

---

---

**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**

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

*AMENDEMENT 3: Transport de vidéos extensibles sur  
Rec. UIT-T H.222.0 | ISO/CEI 13818-1*



Reference number  
ISO/IEC 13818-1:2007/Amd.3:2009(E)

**AVA Documents**  
Document Sharing Hub  
© ISO/IEC 2009

**PDF disclaimer**

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

Adobe is a trademark of Adobe Systems Incorporated.

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



**COPYRIGHT PROTECTED DOCUMENT**

© ISO/IEC 2009

All rights reserved. Unless otherwise specified, no part of this publication may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm, without permission in writing from either ISO at the address below or ISO's member body in the country of the requester.

ISO copyright office  
Case postale 56 • CH-1211 Geneva 20  
Tel. + 41 22 749 01 11  
Fax + 41 22 749 09 47  
E-mail [copyright@iso.org](mailto:copyright@iso.org)  
Web [www.iso.org](http://www.iso.org)

Published in Switzerland

## CONTENTS

|                                                                                                                                       | <i>Page</i> |
|---------------------------------------------------------------------------------------------------------------------------------------|-------------|
| 1) Subclause 1.2.2 .....                                                                                                              | 1           |
| 2) Subclause 2.1 ( <i>Definitions</i> ) .....                                                                                         | 1           |
| 3) Subclause 2.4.2.8 ( <i>T-STD extensions for carriage of ITU-T Rec. H.264   ISO/IEC 14496-10 Video</i> ) .....                      | 2           |
| 4) Subclause 2.4.3.5 ( <i>Semantic definition of fields in adaptation field</i> ) .....                                               | 2           |
| 5) Subclauses 2.4.3.6 ( <i>PES packet</i> ) and 2.4.3.7 ( <i>Semantic definition of fields in PES packet</i> ) .....                  | 3           |
| 6) Subclause 2.4.4.9 ( <i>Semantic definition of fields in Transport Stream program map section</i> ) .....                           | 5           |
| 7) Subclause 2.5.2.7 ( <i>P-STD extensions for carriage of ITU-T Rec. H.264   ISO/IEC 14496-10 Video</i> ) .....                      | 6           |
| 8) Subclause 2.5.3.6 ( <i>Semantic definition of fields in system header</i> ) .....                                                  | 6           |
| 9) Subclause 2.5.5 ( <i>Program Stream directory</i> ) .....                                                                          | 7           |
| 10) Subclause 2.6.1 ( <i>Semantic definition of fields in program and program element descriptors</i> ) .....                         | 7           |
| 11) Subclauses 2.6.6 ( <i>Hierarchy descriptor</i> ) and 2.6.7 ( <i>Semantic definition of fields in hierarchy descriptor</i> ) ..... | 8           |
| 12) Subclause 2.6.11 ( <i>Semantic definition of fields in data stream alignment descriptor</i> ) .....                               | 10          |
| 13) Subclause 2.6.64 ( <i>AVC video descriptor</i> ) .....                                                                            | 10          |
| 14) Subclause 2.6.65 ( <i>Semantic definition of fields in AVC video descriptor</i> ) .....                                           | 11          |
| 15) Subclause 2.6.66 ( <i>AVC timing and HRD descriptor</i> ) .....                                                                   | 11          |
| 16) Subclause 2.6.67 ( <i>Semantic definition of fields in AVC timing and HRD descriptor</i> ) .....                                  | 12          |
| 17) New subclauses 2.6.76 to 2.6.77 ( <i>SVC extension descriptor</i> ) .....                                                         | 13          |
| 18) Subclause 2.7.6 ( <i>Timing constraints for scalable coding</i> ) .....                                                           | 14          |
| 19) Subclause 2.14.1 ( <i>Carriage of ITU-T Rec. H.264   ISO/IEC 14496-10 Video</i> ) .....                                           | 14          |
| 20) Subclause 2.14.3.1 ( <i>T-STD extensions</i> ) .....                                                                              | 15          |
| 21) Subclause 2.14.3.2 ( <i>P-STD extensions</i> ) .....                                                                              | 15          |
| 22) New subclauses 2.14.3.3 to 2.14.3.6 .....                                                                                         | 15          |

## Foreword

ISO (the International Organization for Standardization) is a worldwide federation of national standards bodies (ISO member bodies). The work of preparing International Standards is normally carried out through ISO technical committees. Each member body interested in a subject for which a technical committee has been established has the right to be represented on that committee. International organizations, governmental and non-governmental, in liaison with ISO, also take part in the work. ISO collaborates closely with the International Electrotechnical Commission (IEC) on all matters of electrotechnical standardization.

International Standards are drafted in accordance with the rules given in the ISO/IEC Directives, Part 2.

The main task of technical committees is to prepare International Standards. Draft International Standards adopted by the technical committees are circulated to the member bodies for voting. Publication as an International Standard requires approval by at least 75 % of the member bodies casting a vote.

Attention is drawn to the possibility that some of the elements of this document may be the subject of patent rights. ISO shall not be held responsible for identifying any or all such patent rights.

Amendment 3 to ISO/IEC 13818-1:2007 was prepared by Joint Technical Committee ISO/IEC JTC 1, *Information technology*, Subcommittee SC 29, *Coding of audio, picture, multimedia and hypermedia information*, in collaboration with ITU-T. The identical text is published as ITU-T Rec. H.222.0 (2006)/Amd.3 (03/2009).

This corrected version modifies the title of the amendment to align with that of ITU-T Rec. H.222.0 (2006)/Amd.3 (03/2009).

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 or 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:

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

### 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.