock_frequency divided by K, with N and K integers that have a fixed value within each AVC video sequence, with K greater than or equal to N.

9) Subclause 2.5.5 (Program Stream directory)

Replace in subclause 2.5.5:

Directory entries may be required to reference IDR picture or pictures associated with a recovery point SEI message in an AVC video stream. Each such directory entry shall refer to the first byte of an AVC access unit.

by:

For AVC video streams conforming to one or more profiles defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10, directory entries may be required to reference IDR picture or pictures associated with a recovery point SEI message in an AVC video stream. Each such directory entry shall refer to the first byte of an AVC access unit.

For video sub-bitstreams of AVC video streams conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10, directory entries may be required to reference IDR picture or pictures to be re-assembled from video sub-bitstreams and associated with a recovery point SEI message present in a video sub-bitstream. Each such directory entry shall refer to the first byte of an SVC dependency representation.

10) Subclause 2.6.1 *(Semantic definition of fields in program and program element descriptors)*

Replace Table 2-45 by:

Table 2-45 – Program and program element descriptors

descriptor_tag	TS	PS	Identification
0	n/a	n/a	Reserved
1	n/a	X	Forbidden
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

Table 2-45 – Program and program element descriptors

descriptor_tag	TS	PS	Identification
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	FlexMuxTiming_descriptor
45	X	X	MPEG-4_text_descriptor
46	X	X	MPEG-4_audio_extension_descriptor
47	X	X	auxiliary_video_stream_descriptor
48	X	X	SVC extension descriptor
49-63	n/a	n/a	ITU-T Rec. H.222.0 | ISO/IEC 13818-1 Reserved
64-255	n/a	n/a	User Private

11) Subclauses 2.6.6 *(Hierarchy descriptor)* and 2.6.7 *(Semantic definition of fields in hierarchy descriptor)*

a) Replace Table 2-49 by:

Table 2-49 – Hierarchy descriptor

Syntax	No. of bits	Mnemonic
hierarchy_descriptor() {		
descriptor_tag	8	uimsbf
descriptor_length	8	uimsbf
reserved	1	bslbf
temporal_scalability_flag	1	bslbf
spatial_scalability_flag	1	bslbf
quality_scalability_flag	1	bslbf
hierarchy_type	4	uimsbf
reserved	2	bslbf
hierarchy_layer_index	6	uimsbf
tref_present_flag	1	bslbf
reserved	1	bslbf
hierarchy_embedded_layer_index	6	uimsbf
reserved	2	bslbf
hierarchy_channel	6	uimsbf
}		

b) Replace in subclause 2.6.7:

hierarchy_type – The hierarchical relation between the associated hierarchy layer and its hierarchy embedded layer is defined in Table 2-50.

hierarchy_layer_index – The hierarchy_layer_index is a 6-bit field that defines a unique index of the associated program element in a table of coding layer hierarchies. Indices shall be unique within a single program definition.

hierarchy_embedded_layer_index – The hierarchy_embedded_layer_index is a 6-bit field that defines the hierarchy table index of the program element that needs to be accessed before decoding of the elementary stream associated with this hierarchy_descriptor. This field is undefined if the hierarchy_type value is 15 (base layer).

hierarchy_channel – The hierarchy_channel is a 6-bit field that indicates the intended channel number for the associated program element in an ordered set of transmission channels. The most robust transmission channel is defined by the lowest value of this field with respect to the overall transmission hierarchy definition.

NOTE – A given hierarchy_channel may at the same time be assigned to several program elements.

by:

temporal_scalability_flag – A 1-bit flag, which when set to '0' indicates that the associated program element enhances the frame rate of the bit-stream resulting from the program element referenced by the hierarchy_embedded_layer_index. The value of '1' for this flag is reserved.

spatial_scalability_flag – A 1-bit flag, which when set to '0' indicates that the associated program element enhances the spatial resolution of the bit-stream resulting from the program element referenced by the hierarchy_embedded_layer_index. The value of '1' for this flag is reserved.

quality_scalability_flag – A 1-bit flag, which when set to '0' indicates that the associated program element enhances the SNR quality or fidelity of the bit-stream resulting from the program element referenced by the hierarchy_embedded_layer_index. The value of '1' for this flag is reserved.

hierarchy_type – The hierarchical relation between the associated hierarchy layer and its hierarchy embedded layer is defined in Table 2-50. If scalability applies in more than one dimension, this field shall be set to the value of '8' ("Combined Scalability"), and the flags temporal_scalability_flag, spatial_scalability_flag and quality_scalability_flag shall be set accordingly.

hierarchy_layer_index – The hierarchy_layer_index is a 6-bit field that defines a unique index of the associated program element in a table of coding layer hierarchies. Indices shall be unique within a single program definition. For video sub-bitstreams of AVC video streams conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10, this is the program element index, which is assigned in a way that the bitstream order will be correct if associated SVC dependency representations of the video sub-bitstreams of the same access unit are re-assembled in increasing order of hierarchy_layer_index.

tref_present_flag – A 1-bit flag, which when set to '0' indicates that the TREF field may be present in the PES packet headers in the associated elementary stream. The value of '1' for this flag is reserved.

hierarchy_embedded_layer_index – The hierarchy_embedded_layer_index is a 6-bit field that defines the hierarchy_layer_index of the program element that needs to be accessed and be present in decoding order before decoding of the elementary stream associated with this hierarchy_descriptor. This field is undefined if the hierarchy_type value is 15 (base layer).

hierarchy_channel – The hierarchy_channel is a 6-bit field that indicates the intended channel number for the associated program element in an ordered set of transmission channels. The most robust transmission channel is defined by the lowest value of this field with respect to the overall transmission hierarchy definition.

NOTE – A given hierarchy_channel may at the same time be assigned to several program elements.

c) Replace Table 2-50 by:

Table 2-50 – Hierarchy_type field values

Value	Description
0	Reserved
1	Spatial Scalability
2	SNR Scalability
3	Temporal Scalability
4	Data partitioning
5	Extension bit-stream
6	Private Stream
7	Multi-view Profile
8	Combined Scalability
9-14	Reserved
15	Base layer

12) Subclause 2.6.11 *(Semantic definition of fields in data stream alignment descriptor)*

a) Replace in subclause 2.6.11:

Table 2-54 describes the alignment type for ITU-T Rec. H.264 | ISO/IEC 14496-10 video when the data_alignment_indicator in the PES packet header has a value of '1'. In this case the first PES_packet_data_byte following the PES header shall be the first byte of an AVC access unit or the first byte of an AVC slice, as signalled by the alignment_type value.

by:

Table 2-54 describes the alignment type for ITU-T Rec. H.264 | ISO/IEC 14496-10 video when the data_alignment_indicator in the PES packet header has a value of '1'.

In this case:

- For AVC video streams conforming to one or more profiles defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10, the first PES_packet_data_byte following the PES header shall be the first byte of an AVC access unit or the first byte of an AVC slice, as signalled by the alignment_type value.
- For video sub-bitstreams of AVC video streams conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10, the first PES_packet_data_byte following the PES header shall be the first byte of an SVC dependency representation or the first byte of an SVC slice, as signalled by the alignment_type value.

b) Replace Table 2-54 by:

Table 2-54 – AVC video stream alignment values

Alignment type	Description
00	Reserved
01	AVC slice or AVC access unit
02	AVC access unit
03	SVC slice or SVC dependency representation
04	SVC dependency representation
05-FF	Reserved

13) Subclause 2.6.64 *(AVC video descriptor)*

a) Replace in subclause 2.6.64:

For ITU-T Rec. H.264 | ISO/IEC 14496-10 video streams, the AVC video descriptor provides basic information for identifying coding parameters of the associated AVC video stream, such as on profile and level parameters included in the SPS of an AVC video stream.

The AVC video descriptor also signals the presence of AVC still pictures and the presence of AVC 24-hour pictures in the AVC video stream. If this descriptor is not included in the PMT for an AVC video stream in a transport stream or in the PSM, if present, for an AVC video stream in a program stream, then such AVC video stream shall not contain AVC still pictures and shall not contain AVC 24-hour pictures. (See Table 2-89.)

by:

For AVC video streams, the AVC video descriptor provides basic information for identifying coding parameters of the associated AVC video stream, such as on profile and level parameters included in the SPS of an AVC video stream or in the subset SPS of an SVC video sub-bitstream.

For AVC video streams conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10, there may be one AVC video descriptor associated to each of the video sub-bitstreams identifying coding parameters of the associated re-assembled AVC video streams.

The AVC video descriptor also signals the presence of AVC still pictures and the presence of AVC 24-hour pictures in the AVC video stream. If this descriptor is not included in the PMT for an AVC video stream or a video sub-bitstream in a transport stream or in the PSM, if present, for an AVC video stream or a video sub-bitstream in a program stream, then such AVC video stream shall not contain AVC still pictures and shall not contain AVC 24-hour pictures. (See Table 2-89.)

b) Replace Table 2-89 by:

Table 2-89 – AVC video descriptor

Syntax	No. of bits	Mnemonic
AVC_video_descriptor() {		
descriptor_tag	**8**	**uimsbf**
descriptor_length	**8**	**uimsbf**
profile_idc	**8**	**uimsbf**
constraint_set0_flag	**1**	**bslbf**
constraint_set1_flag	**1**	**bslbf**
constraint_set2_flag	**1**	**bslbf**
constraint_set3_flag	**1**	**bslbf**
AVC_compatible_flags	**4**	**bslbf**
level_idc	**8**	**uimsbf**
AVC_still_present	**1**	**bslbf**
AVC_24_hour_picture_flag	**1**	**bslbf**
reserved	**6**	**bslbf**
}		

14) Subclause 2.6.65 *(Semantic definition of fields in AVC video descriptor)*

Replace in subclause 2.6.65:

profile_idc, constraint_set0_flag, **constraint_set1_flag**, **constraint_set2_flag, AVC_compatible_flags and level_idc** – These fields, with the exception of AVC_compatible_flags shall be coded according to the semantics for these fields defined in ITU-T Rec. H.264 | ISO/IEC 14496-10. The semantics of AVC_compatible_flags are exactly equal to the semantics of the field(s) defined for the 5 bits between the constraint_set2 flag and the level_idc field in the Sequence Parameter Set, as defined in ITU-T Rec. H.264 | ISO/IEC 14496-10. The entire AVC video stream to which the AVC descriptor is associated shall conform to the profile, level and constraints signalled by these fields.

by:

profile_idc, constraint_set0_flag, constraint_set1_flag, constraint_set2_flag, constraint_set3_flag, AVC_compatible_flags and level_idc – These fields, with the exception of AVC_compatible_flags shall be coded according to the semantics for these fields defined in ITU-T Rec. H.264 | ISO/IEC 14496-10. The semantics of AVC_compatible_flags are exactly equal to the semantics of the field(s) defined for the 4 bits between the constraint_set3 flag and the level_idc field in the Sequence Parameter Set, as defined in ITU-T Rec. H.264 | ISO/IEC 14496-10. The entire AVC video stream to which the AVC descriptor is associated shall conform to the profile, level and constraints signaled by these fields.

15) Subclause 2.6.66 *(AVC timing and HRD descriptor)*

Replace in subclause 2.6.66:

The AVC timing and HRD descriptor provides timing and HRD parameters of the associated AVC video stream. For each AVC video stream carried in an ITU-T Rec. H.222.0 | ISO/IEC 13818-1 stream, the AVC timing and HRD descriptor shall be included in the PMT or in the PSM, if PSM is present in the program stream, unless the AVC video stream carries VUI parameters with the timing_info_present_flag set to '1':

- for each IDR picture; and
- for each picture that is associated with a recovery point SEI message.

Absence of the AVC timing and HRD descriptor in the PMT for an AVC video stream signals usage of the leak method in the T-STD is defined in 2.14.3.1 for the transfer from MB_n to EB_n but such usage can also be signalled by the hrd_management_valid_flag set to '0' in the AVC timing and HRD descriptor. If the transfer rate into buffer EB_n can be determined from HRD parameters contained in an AVC video stream, and if this transfer rate is used in the T-STD for the transfer between MB_n to EB_n, then the AVC timing and HRD descriptor with the hrd_management_valid_flag set to '1' shall be included in the PMT for that AVC video stream. (See Table 2-90.)

by:

The AVC timing and HRD descriptor provides timing and HRD parameters of the associated AVC video stream. For each AVC video stream and for each video sub-bitstream carried in an ITU-T Rec. H.222.0 | ISO/IEC 13818-1 stream, the AVC timing and HRD descriptor shall be included in the PMT or in the PSM, if PSM is present in the program stream, unless the AVC video stream or the video sub-bitstream carries VUI parameters with the timing_info_present_flag set to '1':

- for each IDR picture or re-assembled IDR picture; and
- for each picture or re-assembled picture that is associated with a recovery point SEI message.

Absence of the AVC timing and HRD descriptor in the PMT for an AVC video stream or a re-assembled AVC video stream signals usage of the leak method in the T-STD for the transfer from MB_n to EB_n as defined:

- in 2.14.3.1 for an AVC video stream conforming to one or more profiles defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10;
- in 2.14.3.5 for video sub-bitstreams of an AVC video stream conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10.

But such usage can also be signalled by the hrd_management_valid_flag set to '0' in the AVC timing and HRD descriptor. If the transfer rate into buffer EB_n can be determined from HRD parameters contained in an AVC video stream or an AVC video stream re-assembled from video sub-bitstreams, and if this transfer rate is used in the T-STD for the transfer between MB_n to EB_n, then the AVC timing and HRD descriptor with the hrd_management_valid_flag set to '1' shall be included in the PMT for that AVC video stream or for the re-assembled AVC video stream. (See Table 2-90.)

16) Subclause 2.6.67 *(Semantic definition of fields in AVC timing and HRD descriptor)*

Replace in subclause 2.6.67:

hrd_management_valid_flag – This 1-bit field is only defined for use in transport streams.

When the AVC timing and HRD descriptor is associated to an AVC video stream carried in a transport stream, then the following applies. If the hrd_management_valid_flag is set to '1', then Buffering Period SEI and Picture Timing SEI messages, as defined in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10, shall be present in the associated AVC video stream. These Buffering Period SEI messages shall carry coded initial_cpb_removal_delay and initial_cpb_removal_delay_offset values for the NAL HRD. If the hrd_management_valid_flag is set to '1', then the transfer of each byte from MB_n to EB_n in the T-STD shall be according to the delivery schedule for that byte into the CPB in the NAL HRD, as determined from the coded initial_cpb_removal_delay and initial_cpb_removal_delay_offset values for SchedSelIdx = cpb_cnt_minus1. When the hrd_management_valid_flag is set to '0', the leak method as defined in 2.14.3.1 shall be used for the transfer from MB_n to EB_n in the T-STD.

When the AVC timing and HRD descriptor is associated to an AVC video stream carried in a program stream, then the meaning of the hrd_management_valid_flag is not defined.

by:

hrd_management_valid_flag – This 1-bit field is only defined for use in transport streams.

When the AVC timing and HRD descriptor is associated to an AVC video stream or a re-assembled AVC video stream carried in a transport stream, then the following applies. If the hrd_management_valid_flag is set to '1', then Buffering Period SEI and Picture Timing SEI messages, as defined in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10, shall be present in the associated AVC video stream or re-assembled AVC video stream. These Buffering Period SEI messages shall carry coded initial_cpb_removal_delay and initial_cpb_removal_delay_offset values for the NAL HRD. If the hrd_management_valid_flag is set to '1', then the transfer of each byte from MB_n to EB_n in the T-STD shall be according to the delivery schedule for that byte into the CPB in the NAL HRD, as determined from the coded initial_cpb_removal_delay and initial_cpb_removal_delay_offset values for SchedSelIdx = cpb_cnt_minus1. When the hrd_management_valid_flag is set to '0', the leak method for the transfer from MB_n to EB_n in the T-STD shall be used:

- as defined in 2.14.3.1 for AVC video streams conforming to one or more profiles defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10;
- as defined in 2.14.3.5 for video sub-bitstreams of AVC video streams conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10.

When the AVC timing and HRD descriptor is associated to an AVC video stream or a re-assembled AVC video stream carried in a program stream, then the meaning of the hrd_management_valid_flag is not defined.

17) New subclauses 2.6.76 to 2.6.77 *(SVC extension descriptor)*

Add after subclause 2.6.75:

2.6.76 SVC extension descriptor

For video sub-bitstreams of AVC video streams conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10, the SVC extension descriptor provides information about the AVC video stream resulting from re-assembling (up to) the associated video sub-bitstream and provides information about scalability and re-assembly of the associated video sub-bitstream. There may be one SVC extension descriptor associated to any of the video sub-bitstreams of an AVC video stream conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10.

Table AMD3-1 – SVC extension descriptor

Syntax	No. of bits	Mnemonic
SVC_extension_descriptor() {		
descriptor_tag	**8**	**uimsbf**
descriptor_length	**8**	**uimsbf**
width	**16**	**uimsbf**
height	**16**	**uimsbf**
frame_rate	**16**	**uimsbf**
average_bitrate	**16**	**uimsbf**
maximum_bitrate	**16**	**uimsbf**
dependency_id	**3**	**bslbf**
reserved	**5**	**bslbf**
quality_id_start	**4**	**bslbf**
quality_id_end	**4**	**bslbf**
temporal_id_start	**3**	**bslbf**
temporal_id_end	**3**	**bslbf**
no_sei_nal_unit_present	**1**	**bslbf**
reserved	**1**	**bslbf**
}		

2.6.77 Semantic definition of fields in SVC extension descriptor

width – This 16-bit field indicates the maximum image width resolution, in pixels of the re-assembled AVC video stream.

height – This 16-bit field indicates the maximum image height resolution, in pixels of the re-assembled AVC video stream.

frame_rate – This 16-bit field indicates the maximum frame rate, in frames/256 seconds of the re-assembled AVC video stream.

average_bitrate – This 16-bit field indicates the average bit rate, in kbit per second, of the re-assembled AVC video stream.

maximum_bitrate – This 16-bit field indicates the maximum bit rate, in kbit per second, of the re-assembled AVC video stream.

dependency_id – This 3-bit field indicates the value of dependency_id associated to the video sub-bitstream.

quality_id_start – This 4-bit field indicates the minimum value of the quality_id of the NAL unit header syntax element of all the NAL units contained in the associated video sub-bitstream.

quality_id_end – This 4-bit field indicates the maximum value of the quality_id of the NAL unit header syntax element of all the NAL units contained in the associated video sub-bitstream.

temporal_id_start – This 3-bit field indicates the minimum value of the temporal_id of the NAL unit header syntax element of all the NAL units contained in the associated video sub-bitstream.

temporal_id_end – This 3-bit field indicates the maximum value of the temporal_id of the NAL unit header syntax element of all the NAL units contained in the associated video sub-bitstream.

no_sei_nal_unit_present – This 1-bit flag when set to '1' indicates that no SEI NAL units are present in the associated video sub-bitstream.

NOTE – In case the no_sei_nal_unit_present flag is set to '1' for all SVC video sub-bitstreams and is not set to '1' or not present for the AVC video sub-bitstream, any SEI NAL units, if present, are included in the AVC video sub-bitstream. If the SVC extension descriptor is absent for all video sub-bitstreams, SEI NAL units may be present in any SVC dependency representation of an SVC video sub-bitstream, and may require re-ordering to the order of NAL units within an access unit as defined in ITU-T Rec. H.264 | ISO/IEC 14496-10 before access unit re-assembling.

18) Subclause 2.7.6 *(Timing constraints for scalable coding)*

Add in subclause 2.7.6:

For AVC video streams conforming to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10, there are no timing constraints on SVC dependency representations or re-assembled AVC access units of video sub-bitstreams of an AVC video stream.

19) Subclause 2.14.1 *(Carriage of ITU-T Rec. H.264 | ISO/IEC 14496-10 Video)*

Add after the last sentence of subclause 2.14.1:

When an AVC video stream conforms to one or more profiles defined in Annex G of ITU-T Rec. H.264 | ISO/IEC 14496-10, the following constraints additionally apply:

- The AVC video sub-bitstream as defined in 2.1.78 shall be an element of an ITU-T Rec. H.222.0 | ISO/IEC 13818-1 program and the stream_type for this elementary stream shall be equal to 0x1B.
- For each SVC video sub-bitstream as defined in 2.1.79 that is an element of the same ITU-T Rec. H.222.0 | ISO/IEC 13818-1 program, the stream_type for this elementary stream shall be equal to 0x1F.
- All subset Sequence Parameter Sets and Picture Parameter Sets necessary for decoding an SVC video sub-bitstream shall be present within the elementary stream carrying the SVC video sub-bitstream.
- In each elementary stream with stream_type equal to 0x1F, exactly one SVC_drd_nal_unit as defined in 2.14.3.3 may precede all the NAL units of the same SVC dependency representation.

  NOTE 1 – If any SVC_drd_nal_unit is included in any SVC dependency representation then the HRD model should include this SVC_drd_nal_unit in the buffer model as additional non-VCL NAL units. The NAL unit type 24 may be used in a different way by other specifications out of scope of this Specification.
- The TREF field as defined in 2.4.3.7 may be present in the PES headers of elementary streams with stream_type equal to 0x1F. The TREF field shall be set and shall be present in the PES headers as specified in 2.14.3.5 and 2.14.3.6 respectively.

  NOTE 2 – Currently the presence of TREF is only specified for elementary streams with stream_type equal to 0x1F.
- When an ITU-T Rec. H.222.0 | ISO/IEC 13818-1 program includes more than one SVC video sub-bitstream, or more than one AVC video sub-bitstream and at least one SVC video sub-bitstream, a hierarchy descriptor as defined in 2.6.7 shall be used to indicate the dependencies of the related video sub-bitstreams.
- All NAL units of a re-assembled AVC access unit shall be passed to the decoder in the order of NAL units within an access unit as defined in ITU-T Rec. H.264 | ISO/IEC 14496-10.

  NOTE 3 – If SEI NAL units are present in any SVC dependency representation of an SVC video sub-bitstream, these NAL units may require re-ordering to the order of NAL units within an access unit as defined in ITU-T Rec. H.264 | ISO/IEC 14496-10 before access unit re-assembling.
- The profile and level limitations indicated by profile_idc and level_idc syntax elements in the AVC_video_descriptor, if present, and the Type II HRD parameters in the AVC_timing_and_HRD_descriptor, if present, for an AVC video stream resulting from re-assembling (up to) the video sub-bitstream associated to the descriptors shall include NAL units with nal_unit_type syntax element equal to 14 in the AVC video sub-bitstream and, if present in the SVC video sub-bitstream, NAL units with nal_unit_type syntax element equal to 24.

20) Subclause 2.14.3.1 *(T-STD extensions)*

Replace in subclause 2.14.3.1:

The T-STD model includes a transport buffer TB_n and a multiplex buffer MB_n prior to buffer EB_n for decoding of each ITU-T Rec. H.264 | ISO/IEC 14496-10 video elementary stream n. See Figure 2-11.

by:

The T-STD model includes a transport buffer TB_n and a multiplex buffer MB_n prior to buffer EB_n for decoding of each AVC video elementary stream n conforming to one or more profiles defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10. See Figure 2-11.

21) Subclause 2.14.3.2 *(P-STD extensions)*

Replace in subclause 2.14.3.2:

The P-STD model for the decoding of an ITU-T Rec. H.264 | ISO/IEC 14496-10 elementary stream includes a multiplex buffer B_n and a decoder D_n followed by a buffer DPB_n (see Figure 2-12). For each AVC video 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.

by:

The P-STD model for the decoding of an AVC video stream conforming to one or more profiles defined in Annex A of ITU-T Rec. H.264 | ISO/IEC 14496-10, includes a multiplex buffer B_n and a decoder D_n followed by a buffer DPB_n (see Figure 2-12). For each AVC video 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.

22) New subclauses 2.14.3.3 to 2.14.3.6

Add after subclause 2.14.3.2:

2.14.3.3 SVC dependency representation delimiter nal unit

See Table AMD3-2.

Table AMD3-2 – SVC dependency representation delimiter nal unit

Syntax	No. of bits	Mnemonic
SVC_drd_nal_unit() {		
forbidden_zero_bit	**1**	**bslbf**
nal_ref_idc	**2**	**bslbf**
nal_unit_type	**5**	**bslbf**
}		

2.14.3.4 Semantic definition of SVC dependency representation delimiter nal unit

forbidden_zero_bit – shall be equal to 0x0

nal_ref_idc – shall be equal to 0x0

nal_unit_type – shall be equal to 0x18

2.14.3.5 T-STD extensions for SVC

The T-STD model described in 2.14.3.1 is applied if the received elementary stream is a video sub-bitstream of stream_type 0x1B, i.e. only the AVC video sub-bitstream is received and decoded.

When there is a set of received video sub-bitstreams in an ITU-T Rec. H.222.0 | ISO/IEC 13818-1 program, for which dependencies may be signaled in the hierarchy_descriptor as defined in 2.6.6, and when there is at least one of the video sub-bitstreams in the set of received elementary streams having the value of stream_type equal to 0x1F, the T-STD model as described in 2.14.3.1 is extended as illustrated in Figure AMD3-1 and as specified below.

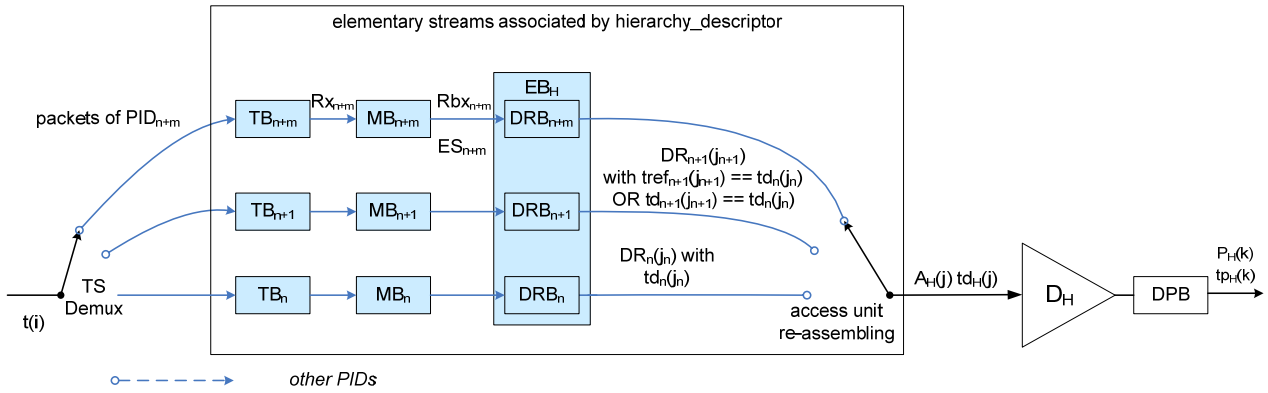

Figure AMD3-1 – T-STD model extensions for ITU-T Rec. H.264 | ISO/IEC 14496-10 Video with scalable video sub-bitstreams

The following additional notations are used to describe the T-STD extensions and are illustrated in Figure AMD3-1 above.

ES_n	is the received elementary stream associated to dependency_id value equal to n
ES_H	is the received elementary stream associated to the highest value H of dependency_id present in the set of received elementary streams
j	is an index to the re-assembled access units
j_n	is an index to the SVC dependency representations of the elementary stream associated to dependency_id value equal to n
$DR_n(j_n)$	is the j_n-th SVC dependency representation of video sub-bitstream associated to dependency_id value equal to n
$A_H(j)$	is the j-th access unit resulting from re-assembling (up to) the j_H-th SVC dependency representations with dependency_id value equal to H
$td_n(j_n)$	is the decoding time, measured in seconds, in the system target decoder of the j_n-th SVC dependency representation of the video sub-bitstream associated to dependency_id value equal to n
$td_H(j)$	is the decoding time, measured in seconds, in the system target decoder of the j-th access unit $A_H(j)$ resulting from re-assembling (up to) the j_H-th SVC dependency representations $DR_H(j_H)$
$tref_n(j_n)$	is equal to the decoding time value $td_{n-1}(j_{n-1})$ associated to the j_{n-1}-th SVC dependency representation $DR_{n-1}(j_{n-1})$, indicated by the TREF field of the PES header of the j_n-th SVC dependency representation $DR_n(j_n)$ of the same access unit
TB_n	is the transport buffer for elementary stream associated to dependency_id value equal to n
TBS_n	is the size of the transport buffer TB_n, measured in bytes
MB_n	is the multiplexing buffer for elementary stream associated to dependency_id value equal to n
MBS_n	is the size of the multiplexing buffer MB_n, measured in bytes
DRB_n	is the dependency representation buffer for elementary stream ES_n
$DRBS_n$	is the size of dependency representation buffer DRB_n, measured in bytes
EB_H	is the elementary stream buffer for all video sub-bitstreams
EBS_H	is the size of elementary stream buffer EB_H, measured in bytes
Rx_n	is specified in 2.14.3.1

Carriage in PES packets

For correct re-assembling of the SVC dependency representations to an AVC access unit, if there is both an SVC dependency representation with any dependency_id equal to n and an SVC dependency representation with dependency_id equal to (n+1) in the same AVC access unit, the following applies:

- a PES packet per SVC dependency representation start shall be present, i.e. at most one SVC dependency representation may commence in the same PES packet;

- the PTS and, if applicable, the DTS value shall be provided in the PES header of each SVC dependency representation;
- if the DTS value of the SVC dependency representation with dependency_id equal to n is different from the DTS value of the SVC dependency representation with dependency_id equal to (n+1) of the same access unit, the TREF field as defined in 2.4.3.7 shall be present in the PES header extension of the SVC dependency representation with dependency_id equal to (n+1) and the TREF field value shall be equal to the DTS value of the SVC dependency representation with dependency_id equal to n.

DPB buffer management

The DPB buffer management for the re-assembled AVC video stream shall conform to 2.14.3.1 using AVC access unit timing values, as DTS or CPB removal time, and PTS or DPB removal time, associated with the SVC dependency representations of the video sub-bitstream in elementary stream ES_H.

TB_n, MB_n, EB_n buffer management

The following applies:

- There is exactly one transport buffer TB as defined in 2.14.3.1 for each received elementary stream in the set of received video sub-bitstreams contained in elementary streams as shown in Figure AMD3-1.
- There is exactly one multiplexing buffer MB_0 for the AVC video sub-bitstream in elementary stream ES_0, where the size of the multiplexing buffer MBS_0 is constrained as follows:

$$MBS_0 = BS_{mux,0} + BS_{oh,0} + 1200 \times MaxCPB[level]_0 - cpb\_size_0$$

where $BS_{mux,0}$, $BS_{oh,0}$ are defined in 2.14.3.1 for the AVC video sub-bitstream in elementary stream ES_0;

where $MaxCPB[level]_0$ and cpb_size_0 for the elementary stream ES_0 are defined as in 2.14.3.1.

NOTE 1 – If HRD parameters are present in at least one of the video sub-bitstreams, those parameters have to be carefully handled in order to not unnecessarily increase the multiplexing buffers allocation.

- There is exactly one multiplexing buffer MB_n for each received elementary stream associated to dependency_id value not equal to 0, where the size of each multiplexing buffer MBS_n is constrained as follows:

$$MBS_n = BS_{mux,n} + BS_{oh,n}$$

where $BS_{mux,n}$, $BS_{oh,n}$ are defined in 2.14.3.1 for the AVC video stream resulting from re-assembling (up to) the SVC video sub-bitstream in elementary stream ES_n.

- There is exactly one elementary stream buffer EB_H for all the elementary streams in the set of received elementary streams as shown in Figure AMD3-1, of which the size EBS_H has the following value:

$$EBS_H = cpb\_size_H$$

where cpb_size_H is the cpb_size for the SVC video sub-bitstream in elementary stream ES_H as defined in 2.14.3.1 for the re-assembled AVC video stream.

- There is exactly one dependency representation buffer DRB_n for each elementary stream in the set of received elementary streams as shown in Figure AMD3-1, where each dependency representation buffer DRB_n in the set of received elementary streams is allocated within EB_H. Even though the size $DRBS_n$ of individual DRB_n is not constrained, the sum of the sizes $DRBS_n$ is constrained as follows:

$$EBS_H = \sum_n (DRBS_n)$$

- Transfer from MB_n to DRB_n is applied as follows:

  If the AVC_timing_and_HRD_descriptor is present with the hrd_management_valid_flag set to '1' for elementary stream ES_H, then the transfer of data from MB_n to DRB_n shall follow the HRD defined scheme for data arrival in the CPB of elementary stream ES_H as defined in Annex C of ITU-T Rec. H.264 | ISO/IEC 14496-10.

  Otherwise, the leak method shall be used to transfer data from MB_n to DRB_n as follows:

  Rate Rbx_n:

$$Rbx_n = 1200 \times MaxBR[level]_n$$

where $MaxBR[level]_n$ is defined for the byte stream format in Table A.1 (Level Limits) in ITU-T Rec. H.264 | ISO/IEC 14496-10 for the level of the AVC video stream resulting from re-assembling (up to) the associated video sub-bitstream n in elementary stream ES_n. If there is PES packet payload data in MB_n, and buffer EB_H is not full, the PES packet payload is transferred from MB_n to DRB_n at a rate equal to Rbx_n. If EB_H is full, data are not removed from MB_n. When a byte of data is transferred from MB_n to DRB_n, all PES packet header bytes that are in MB_n and precede that byte are instantaneously removed and discarded. When there is no PES packet payload data present in MB_n, no data is removed from MB_n. All data that enters MB_n leaves it. All PES packet payload data bytes enter DRB_n instantaneously upon leaving MB_n.

Access unit re-assembling and EB removal

The following specifies the access unit re-assembling that results in AVC access unit $A_H(j)$:

i) Collect all SVC dependency representations $DR_n(j_n)$ starting with the highest value of dependency_id n, equal to H, to the lowest value of dependency_id n, equal to m, present in access unit $A_H(j)$ following the rule below:

- For each two corresponding $DR_{y+1}(j_{y+1})$ and $DR_y(j_y)$ of the SVC dependency representations collected for access unit $A_H(j)$, if TREF field is present for $DR_{y+1}(j_{y+1})$, the TREF value of $DR_{y+1}(j_{y+1})$ $tref_{y+1}(j_{y+1})$ shall be equal to the DTS value $td_y(j_y)$ of $DR_y(j_y)$, otherwise the DTS value of $td_{y+1}(j_{y+1})$ of $DR_{y+1}(j_{y+1})$ shall be equal to DTS value $td_y(j_y)$ of $DR_y(j_y)$.

ii) Assemble the SVC dependency representations in consecutive order of dependency_id n starting from 'm' to 'H' for the j-th access unit $A_H(j)$. If SEI NAL units are present in any SVC dependency representation with dependency_id not equal to 0, these NAL units shall be re-ordered to the order of NAL units within an access unit as defined in ITU-T Rec. H.264 | ISO/IEC 14496-10 before access unit re-assembling.

NOTE 2 – The number of SVC dependency representations for each access unit $A_H(j)$ may vary depending on value of j. If m represents the lowest value of dependency_id for access unit AH(j), then the collected and assembled SVC dependency representations include $DR_m(j_m)$, $DR_{m+1}(j_{m+1})$, …, $DR_H(j_H)$.

The following specifies the removal of access unit $A_H(j)$ from buffer EB_H:

At time $td_H(j)$ the AVC access unit $A_H(j)$ shall be re-assembled and available for removal from buffer EB_H. The decoding time $td_H(j)$ is specified by the DTS or by the CPB removal time that is associated to the SVC dependency representations in elementary stream ES_H, as derived from information in the re-assembled AVC video stream.

STD delay

The STD delay for re-assembled AVC access units shall follow the constraints specified in 2.14.3.1.

Buffer management conditions

Transport streams shall be constructed so that the following conditions for buffer management are satisfied:

- Each TB_n shall not overflow and shall be empty at least once every second.
- Each MB_n, EB_H, and DPB shall not overflow.
- EB_H shall not underflow, except when VUI parameters are present for the AVC video sequence of the re-assembled AVC video stream with the low_delay_hrd_flag set to '1'. Underflow of EB_H occurs for AVC access unit $A_H(j)$ when one or more bytes of $A_H(j)$ are not present in EB_H at the decoding time $td_H(j)$.

2.14.3.6 P-STD extensions for SVC

The P-STD model described in 2.14.3.2 is applied if the decoded elementary stream is a video sub-bitstream of stream_type 0x1B, i.e. only the AVC video sub-bitstream is decoded.

When there is a set of decoded video sub-bitstreams in an ITU-T Rec. H.222.0 | ISO/IEC 13818-1 program, for which dependencies may be signaled in the hierarchy_descriptor as defined in 2.6.6, and when there is at least one of the video sub-bitstreams in the set of decoded elementary streams having the value of stream_type equal to 0x1F, the P-STD model as described in 2.14.3.2 is extended as illustrated in Figure AMD3-2 and as specified below.

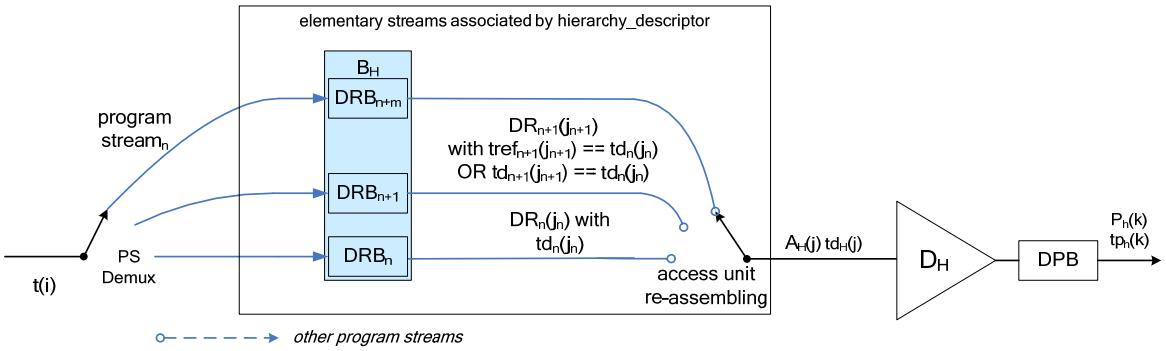

Figure AMD3-2 – P-STD model extensions for ITU-T Rec. H.264 | ISO/IEC 14496-10 Video with scalable video sub-bitstreams

The following additional notations are used to describe the P-STD extensions and are illustrated in Figure AMD3-2 above.

ES_n	is the decoded elementary stream associated to dependency_id value equal to n
ES_H	is the decoded elementary stream associated to the highest value H of dependency_id in the set of decoded elementary streams
j	is an index to the re-assembled access units
j_n	is an index to the SVC dependency representations of the elementary stream associated to dependency_id value equal to n
$DR_n(j_n)$	is the j_n-th SVC dependency representation of video sub-bitstream associated to dependency_id value equal to n
$A_H(j)$	is the j-th access unit resulting from re-assembling (up to) the j_H-th SVC dependency representations with dependency_id value equal to H
$td_n(j_n)$	is the decoding time, measured in seconds, in the system target decoder of the j_n-th SVC dependency representation of the video sub-bitstream associated to dependency_id value equal to n
$td_H(j)$	is the decoding time, measured in seconds, in the system target decoder of the j-th access unit $A_H(j)$ resulting from re-assembling (up to) the j_H-th SVC dependency representations $DR_H(j_H)$
$tref_n(j_n)$	is equal to the decoding time value $td_{n-1}(j_{n-1})$ associated to the j_{n-1}-th SVC dependency representation $DR_{n-1}(j_{n-1})$, indicated by the TREF field of the PES header of the j_n-th SVC dependency representation $DR_n(j_n)$ of the same access unit
B_H	is the input buffer for all decoded video sub-bitstreams
BS_H	is the size of the input buffer B_H, measured in bytes
DRB_n	is the dependency representation buffer for elementary stream ES_n
$DRBS_n$	is the size of dependency representation buffer DRB_n, measured in bytes

Carriage in PES packets

For correct re-assembling of the SVC dependency representations to an AVC access unit, if there is both an SVC dependency representation with any dependency_id equal to n and an SVC dependency representation with dependency_id equal to (n+1) in the same AVC access unit, the following applies:

- a PES packet per SVC dependency representation start shall be present, i.e. at most one SVC dependency representation may commence in the same PES packet;
- the PTS and, if applicable, the DTS value shall be provided in the PES header of each SVC dependency representation;
- if the DTS value of the SVC dependency representation with dependency_id equal to n is different from the DTS value of the SVC dependency representation with dependency_id equal to (n+1) of the same access unit, the TREF field as defined in 2.4.3.7 shall be present in the PES header extension of the SVC dependency representation with dependency_id equal to (n+1) and the TREF value shall be equal to the DTS value of the SVC dependency representation with dependency_id equal to n.

DPB buffer management

The DPB buffer management for the re-assembled AVC video stream shall conform to 2.14.3.1 using AVC access unit timing values, as DTS or CPB removal time, and PTS or DPB removal time, associated with the SVC dependency representations of the video sub-bitstream in elementary stream ES_H.

B_n buffer management

The following applies:

- There is exactly one elementary stream buffer B_H for all the elementary streams in the set of decoded elementary streams as shown in Figure AMD3-2, where the size of BS_H is defined by the P-STD_buffer_size field in the PES packet header of elementary stream ES_H.
- There is exactly one dependency representation buffer DRB_n for each elementary stream in the set of decoded elementary streams as shown in Figure AMD3-2, where each dependency representation buffer DRB_n in the set of decoded elementary streams is allocated within BS_H. Even though the size $DRBS_n$ of individual DRB_n is not constrained, the sum of the sizes $DRBS_n$ is constrained as follows:

$$BS_H = \sum_n (DRBS_n)$$

  where BS_H is the size of the input buffer for the SVC video sub-bitstream in elementary stream ES_H as defined in 2.14.3.2 for the re-assembled AVC video stream.

Access unit re-assembling and B removal

The following specifies the access unit re-assembling that results in AVC access unit $A_H(j)$:

i) Collect all SVC dependency representations $DR_n(j_n)$ starting with the highest value of dependency_id n, equal to H, to the lowest value of dependency_id n, equal to m, present in access unit $A_H(j)$ following the rule below:
   - For each two corresponding $DR_{y+1}(j_{y+1})$ and $DR_y(j_y)$ of the SVC dependency representations collected for access unit $A_H(j)$, if TREF field is present for $DR_{y+1}(j_{y+1})$, the TREF value of $DR_{y+1}(j_{y+1})$ $tref_{y+1}(j_{y+1})$ shall be equal to the DTS value $td_y(j_y)$ of $DR_y(j_y)$, otherwise the DTS value of $td_{y+1}(j_{y+1})$ of $DR_{y+1}(j_{y+1})$ shall be equal to DTS value $td_y(j_y)$ of $DR_y(j_y)$.

ii) Assemble the SVC dependency representations in consecutive order of dependency_id n starting from 'm' to 'H' for the j-th access unit $A_H(j)$. If SEI NAL units are present in any SVC dependency representation with dependency_id not equal to 0, these NAL units shall be re-ordered to the order of NAL units within an access unit as defined in ITU-T Rec. H.264 | ISO/IEC 14496-10 before access unit re-assembling.

   NOTE – The number of SVC dependency representations for each access unit $A_H(j)$ may vary depending on value of j. If m represents the lowest value of dependency_id for access unit $A_H(j)$, then the collected and assembled SVC dependency representations include $DR_m(j_m)$, $DR_{m+1}(j_{m+1})$, …, $DR_H(j_H)$.

The following specifies the removal of access unit $A_H(j)$ from buffer B_H:

At time $td_H(j_H)$ the AVC access unit $A_H(j_H)$ shall be re-assembled and available for removal from buffer B_H. The decoding time $td_H(j)$ is specified by the DTS or by the CPB removal time that is associated to the SVC dependency representations in elementary stream ES_H, as derived from information in the re-assembled AVC video stream.

STD delay

The STD delay for the re-assembled AVC access units shall follow the constraints specified in 2.14.3.2.

Buffer management conditions

Program streams shall be constructed so that the following conditions for buffer management are satisfied:

- B_H shall not overflow.
- B_H shall not underflow, except when VUI parameters are present for the AVC video sequence of the re-assembled AVC video stream with the low_delay_hrd_flag set to '1' or when trick_mode status is true. Underflow of B_H occurs for AVC access unit $A_H(j)$ when one or more bytes of $A_H(j)$ are not present in B_H at the decoding time $td_H(j)$.

SERIES OF ITU-T RECOMMENDATIONS

Series A	Organization of the work of ITU-T
Series D	General tariff principles
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	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	Terminals and subjective and objective assessment methods
Series Q	Switching and signalling
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 and next-generation networks
Series Z	Languages and general software aspects for telecommunication systems

Printed in Switzerland
Geneva, 2009